

**Институт информационных и вычислительных технологий
МОН РК**



МАТЕРИАЛЫ
научной конференции
ИИВТ МОН РК
«Современные проблемы информатики и
вычислительных технологий»
1 - 4 июля 2019 года

Алматы 2019

УДК 004(063)
ББК 32.973
С56

Главный редактор:
Калимолдаев М.Н. - генеральный директор ИИВТ, академик НАН РК,
доктор физико-математических наук, профессор

Ответственные редакторы:
Юничева Н.Р. - ученый секретарь ИИВТ МОН РК, кандидат технических наук, доцент
Мамырбаев О.Ж. - заместитель генерального директора ИИВТ, доктор PhD
Калижанова А.У. - заместитель генерального директора ИИВТ, кандидат физико-математических наук, доцент

С56 Современные проблемы информатики и вычислительных технологий:
Мат. науч. конф. (1 - 4 июля 2019 г). – Алматы: ИИВТ МОН РК, 2019. – 407 с.

ISBN 978-601-332-346-6

В сборнике представлены материалы конференции ИИВТ МОН РК «Современные проблемы информатики и вычислительных технологий».

В сборнике опубликованы научные материалы сотрудников института, докторантов PhD и магистрантов, а также научных работников, принимающих участие в проектах по грантовому и программно-целевому финансированию.

Рассмотрены актуальные вопросы в области информационных и вычислительных технологий, информационной безопасности, проблем распознавания образов, классификации и теории принятия решений, оптимизации и оптимального управления динамическими системами, разработки моделей методов и информационной технологии построения интеллектуальных экспертных систем прогнозирования и управления сложными объектами, математического моделирования и управления динамическими, техническими и экономическими системами, теории моделей и спецификаций информационных систем, робототехнических систем, проблем синтеза и распознавания казахской речи.

Материалы сборника предназначены для научных работников, докторантов и магистрантов, а также студентов старших курсов.

УДК 004(063)
ББК 32.973

ISBN 978-601-332-346-6

© Институт информационных и
вычислительных технологий
МОН РК, 2019

Программный комитет

Председатель:

Калимолдаев М.Н., академик НАН РК, д.ф.-м.н., профессор, генеральный директор ИИВТ МОН РК

Сопредседатели:

Мамырбаев О.Ж., доктор PhD, зам. ген. директора ИИВТ МОН РК

Калижанова А.У., к.ф.-м.н., доцент, зам. ген. директора ИИВТ МОН РК

Ученый секретарь:

Юничева Н.Р., к.т.н., доцент, ученый секретарь ИИВТ МОН РК

Члены Программного комитета:

Амиргалиев Е.Н., д.т.н., профессор

Арсланов М.З., д.ф.-м.н., профессор

Ашимов А.А., академик НАН РК, д.т.н.

Бияшев Р.Г., д.т.н., профессор

Кудайкулов А.К., д.ф.-м.н., профессор

Мазаков Т.Ж., д.ф.-м.н., профессор

Мусабаев Р.Р., к.т.н.

Мустафин С.А., к.т.н., доцент

Нуртазин А.Т., к.т.н., доцент

Нысанбаева С.Е., д.т.н., профессор

Пак И.Т., академик РАЕН, д.т.н., профессор

Самигулина Г.А., д.т.н., профессор

Тлеубергенов М.И., д.ф.-м.н., профессор

Утепбергенов И.Т., д.т.н., профессор

Организационный комитет

Председатель:

Ахметжанов М.А., доктор PhD

Члены Организационного комитета:

Шаяхметова А.С., доктор PhD

Козбакова А.Х., доктор PhD

Ибраева В.М.

Анищенко Л.Н.

Калиева Г.С.

Капалова Н.А., к.т.н.

Кулемзин А.А.

Мажитов Ш.С.

Шахмаев Р.А.

Меркебаев А.

RELIABLE GESTURE RECOGNITION METHODS

Aitimov A., Turan C., Duisebekov Zh., Amirgaliyev Ye.

*e-mail: askhat.aitimov@sdu.edu.kz, cemil.turan@sdu.edu.kz,
zhasdauren.duisebekov@sdu.edu.kz, yedilkhan.amirgaliyev@sdu.edu.kz*

*Suleyman Demirel University, Kaskelen, Kazakhstan
Satbayev University, Almaty, Kazakhstan
Institute of Information and Computational Technologies CS MES RK,
Kazakhstan*

Abstract. *Due to the variety of feature extractions and classifiers, many different algorithms have been proposed for gesture recognition. In this paper, we work to increase the recognition performance in terms of recognition rate and execution time by using recently proposed modified sparse representation classifier based on intensity of images. Sparsity based classifier is compared with two conventional ones as K-nearest neighbor and random forest classifiers on gesture recognition. Simulation results showed that, our recognition algorithm based on sparsity has a higher performance than that of the others for both recognition rate and execution time.*

Introduction. Gestures are the most natural ways of communication and take an important role in everyday life. Hand gesture recognition has always attracted a lot of attention in academia and industry [1]. From assistive technologies to office appliances, mobile applications and autonomous cars, all kinds of technologies today are trying to integrate hand gestures to make the interaction process between human and machine much easier, faster, and more natural [2], [3]. By an efficient recognition algorithm, it is possible that smart devices such as cellphones or computers can be controlled by only human gestures instead of fingers or voice.

A typical gesture recognition system will usually include a pre-processing step for detection, and a classifier for recognition [2]. There are different methods and technologies used for pre-processing and feature extraction. While some apply various filters and algorithms, others use sensors and extra devices to get more accurate data and results [3]. Because we work with image intensities as feature vectors, in this paper, we focus only on the classification process. Specifically, three different algorithms are used for recognition, and results are compared. The aim of the work is to apply the recently proposed Modified Sparse Representation-based classification algorithm (MSRC) [4] in hand gesture recognition, and compare its results against existing common algorithms used in gesture recognition such as k-Nearest Neighbors (kNN) [3], [5] and Random Forest (RF) algorithms [6]. Images, which are static hand postures taken from a ready dataset [7], are processed, and each algorithm is applied to measure its performance in terms of time and accuracy of recognition.

This paper is organized as follows: Section II describes the specifics of the dataset used for comparison. In Section III, the © 2018 IEEE above mentioned three algorithms are briefly introduced to carry the experiments. Section IV contains the experimental results and comparisons are made. Finally, the last section concludes with a brief summary and future work.

Hand Posture Dataset. Static hand posture images from “Hand Reader Dataset” by Ghassem Tofighi [7] are used to test and compare the performance of the

algorithms. This section explains some of the important details and specifics of this open source dataset. “Hand Reader Dataset” contains 500 images of 10 different hand postures, which are 10 American Sign Language alphabet letters. The current study uses only 4 of those signs (A, B, C, D). For every sign there are 50 images to give the total of 200 images. All the images are in JPG format, and are 320x240 pixels in size. Every image is unique, and shows a different representation of a single sign in terms of scale and form of hand posture. All images have uniform black background, but may have slight differences in lighting conditions.



Fig. 1. Sample images from Hand Reader Dataset

Review of classifiers for recognition.

k-Nearest Neighbor Classifier. *k*-nearest neighbor (kNN) is a non-parametric and commonly used classification method in any type of recognition process. It has been first introduced by Cover and Hart in 1967 [8]. Effectiveness, simplicity, stability and good performance of kNN take attention of many researchers in recognition and machine learning algorithms [9]. In addition, Wu et. al mentioned in their work that kNN is one of the top 10 algorithms in data mining [10].

The number of neighbor’s *k* and the distance between test and training samples are the main components of kNN. Consider $T = \{q_{1,1}, q_{1,2}, \dots, q_{2,1}, q_{2,2}, \dots, q_{m,n}\}$ as training sample which is formed by *n* number of feature vectors for *m* number of classes. For a given test input $I = \{r_1, r_2, \dots, r_n\}$ and a number of *k* neighbors (here used 1), *k* nearest neighbor elements are calculated by Euclidean formula (other distance formulas such as Manhattan or Minkowski can be used),

$$D = \sqrt{\sum_{i=1}^n (T_i - I_i)^2} \quad (1)$$

kNN algorithm first calculates the distance from k -nearest neighbor training samples from the given test sample and determines the class of test sample as the most repeated neighbor's class.

Random Forrest Classifier

The second and also widely used decision tree based classification method is Random Forrest (RF) classifier. Breiman developed RF using Bagging and random selection of variable in 2001. This machine-learning algorithm has good performance in classification on big databases with a fast decision [11]. Currently, RF has been implemented in many spheres such as image classification, real-time pose recognition, sign language recognition with a significant performance [12].

The idea of RF is to create many random decision-trees from data and obtain sample used out-of-bag samples to test performance of model. Then, test data is compared with each decision-tree to obtain the prediction and most repeated prediction will be the result.

Sparse Representation Classifier

[13] proposed that any new test image can be classified by training images matrix and sparse reconstruction vector and called it as sparse representation based classification (SRC) algorithm. However, they used a sparse vector including an entry 1 corresponded to each image index in training matrix to recover the test image. Because that needs the same number of different reconstruction vectors as the number of the image vectors in training matrix, recognition process takes very long time. Based on this sparse representation, the authors in papers [4], [14] proposed a modified and robust algorithm which exploits the l_1 - minimization of the reconstruction vector as following:

Consider an $m \times n$ training matrix $A = [v_{1,1}, v_{1,2}, \dots, v_{i,n}]$ formed by the columns of image vectors of $v_{i,n}$ where i represents the index of image classes and n is the number of images in each class. To classify a given test sample image vector $\in \mathbb{R}^m$, solve l_1 -minimization problem for $x \in \mathbb{R}^n$ that minimizes $\|y - Ax\|_1$. Compute the sum of entries $\delta_i(x)$ which correspond to each i - th class and obtain $R_i = \text{sum}(\delta_i(x))$.

It is clear that the maximum sum gives the corresponding class because of high correlation between the test sample and training class as: $\text{identity}(y) = \arg \max_i \{R_i\}$. Because we use only one sparse reconstruction vector found by l_1 - minimization instead of many different sparse vectors, the modified algorithm takes very less time compared with the original SRC algorithm.

Figure 2 clearly explains the test image reconstruction by the product of training image matrix and sparse reconstruction vector. Each column with different colors represents a different image. The same class images have the same color and placed in the training matrix (**A**) group by group. To reconstruct the test image (**y**), the sparse reconstruction vector (**x**) takes the significant entries from the related columns of A because of high correlation between the same class images. To classify the test image, we just calculate the sum of entries group by group (here two by two) in **x**. Finally, the place of maximum sum gives us the corresponding class and thus the recognition is completed.

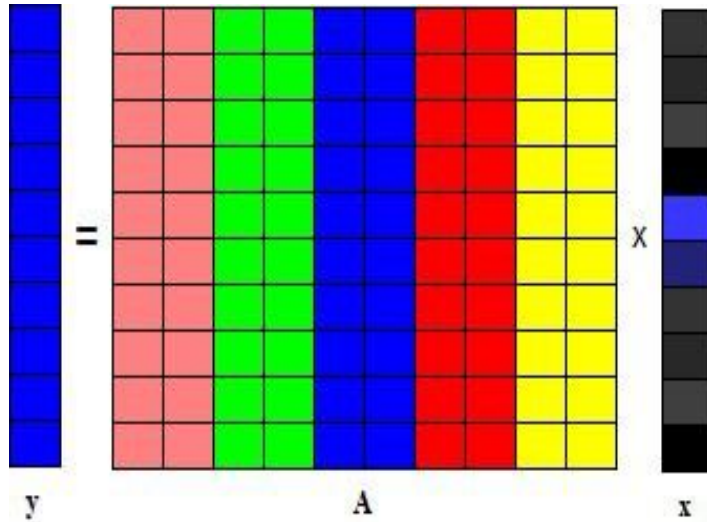


Fig. 2. Sparse reconstruction

Simulation Results.

In this section, we have done experiments with three algorithms and the sparse based classifier has been compared with kNN and RF in terms of recognition rate and execution time for recognition of different size images. Simulations were performed on Python platform. All 200 images in dataset were cropped to get rid of the redundant parts at the edges of the hands, obtaining a new rectangular frame and every cropped image was converted into grayscale image. Then, they are rescaled as six different sizes with 20x15, 24x18, 28x21, 32x24, 36x27 and 40x30 obtaining six different groups of size. After that, each group was divided into two sets to be used as training set with 120 images and testing set with 80 images. Simulations were categorized into two parts. In the first part, we compared the accuracy of recognition for all three algorithms and carried the results to a bar graph as shown in Fig. 3.

Table 1. Recognition rate (rr) and execution time for knn, rf and msrc algorithms for different image sizes

SIZE	kNN		RF		MSRC	
	RR(%)	TIME(s)	RR(%)	TIME(s)	RR(%)	TIME(s)
20x15	95	0.104	87.5	0.216	97.5	0,054
24x18	95	0.104	92.5	0.222	95.1	0,056
28x21	95	0.104	88.75	0.228	96.25	0,056
32x24	95	0.105	88.75	0.233	96.25	0,061
36x27	95	0.105	87.5	0.236	95	0,061
40x30	95	0.107	91.25	0.281	97.5	0,081

It shows that kNN is very stable algorithm giving always 95 % of recognition rate and approximately constant time of 0.1 s (in Fig. 4). Compared with kNN, RF has less RR (between 87.5 and 92.5) and needs more time (between 0.216 and 0.281). Although the time increases with increasing size of images as shown in Fig. 4, MSRC has the best performance with a recognition rates more than 95 % and less than 0.1 second of execution time for all sizes. As a limitation of MSRC, we should say that, it

is better to use for small size images and it will not be useful for big data because of very large size of training matrix. All data obtained in simulations are given in Table I.

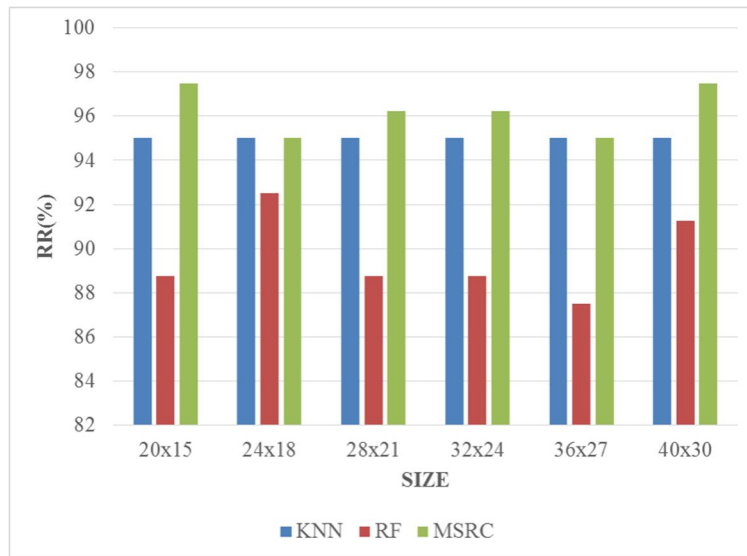


Fig. 3. Recognition rate vs image size for all algorithms of kNN, RF and MSRC

Conclusions

Gesture recognition is one of the most interesting areas in computer vision. Recent researches show that, by successful gesture recognition, many smart devices can be controlled without touching or speaking but just with human gestures. In this work, we tried to contribute into this area by increasing the recognition performance and applied the recently proposed sparsity based recognition algorithm MSRC. It has been compared with kNN and RF classifiers to test its appropriateness. “Hand Reader Dataset with different gesture representation and sizes are used” to compare the performances in terms of recognition rate and execution time. Simulation results show that the MSRC has significantly better performance than the other two classifiers in terms of recognition rate and execution time. As a limitation, it is noticeable that MSRC algorithm gets slower with increasing size. So, we recommend that it is better to use for small size images or less number of training sample. As a future work, we plan to work on a combination of MSRC and a dimensionality reduction algorithm, in order to decrease the execution time.

References

1. C. Quan and J. Liang, “A Simple and Effective Method for Hand Gesture Recognition”, IEEE International Conference on Network and Information Systems for Computers, pp. 302-305, China, 2016.
2. R. Lionnie, I.K. Timotius, and I. Setyawan, “An Analysis of Edge Detection as a Feature Extractor in a Hand Gesture Recognition System based on Nearest Neighbor”, IEEE International Conference on Electrical Engineering and Informatics, pp. 1-4, Indonesia, 2011.
3. T.P. Salunke and S.D. Bharkad, “Power Point Control Using Hand Gesture Recognition Based on Hog Feature Extraction And kNN Classification” IEEE International Conference on Computing Methodologies and Communication, pp. 1151-1155, India, 2017.

4. C. Turan, S. Kadyrov, and D. Burissova, "An Improved Face Recognition Algorithm Based on Sparse Representation", *IEEE International Conference on Computing and Network Communications*, pp. 32-35, Kazakhstan, 2018.
5. B. Gupta, P. Shukla, and A. Mittal, "K-Nearest Correlated Neighbor Classification for Indian Sign Language Gesture Recognition using Feature Fusion", *IEEE International Conference on Computer Communication and Informatics*, pp. 1-5, India, 2016.
6. M.F. Wahid, R. Tafreshi, M. Al-Sowaidi, and R. Langari, "An Efficient Approach to Recognize Hand Gestures Using Machine-Learning Algorithms", *IEEE 4th Middle East Conference on Biomedical Engineering*, pp. 171-176, Tunisia, 2018.
7. Hand-Reader-Dataset, GitHub repository. [Online]. Available: <https://github.com/tofighi/Hand-Reader-Dataset>
8. Yanhui Guo, Siming Han, Ying Li, Cuifen Zhang, Yu Bai, "K-Nearest Neighbor combined with guided filter for hyperspectral image classification", *International Conference on Identification, Information and Knowledge in the Internet of Things*, vol., pp. 159-165, China, 2018.
9. Nimagna Biswas, Saurajit Chakraborty, Sankha Subhra Mullick, Swagatam Das, "A parameter independent fuzzy weighted k-Nearest neighbor classifier", *Pattern Recognition Letters*, vol.101, pp. 80-87, India, 2018.
10. Jianping Gou, Hongxing Ma, Weihua Ou, Shaoning Zeng, Yunbo Rao, Hebiao Yang, "A parameter independent fuzzy weighted k-Nearest neighbor classifier", *Pattern Recognition Letters*, vol.101, pp. 80-87, India, 2018.
11. Robin Genuer, Jean-Michel Poggi, Christine Tuleau-Malot, Nathalie Villalaneix, "Random Forests for Big Data", *Big Data Research*, vol.9, pp. 28-46, France, 2017.
12. Ajjen Joshi, Camille Monnier, Margrit Betke, Stan Sclaroff, "Comparing random forest approaches to segmenting and classifying gestures", *Image and Vision Computing*, vol.58, pp. 86-95, USA, 2017.
13. J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 2102-27, 2009.
14. C. Turan, "Robust Face Recognition via Sparse Reconstruction Vector", *13th International Conference on Electronics, Computer and Computation (ICECCO)*, 2017.

METHODS FOR ACCURATE GESTURE RECOGNITION

Amirgaliyev Ye., Jantayev R.

e-mail: yedilkhan.amirgaliyev@sdu.edu.kz, ruslan.jantayev@sdu.edu.kz

*Institute of Information and Computational Technologies CS MES RK,
Kazakhstan*

Suleyman Demirel University, Kaskelen, Kazakhstan

Abstract. *One of the essential problems in Computer Vision is identification of important objects. While exhaustive work done image processing and its performance is still being limited by ambiguity. In current work we compared traditional machine*

learning method versus Deep Learning method on Handwritten Digit Recognition using MNIST dataset. We showed that Convolutional Neural Network(CNN) algorithm reaches higher recognition accuracy than Support Vector Machine(SVM).

Introduction. Handwritten Digit Recognition technologies are widely used due to the technological innovation advancements as well as stability motives [1]. As a crucial Handwritten Digit Recognition has brought up considerable attention – due to the fact its trustworthiness as well as functionality. Both individuals and social industries are in demand of pragmatic as well as coercion-less computerized recognition affirmation technique [2]. Consequently, Optical Character Recognition(OCR) programs are being formulated speedily and also as a result almost any OCR Digit under various different hand-written presents such as left-handwritten as well as right-handwritten are classified as well [1]. Below fig-1 is an example of different hand-writings. It's proven fact that, the position from the left-handwritten and right-handwritten figures make crucial effect on performance to the recognition rate. The pictures within the digit data sets are necessary to be pre-processed to get a much better performance of identification [3].

There are various types of HDR techniques as reported by scientists: Saxena et al, suggested a new neural network product regarding classification associated with handwritten numbers these people boosted their own approaches utilizing ensemble classification [5]. Das et al, selected neighborhood features within hand-written numbers making use of anatomical system after which categorized features along with SVM [6]. Cardoso along with Wichert recommended a new bio-chemically influenced design with regard to HDR additionally utilized a straight collection SVM [7]. A new hybrid design is suggested through Niu as well as Suen, in which including the particular synergy associated with a couple of exceptional classifiers: Convolutional Neural Network (CNN) as well as Support Vector Machine (SVM) which was the primary factor of writers intended for enhancing hand-written number recognition [8].

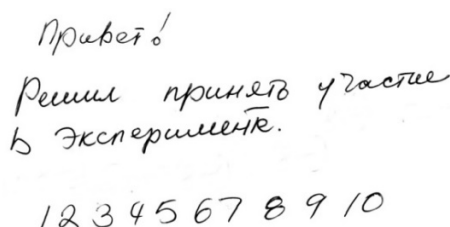


Fig. 1. Example of Handwritten Digits (left-handwritten and right-handwritten)

Mnist Database. Representative examples originated from MNIST(Modified National Institute of Standards and Technology) dataset incorporates handwritten digits overall 70,000 images containing 60,000 samples in training set and 10,000 samples in testing set, both with marked images from 10 digits including 0-9. These examples are the only little part of the broad set of NIST datasets where size was normalized to suit a 20*20-pixel box and not changing the aspect ratio. Handwritten digits are images which fits to the form of 28*28 gray scale strength of images appearing for an image along with the first column to be label (0-9) for every single image. The same has chosen for the situations of the testing set as 10,000 images with a label of 0 to 9. Below, fig-2 is the MNIST Handwritten Digits in row table.

MNIST database has four files in it: 1. Training set images as a source data 2. Training set labels from 0 to 9 3. Test set images as a sources data 4. Test set labels from 0 to 9

Review of Algorithms. Within this section, brief details are given for suggested approach and algorithms utilized in the experiments. Before you apply very good algorithms, a pre-processing step is conducted for those images within the database by Histograms of Oriented Gradients (HOG) to obtain remove noise and obtain better result. After that, SVM and CNN are tested for conjecture and in contrast to one another when it comes to recognition rates. Within the followings, the filter and algorithms are briefly described.



Fig. 2. Different samples of handwritten digits in MNIST

Histograms of Oriented Gradients.

HOG is commonly applied being a feature detailed picture location for subject detection such as Handwritten Digit Recognition or any object identification. To improve the effectiveness in the object searching, colors and gamma from the picture must be normalized. The item search will depend on the diagnosis strategy applied for the small images based on slipping detector windowpane that probes area by location of the authentic enter appearance as well as its scaled versions. [3]

The initial step in HOG recognition would be to divide the origin image into blocks (for instance 1616 pixels). Each block is split by small regions, known as cells (for instance 88 pixels). Usually blocks overlap one another, so the same cell might be in a number of blocks. For every pixel inside the cell the vertical and horizontal gradients are acquired. The easiest method to achieve that is by using 1-D Sobel vertical and horizontal operators:

$$G_x(y, x) = Y(y, x + 1) - Y(y, x - 1); \tag{1}$$

$$G_y(y, x) = Y(y + 1, x) - Y(y - 1, x)$$

$Y(y,x)$ is considered to be the pixel depth located at coordinates x as well as y . $G_x(y,x)$ is the horizontal gradient, and also $G_y(y,x)$ is the vertical gradient. The particular magnitude and also phase of the slope tend to be established as following: s

$$G_x(y, x) = \sqrt{G_x^2(y, x) + G_y^2(y, x);}$$

$$\theta(y, x) = \arctan\left(\frac{G_y(y, x)}{G_x(y, x)}\right) \tag{2}$$

Then, for each cell HOG is produced correspondingly which will cover whole image. As for Histogram, nine Q bins are created, each bins being set 20 degrees which makes 180 degrees in total to cover histogram. Fig-3 shows overview of HOG.

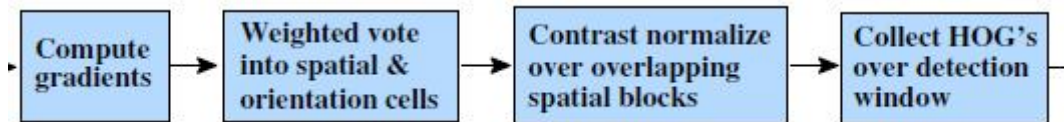


Fig. 3. An overview of our feature extraction and object detection chain

Support Vector Machine. Support vector machine is a vital model in machine learning. Some general applications which are commonly used are: face detection, text and hypertext categorization, classification of images and bioinformatics [9]. SVM sorts the data into the categories using supervised learning method. In graphical representation we can see the line which splits the data and it is known as hyperplane. SVM model implies that only support vector is important whereas other training examples are ignorable. The visual representation of functioning SVM classifier presented below:

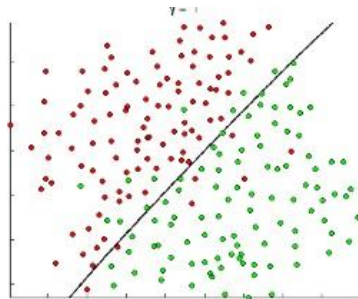


Fig. 4. An overview of Support Vector Machine Classification

For classification the handwritten digits, the SVM algorithm which we used works as follows:

First, loading the data - which are samples of handwritten digits, then we divide it into the testing and training Image and labels (train-labels). After that we divide the training data into training and testing in order to start training the classifier. We created a classifier with gamma value 0.001

```
classifier = svm.SVC(gamma=0.001)
```

We can start train the SVM classifier, providing all needed data (labels, train and testing data) from above as an input. This training is typically what takes the most computational time. Based on a learning from provided labelled data SVM classifies the unknown images. To get the accuracy of trained classifier, the digits identified with SVM matched with provided training labels.

The classification report for classifier which we used is:

```
SVC (C=1.0, cache-size=200, class-weight=None, coef=0.0, decision-function-
shape='ovr', degree=3, gamma=0.001, kernel='rbf', max-iter=-1, probability=False,
random-state=12 shrinking=True, tol=0.001, verbose=False)
```

This trained classifier is used once more on the data which defined as testing. Test image data is applied to estimate the digit's labels and is contrasted with the supplied test labels to identify the accuracy of the CNN algorithm. The Confusion Matrix which we are printed gives the percentage of accuracy on every digit.

Convolutional Neural Network

Now we are going to dive into deep learning, exactly CNN. Its a deep learning algorithm which most commonly used for image classification [8]. It is getting an image as an input, specify values as weights and biases to features into the image, so it can differentiate one from another. CNNs use a variety of multi layer perceptions which are suppose to need to have minor preprocessing. Also called as shift invariant or area invariant artificial neural networks, according to their shared- weights architecture and transformation in-variance properties. In deepnn function we created a neural network which builds a graph for classifying the images. It takes a tensor as an input with dimensions and 784 pixels each. We started with the downloading the data and reshaping it into the array of images with 4 dimensions. Now we can setup the CNN architecture and start to train our model to recognize digits from images. Our neural network consists of two hidden layers. The output layer will have 10 nodes, each node will take a value between 0 to 1 and it will be representing each number. The node with the highest value will be predicted output. The input layer has as many nodes as numbers representing each image. This is going to be about 784 because each image is represented by 28x28 pixels

First convolutional layer stands for transforming grayscale image to 32 dimensional feature vector. Second convolutional layer transforms 32 dimensional feature vector into 64. Pooling layers apply 2x2 non-overlapping combining, so the down-sampling takes place here. Then, in the fully connected layers every input will be connected to every node in the next layer, instead of that some random edges should be dropped. This is done in dropout layer, and important because overfitting must be avoided. Last convolutional layer is connected to a layer that reshapes the vector, i.e. 28x28 shaped images transforms to 7x7x64 dimensional feature map. That layer is after connected to next 1024 dimensional layer. Consequently, fully connected layers have 1024 and 10 neurons, respectively. After, a function, that returns a 2d convolution layer with full stride, downsamples a feature map by 2 times. It stimulates a weight and bias variable of a given image. Below fig-5 is the conceptual topology of CNN.

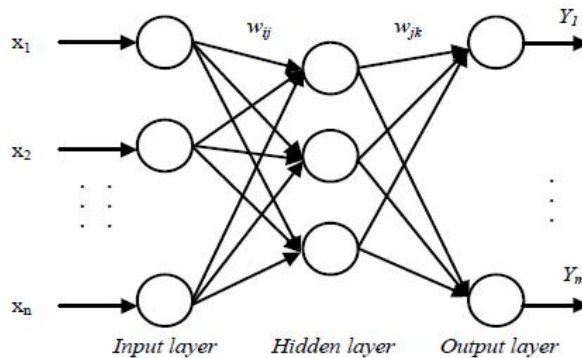


Fig. 5. The Topology of Convolutional Neural Network

Convolution Neural Network. In the study of Handwritten Digit Recognition, the CNN algorithm impressed with 100 percent recognition rate accuracy. The total number of steps were 4000, however in step 3800 it started giving 100 percent accuracy rate. Fig-6 is demonstration result of CNN result.

```
step 3400, training accuracy 0.98
step 3500, training accuracy 0.98
step 3600, training accuracy 0.98
step 3700, training accuracy 0.98
step 3800, training accuracy 1
step 3900, training accuracy 1
step 4000, training accuracy 1
```

Fig. 6. CNN result on Handwritten Digit Recognition

Overall we analyzed two Handwritten Digit Recognition algorithm, mainly SVM and CNN. SVM with HOG preprocessing step gave 94 percent accuracy. On the other hand, CNN impressed with 100 percent recognition rate within its 4000 step. In both cases MNIST database was utilized. Above figure-x and figure-xx show the detailed recognition rate. It was not pointed out about the performance times regarding recognition procedures in this work, however we should say that the particular CNN is much faster than SVM. It is currently being organized for the future work.

CNN studying has been effectively applied to a large, real-world activity. Gained outcomes appear to be at the state of the art within hand-written number identification. The particular network experienced numerous connections however fairly several cost-free variables. The actual system structures and the limitations on the weights were designed to incorporate mathematical knowledge about the task into the method. Because of its buildings, this system could be trained on a low-level manifestation of data that had minimum preprocessing (as opposed to sophisticated attribute elimination). Because of the obsolete dynamics of the data and because of the constraints enforced on the system, this mastering time was comparatively quick considering the size of the training set. Scaling attributes have been far better than one would expect simply coming from extrapolating results of back-propagation upon smaller sized, artificial issues. [4]

Conclusions

This particular paper is a picture of continuing work. Even though we expect ongoing changes in all aspects of identification technology, there are some findings that are likely to stay legitimate for some time.

The Handwritten Digit Recognition problem was main issue, in current work. We compared two major algorithms namely, SVM from Machine Learning and CNN from Deep Learning. Both experiments were done to MNIST database. SVM with HOG preprocessing step showed 94 percent of accuracy and CNN impresses with 100 percent accuracy with only 4000 steps. We successfully presented that CNN algorithm has higher accuracy rate than SVM in our current work. As a upcoming job, the time operation of calculations can be analyzed with different data source in order to get strong evaluations.

References

1. D. Hu, Research and application of handwritten numeral recognition method. China: Nanchang, 2012.
2. N.V. Boulgouris, Konstantinos N. Plataniotis and Evangelia Micheli-Tzanakou, Biometrics: Theory, Methods, and Applications. Canada: Wiley, 2010.
3. R. Ebrahimzadeh, M. Jampour, Efficient Handwritten Digit Recognition based on Histogram of Oriented Gradients and SVM. Austria: Graz, 2014.

4. Y.L. Cun, B. Boser, J.S. Denker, D. Henderson Handwritten Digit Recognition with a Back-Propagation Network. USA: New Jersey, 2003.
5. N. Saxena, Q. Abbas Kazmi, C. Pal and O.P. Vyas Employing Neocognitron Neural Network Base Ensemble Classifiers To Enhance Efficiency of Classification In Handwritten Digit Datasets. USA: D.C. 2011.
6. N. Das, R. Sarkar, S. Basu, M. Kundu, M. Nasipuri, D. Kumar A genetic algorithm based region sampling for selection of local features in handwritten digit recognition application. Appl. Soft Comput. 2012.
7. Cardoso, A. Wichert Handwritten digit recognition using biologically inspired features. Neurocomputing, 2013.
8. X. Niu, C. Y. Suen A novel hybrid CNN-SVM classifier for recognizing handwritten digits. China: Shanghai, 2015.
9. L. Botton, C. Cortes, J. S. Denker Comparison of Classifier Methods: Case Study in Handwritten Digit Recognition. USA: California, 2014.

DECOMPOSITION APPROACH FOR CLUSTERING PROBLEM

^{1,2}Mladenovic N., ²Krassovitskiy A., ²Mussabayev R.

e-mail: nenad@mi.sanu.ac.rs, akrassovitskiy@gmail.com, rmusab@gmail.com

¹Emirates College of Technology, Abu Dhabi, United Arab Emirates

²Institute of Information and Computational Technologies, Almaty, Kazakhstan

***Abstract.** Clustering of large scale datasets(BigData) is widely discussed and currently a topic of high demand. The quality of data clustering is commonly estimated by Sum of Squared Distance(SSD) criteria. In this study we show that it is possible to achieve high quality results in the terms of SSD criteria by applying iteratively the clustering procedure on subsets of the input dataset. Also, we demonstrate that the clustering time can be used as an additional parameter that allows to rig the quality of the clustering.*

Introduction

Clustering methods receive a lot of attention recently as effective tools in Machine Learning that allow discovering phenomena in the raw data. Another motivation comes from the necessity to process Big datasets in order to obtain natural grouping in the data. Hence, one of the main aspects of the clustering methods is scalability.

There is a number studies that are directed on improvement of clustering quality in cost of time complexity [7, 8]. These type of methods usually have an essential drawback: it is practically impossible to cluster medium and large datasets (about 10^5 elements and more). These methods cannot work on huge data bases because their time/space complexity grows (so, polynomially) very fast. Hence, it is reasonable to search for an algorithms having a property of a balance between scalability and clustering quality [1-4]. One of the well-known methods in data clustering is k -means algorithms which is widely used due to its simplicity and good characteristics. A number of techniques are directed on partitioning clustering methods [6] of k -means. There are algorithms which use data decomposition approach, e.g., minibatch k -means [11]. A weighted version of the k -means algorithm is applied in [3]. Meta-heuristics may help

essentially when the exact solution is hard or expensive in terms of time/space. There are several heuristics implemented in k -means that speed up computation:

- removing at each iteration patterns that are unlikely to change their membership thereafter as in [4];
- using triangle inequality in [12];
- mixture of various techniques [13].

We use a modification of k -means++ in order to build meta-heuristics that is executed on some subsets of the entire dataset. The goal of the study is to build a dataset decomposition method for centroid initialization of k -means clustering in order to produce competing results regarding MSSD (Minimum Sum of Square Distance) criteria. I.e., we produce a method for finding k -means initialization such that it is close to the optimal one, while having fast computation speed. Here we use meta-heuristics over k -means clustering by processing obtained data into secondary (high-level) clustering procedure. Another purpose of our research is to determine the efficiency of such heuristics and in different behaviours regarding meta-parameters.

2. Algorithm

By our approach, first, we decompose (split) the entire dataset shuffled randomly on subsets of fixed size. We are taking either all elements, or some portion of the dataset such that it remains representative. Next step is to do k -means clustering on some of these subsets (=windows). By estimating SSD on their corresponding clusterings we make a heuristic for a good initialization of the algorithm on the entire dataset.

(Meta-) parameters:

- k is number of required clusters
- N is number of objects in the entire data set
- d is window size (number of objects in one window). The sizes of the windows are chosen in proportion to the entire dataset. *E.g.*, taking 5 wins decomposition means taking the size of the windows equal to $N/5$.
- n is number of windows used for independent initialization of k -means during Phase 1
- $m(\geq n)$ is total number of windows used for the clustering. The union of m windows may or may not cover the entire dataset.

We considered the following two modes for the windows (wins) generation:

1. Segmentation of the entire data set on windows: A random permutation of objects in the data set is created. The data set is segmented into successive windows of size d . We refer to it as uniform window decomposition mode.

2. For each window, d random objects are selected from the entire set. By repeating this, the required number of windows is generated (objects may be picked repeatedly in different wins). We refer to it as random window generation mode.

PHASE 1: An independent application of k -means++ algorithm on a fixed number n of windows in order to find aggregated set of centroids. The scheme for the algorithms is shown in Fig. 1. For this, the centroids of all n windows are joined into one new data set. Each object is assigned the weight corresponding to the normalized SSD value for the window in which it is calculated as centroid. The weight of i -th object is calculated as follows:

$$w_i = 1 - (SSD_i - SSD_{min}) / (SSD_{max} - SSD_{min}),$$

where SSD_i is SSD value for window from which the i -th centroid is taken as object. Then, using k -means, the new dataset is divided into k -clusters, taking into account the

weights of the w_i objects. In the case of degeneration, k -means is reinitialized. The resulting centroids are used for:

- 1.1. Initialization of k -means on the entire data set
- 1.2. Evaluation of the SSD on the entire data set
- 1.3. Initialization Phase 2 of the algorithm described in the following sections.

PHASE 2 During processing subsequent windows $n + 1, n + 2, \dots, m$ we have considered two options: parallel and sequential.

2.1. Parallel option

The centroids obtained in the previous Phase 1 are used to initialize k -means on each subsequent window $m + 1, m + 2, \dots, m$. The resulting centroids and SSD are stored if there is no centroid degeneration. The stopping condition is the specified limit either on the computation time or on the number of windows being processed. Similar to Phase 1 we do the clustering on the joined set of centroids. Subsequent use of its results is similar to clauses 1.1 and 1.2 of Phase 1.

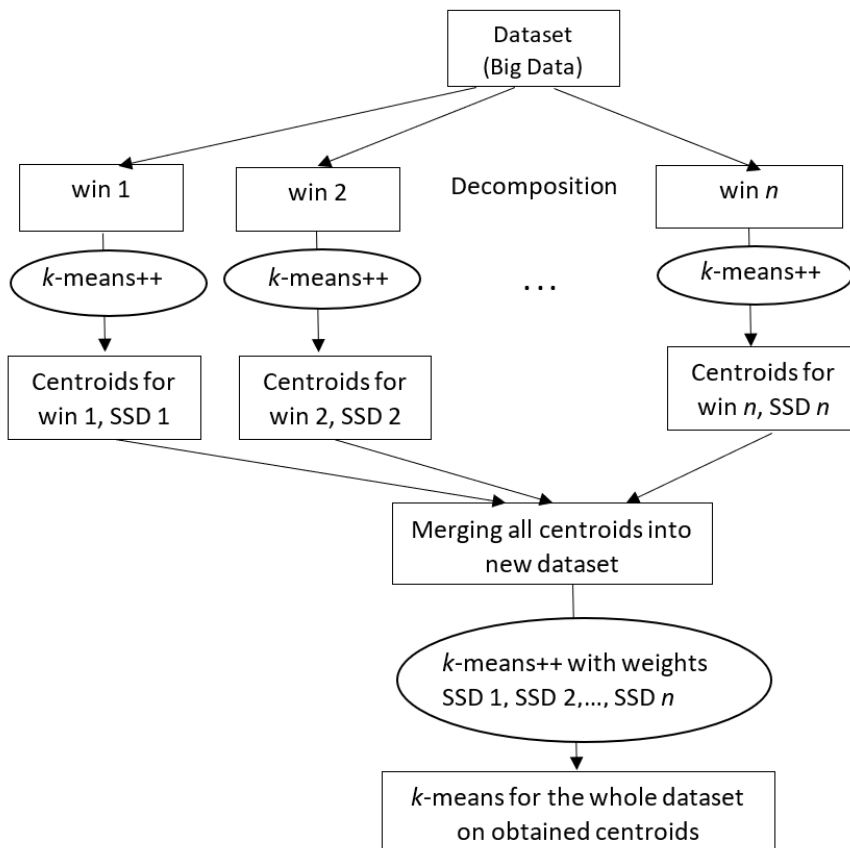


Figure 1. Scheme for parallel decomposition clustering on windows $n + 1, \dots, m$.

2.2. Sequential option

The following is the sequential version of the algorithm. It is schematically represented in Fig. 2.

Centroids from phase 1 are used for k -means clustering on the window $m + l, m + l + 1, \dots$. If there is no degeneration during clustering then memorize the resulting centroids and the corresponding SSD values. In order to obtain new centroids, we carry

out clustering with weights on the united set of centroids (similarly to Phase 1). If the time limit has not been exhausted then use centroids obtained in step 4, and repeat the procedure. Finally, when time limit is reached then the usage of obtained results on the last calculated centroids is similar to clauses 1.1 and 1.2 of Phase 1.

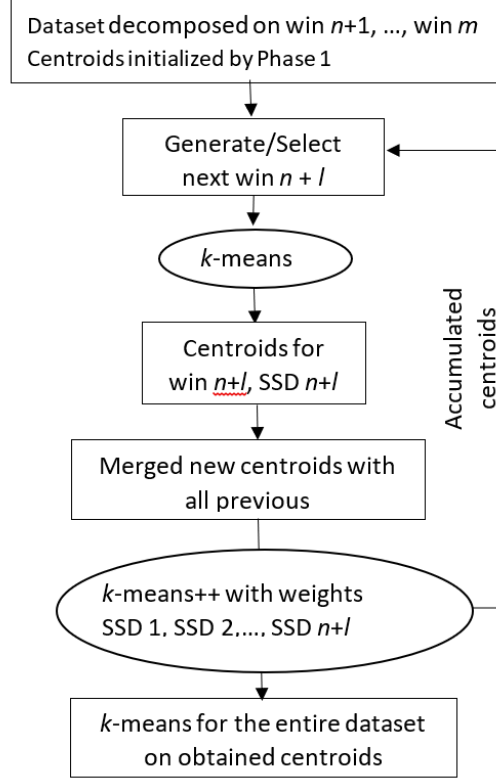


Figure 2: Sequential accumulation of the heuristically optimal SSD and centroids

3. Experiments

In Table 1 we show the testing result of our algorithm on datasets of various sizes. Each dataset is decomposed on windows of sizes $1/5 \times N$, $1/10 \times N$, ..., $1/100 \times N$.

Table 1: Experiments on different datasets. SSD criteria and computation time are considered

dataset size	num. of experiments	improves k-means++	$\frac{SSD(kmeans++)}{SSD(ours)}$	num. of overtime	$\frac{time(ours)}{time(kmeans++)}$
1×10^3	1590	33 (2.08%)	0.964	0	0.330
1×10^3	1186	35 (2.95%)	0.989	17	0.433
1×10^3	312	14 (4.49%)	0.992	15	0.051
1×10^6	1140	345 (30.26%)	1.000	10	0.386
1×10^6	300	160 (53.33%)	1.000	12	0.394
1×10^6	82	40 (48.78%)	0.996	41	0.971
$1 \times 10^{6*}$	30	26 (86.67%)	1.000	30	1.650
$1 \times 10^{6**}$	260	165 (63.46%)	1.000	0	0.160

Datasets description:

- Our first experiments are conducted on syntactic datasets from 10×10^3 elements to 2×10^6 and number of attributes from 20 to 100. Also different sizes of windows are considered. The structure of the data: some number of synthetic blobs of Gaussian distribution.

- For size 50×10^3 dataset the number of clusters was considered in the range from 5 to 70 with the step 2. The number of blobs was picket in the range from 30 to 140. The number of chunks was in the range from 5 to 100 with the step 5.

- For size 5×10^6 dataset SUSY is obtained from open UCI database.

- *Dataset is preprocessed by normalization.

- **2-means clustering problem, decomposition of dataset on windows $1/5 \times N, \dots, 1/150 \times N$.

Fig. 3-4 demonstrate the advantage of our algorithm in terms of SSD for various window sizes and k for 560 experiments on syntetic dataset of size 1×10^6 . While Fig. 5 shows the difference in terms of computation time. X-axis corresponds to combination of the number of clusters k and the window sizes l , i. e., each value has form $k \times 20 + l$.

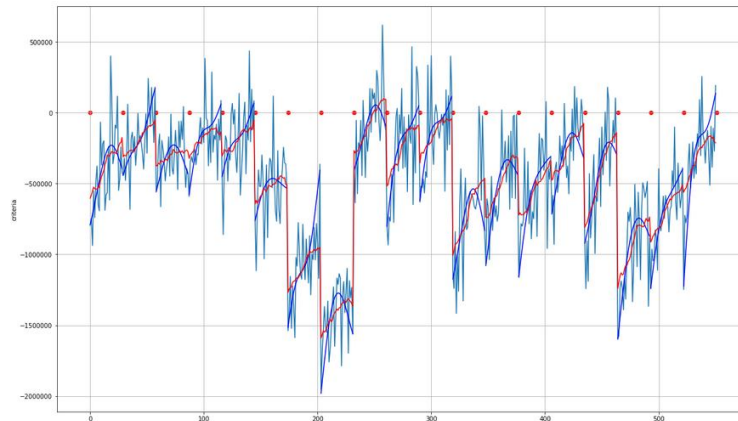


Figure 3. Computed difference between SSD criterion for our algorithm and kmeans++.
Random window generation mode

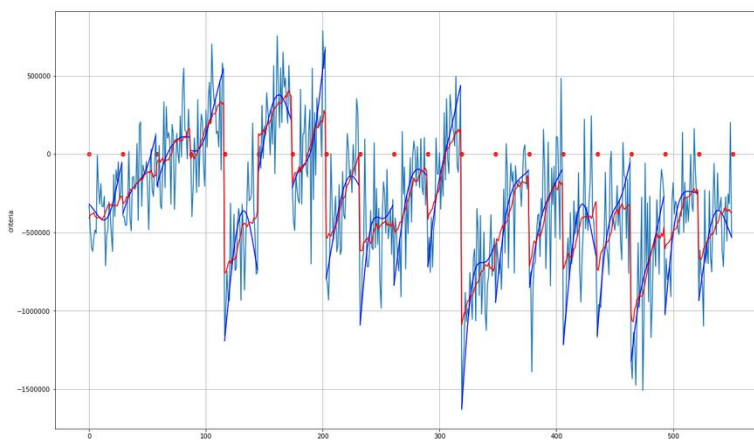


Figure 4. Computed difference between SSD criterion for our algorithm and k -means++.
Uniform window decomposition mode

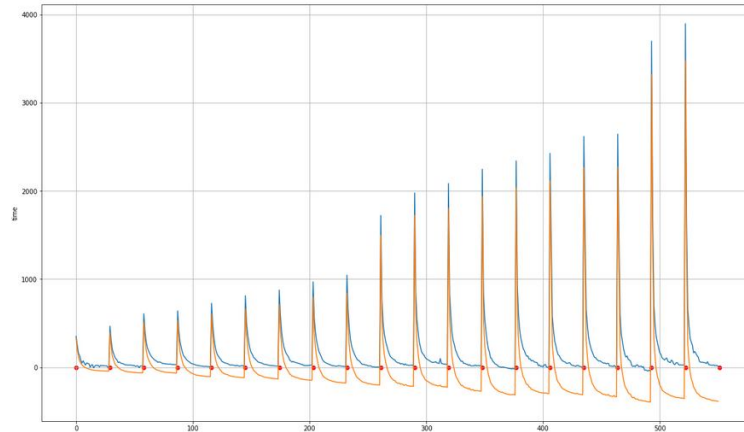


Figure 5. Comparing the computation times for our algorithm and k-means++ (the difference between two algorithms is shown). The lower chart represent pure decomposition clustering without initialization steps and step of evaluation on entire dataset. The peaks correspond to the largest sizes of windows ($1/5 \times N$). Monotonic declines correspond to window sizes $1/10 \times N, 1/15 \times N, \dots, 1/100 \times N$

Conclusions

In this approach we show that it is possible to achieve better results in the meaning of SSD criteria by applying iteratively the clustering procedure on subsets of the dataset. Obtained centroids are processed again by meta-clustering, resulting to the final solution.

Some observations:

- Centroids calculated by this method on large datasets provide reasonable SSD values even without clustering on the whole dataset.
- Window sizes should be representative, hence there is no sense in splitting the dataset for huge number of windows. In this case the number of degenerated clusters grows essentially.
- Slight improvement is detected on normalized data and in cases number of clusters k is small.
- Our experiments mostly support the idea that the precision of clustering results are highly dependant on the dataset-size, while it does not strongly depend on the clustering window size. Hence, the latter one may be optimized.

So the idea of clustering decomposition is well known we may generalize it to build a framework with meta-heuristic that allows to use well defined clustering algorithms and provide estimates to corresponding time/SSD quality reduction.

Acknowledgments

This research is conducted within the framework of the grant num. BR05236839 "Development of information technologies and systems for stimulation of personality's sustainable development as one of the bases of development of digital Kazakhstan".

References

1. Bagirov, A.M.: Modified global k -means algorithm for minimum sum-of-squares clustering problems. Pattern Recognition 41 (10), 3192--3199 (2008). DOI <https://doi.org/10.1016/j.patcog.2008.04.004>. URL <http://www.sciencedirect.com/science/article/pii/S0031320308001362>. (09.06.2019)
2. Bahmani B., Moseley B., Vattani A., Kumar R., Vassilvitskii S., Scalable k -means++. Proc. VLDB Endow. 5(7), 622–633, 2012.

3. Capó M., Pérez A., Lozano J.A., An efficient approximation to the k -means clustering for massive data. *Knowl.-Based Syst.* 117, 56–69, 2017.
4. Chiang M.C., Tsai C.W., Yang C.S., A time-efficient pattern reduction algorithm for k -means clustering. *Inform. Sci.* 181(4), 716–731, 2011.
5. Comparing different clustering algorithms on toy datasets. URL http://scikit-learn.org/stable/auto_examples/cluster/plot_cluster_comparison.html. (09.06.2019)
6. HajKacem M.A.B., N'Cir CE.B., Essoussi N. Overview of Scalable Partitional Methods for Big Data Clustering. In: Nasraoui O., Ben N'Cir CE. (eds) *Clustering Methods for Big Data Analytics. Unsupervised and Semi-Supervised Learning*. Springer, Cham, 2019.
7. Karlsson, C. (ed.): *Handbook of Research on Cluster Theory*. Edward Elgar Publishing, 2010.
8. Krassovitskiy A., Mussabayev R. Energy-Based Centroid Identification and Cluster Propagation with Noise Detection. In: Nguyen N., Pimenidis E., Khan Z., Trawiński B. (eds) *Computational Collective Intelligence. ICCCI 2018. Lecture Notes in Computer Science*, vol 11055. Springer, Cham, 2018.
9. Online clustering data sets UCI. URL <https://archive.ics.uci.edu/ml/datasets.html>. (09.06.2019)
10. Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G.J., Ng, A., Liu, B., Yu, P.S., Zhou, Z.-H., Steinbach, M., Hand, D.J. and Steinberg, D. (2007) Top 10 Algorithms in Data Mining. *Knowl. Inf. Syst.*, 14, 37. DOI 10.1007/s10115-007-0114-2, 2008.
11. Comparison of the K-Means and MiniBatchKMeans clustering algorithms. https://scikit-learn.org/stable/auto_examples/cluster/plot_mini_batch_kmeans.html. (09.06.2019)
12. S.J. Redmond, C. Heneghan, A method for initialising the K-means clustering algorithm using kd-trees. *Pattern Recogn. Lett.* 28(8), 965–973, 2007.
13. X. Cui, P. Zhu, X. Yang, K. Li, C. Ji, Optimized big data K-means clustering using MapReduce. *J. Supercomput.* 70(3), 1249–1259, 2014.

AN APPROACH FOR DETECTING SENTIMENTS IN MASS MEDIA PUBLICATIONS USING CONTENT SPECIFIC INFORMATIVE FEATURES

**Nurumov K., Baimakhanbetov M., Buldybayev T., Akoyeva I.,
Ospanova U.**

e-mail: Timur.Buldybayev@iac.kz

*Information-Analytical Center JSC under the Ministry of Education and Science
of the Republic of Kazakhstan*

Abstract. *The paper introduces a new approach of detecting in text sentiments using content specific features of mass media publications such as text objectivity, politicization, call to action, manipulation and public resonance. Moreover, the authors used parts of speech and seventeen topic dictionaries as additional features. The measurement of the features is based on the dictionary approach and tested using support vector machine, random forest and XGboost supervised learning algorithms with*

parameter grid search. The results indicate that the proposed methodology works as it indicates 83.86% of accuracy.

1. Introduction

Nowadays, textual data is one of the biggest sources for big data analytics. With the spread and availability of open source mass media publications and twitter data, rises the necessity of training robust machine learning algorithms that can analyze these textual data with an acceptable level of accuracy. Not surprisingly, text analytics and text mining became one of the most popular strands of research [1, 2]. Among many research topics, a particular interest attracted the topic of sentiment detection. Simply saying, sentiment analysis deals with the detection of positive or negative sentiments in the text. In terms of machine learning the actual detection can be in different ways, one can turn the problem into classification and/or regression tasks, or one can use manually annotated dictionaries and measure with them the actual frequency of positive and negative words in the text. Though, in general, text mining and sentiment analysis, in particular, are very popular, yet they are somewhat understudied in the Russian speaking scientific community. Despite the existence of successful attempts to create systems of sentiment detection [3, 4], there is much potential for increasing the accuracy of sentiment detection using content specific features of text publications. Our methodology proposes a combined dictionary, parts of speech and machine learning approach to detect sentiments in mass media publications. The core of the approach is domain specific content features of publications such as politicization, manipulation, objectivity, call to action and public resonance. We use three popular machine learning supervised algorithms to detect positive or negative sentiments. As a result, our proposed methodology indicates 83.86% of accuracy.

This paper is organized in the following way, in the next section we discuss related work, section three deals with the proposed methodology and measurement, section four presents the results of the machine learning procedures and section five concludes the paper with discussion.

2. Related work

Sentiment classification received huge interest from the scientific community. Early attempts in machine learning and statistical analysis focused on the binary classification of movie reviews [5]. Using several machine learning methods such as Naïve Bayes, support vector machine and maximum entropy, the authors tried to classify a collection of reviews falling into either positive or negative category. Subsequent research significantly expanded the topic and improved the overall accuracy while increasing the general complexity of the algorithms. For instance, Pang and Lee used only subjective parts of documents to model sentiment polarity, improving the accuracy from 82.8% to 86.4% [6]. Another attempt moved from the binary case to multiclass classification of the movie reviews using five-point rating scales [7]. McDonald et al. proposed a joint classification of documents and sentences which reduced the error of the sentiment detection [8]. Ikeda et al. modelled polarity shifters of words in sentences, the method improved the accuracy of the algorithm, particularly on the sentence level [9]. Maas et al. proposed a mix of unsupervised and supervised methods to capture the sentiment level of document-word level [10]. The authors employed a probabilistic model of words representation in documents and then used logistic regression to capture sentiments of the words. Xia et al. focused on the problem of polarity shift detection [11]. In particular, they elaborated a three-stage procedure which firstly divides documents into subsentences and statistically detects polarity shifts, secondly eliminates negations and

thirdly employs different combinations of weighted classifiers for sentiment classification. More recent research in sentiments analysis shifted to the incorporation of deep learning technics as well as sentence-level classification [12, 13, 14, 15].

In parallel with the machine learning approaches, scientific literature elaborated the lexicon-based approach. Instead of relying on rigorous statistical techniques, the proponents of lexicon approach extensively used annotated word dictionaries. For example, Turney and Littman used General Inquirer Lexicon to check the results of the unsupervised learning algorithm of the semantic orientation of words [16]. Denecke, using the SentiWordNet, proposed a methodology of polarity classification in a multilingual framework [17]. Ding et al. analyzed context-dependent words and particular linguistic constructs that are hardly detected by learning algorithms [18]. Taboada et al. developed a new measurement – semantic orientation that detects a measure of subjectivity and sentiment in texts [19]. Other research attempts Tan et al. and Mudinas et al. combined lexicon and machine learning methods to create a mixed procedure that could perform better than any of the two approaches separately [20, 21]. Moreover, despite the existence of robust sources of lexicons (SentiWordNet, SentiNet, LIWC) scholars elaborated new models such as VADER, which signalled classification accuracy of 0.96 (F1) [22]. The lexicon does not only relatively precisely detect sentiments in social media texts but can be generalized to other domains as well.

The topic of sentiment detection is not new, and the previous research achieved impressive results in the accuracy of machine learning algorithms and lexicon-based dictionaries. Our purpose in this work is not to obtain results of a higher accuracy or to invent a new algorithm, but to combine machine learning and lexicon approaches to test the new set of features derived from the content of media texts. In this respect our attempt to detect sentiments is close to the work of Mudinas et al. [21], however, we see our contribution in using a larger set of domain-specific features as well as applying more supervised learning algorithms to check the accuracy of the elaborated methodology.

3. Methodology

The proposed approach consists of using domain-specific informative features of mass media publications. They include the degree of objectivity, call to action, manipulation, degree of politicization and public response. Apart from that, we derived part of speech-specific features such as the number of adjectives, nouns, verbs, adverbs and pronouns. In order to measure our domain-specific features, we used part of speech tagging procedure and manually annotated dictionaries, created with the help of the experts from political science, communication and linguistics. After the PoS tagging, we used the dictionary of sentiment words together with negators, amplifiers and diminishers and labelled the corpus of 35 thousand publications. In the same way, the rest four dictionaries, containing content-specific words and phrases, were applied to derive the set of learning features. Next, three supervised learning algorithms were used to train the machine to detect the level of sentiment polarity; they include the-state-of-the-art support vector machine, random forest classifier and XGboost algorithm. The next subsections elaborate the proposed methodology in details.

3.1 Domain specific dictionaries

Content specific informative features are the cornerstone of the proposed methodological approach. They represent necessary information about the content of the publication, allows accessing the degree of objectivity and whether the author tried to politicize an issue which is not connected with politics. Moreover, they detect whether the information in the document is reliable and whether the author tried to use

manipulative words. In order to detect in text presence of these features, experts from linguistics manually annotated dictionaries of unbiasedness, call to action, politization, manipulation, connections, sources, persons and sentiments. To complete the dictionaries, we launch the procedure of dictionary expansion, which is based on an algorithm of finding similar word forms from the matrix of word co-occurrence.

The algorithm is simple, however, it allowed us to significantly expand our dictionaries. However, it must be noted that the algorithm returned the data in a raw format, or put differently it returned the words and word-forms in their original in text format. Thus, the only reasonable option was to lemmatize the resulting dictionaries and drop all duplicated entries. In addition to content features, we also used topic dictionaries, which were manually annotated by experts from the fields of political science, education, sociology, public policy and journalism.

Table 1. A short summary of the obtained topic dictionaries

Dictionary	Number of manually annotated words	Number of words after extension
objectivity	Reliable sources – 674 Words - 89	Reliable sources – 674 Words - 111
sentiment	positive – 979 negative - 2111	positive – 1272 negative - 2526
politization	30	34
call to action	89	145
civil society	184	232
medicine	141	174
industrial development and technologies	120	143
defence and security	76	89
education and science	154	169
government	96	121
crime and incidents	307	408
agriculture	62	69
social security	30	38
sport, culture and entertainment	242	307
labor	24	30
energy	69	87
ecology and tourism	40	52
emergency situations	53	73
administrative management	123	161
economics and finance	235	306
foreign policy	94	115

As one can notice in the negative and the positive sentiments, the number of words increased by 19.6% and 29% respectively. The algorithm added only four words to the politicize dictionary, however significantly expanded all 17 topic dictionaries.

3.2 Measurement procedure

Sentiment polarity. We first start by applying part of speech tagging on the corpus of 35645 media texts from five news websites. PoS tagging is somewhat problematic for the Russian language, as there are not many reliable libraries and open soft solutions of trained dictionaries, however, we used `udpipe` library¹ freely available for the C++ and R statistical software. After the PoS tagging, the texts were labelled with the help of the sentiment dictionary. The goal was to obtain an indicator of the sentiment in the range of $[-1:1]$, the indicator could be then used as a continuous or a categorical variable. In case of a categorical variable, one can either use a binary variable or can assign multiple categories by dividing the range into 11 groups with the middle group being the neutral category. In order to cover all scenarios, we labelled the texts by creating 3 variables that represented the overall sentiment polarity of the document. The document polarity is the constrained sum of sentence-level polarities, whereas the sentence level polarity was calculated using negators, amplifiers and diminishers. The negators were used to inverse polarity of the words if the negator was found before and after 4 words of the sentiment it would automatically reverse the sign of the sentiment word from +1 to -1. Consider the following simple sentence:

*The weather was **good** (1)*

*The weather was **not good** (-1)*

Without considering the negator *not*, one could mistakenly count the second sentence as positive, but the presence of the negator clearly inverses polarity of the sentiment. In the same manner, we used amplifiers and diminishers to capture the intensity of the sentiments as precisely as possible.

*The weather was **very good** (1)*

The algorithm detects amplifier *very* and assigns the weight of 0.5², in the same way, the algorithms identify words that diminish polarity of the words in each sentence. Thus, the resulting polarity obtained at the sentence level was summed at the document level in the range of $[-1:1]$.

Objectivity. Instead of using the feature of subjectivity, we use only objectivity. The feature presents several objective sentences in the text, whereas a sentence is considered to be objective if it contains at least one criteria of reliability. In general, there are four criteria of reliability: number, reference to website, reference to a public person and reference to public institute (organization). Only those sentences that contain at least one of the above-stated criteria can be considered as objective and receive the score of 1. We calculated objective sentences in a document as well as used dictionary of connections, reliable sources and persons. Thus, the overall criteria of objectivity were presented with the four derived features.

Politicization. A text is considered to be politicized if the author tried to politicize an issue which is not connected with politics. For example, it can be hardly imagined that the topic on healthy food can contain words, word forms or paragraphs on political issues. However, if such a case presents itself, then the text undoubtedly has some degree of politicization. Though we do not estimate the degree of politicization, nevertheless we estimate the frequency count of politicized words in the text.

¹ Wijffels, J. Tokenization, Parts of Speech Tagging, Lemmatization and Dependency Parsing with the “UDPipe”, “NLP” Toolkit. Retrieved from: <https://cran.r-project.org/web/packages/udpipe/index.html>

² The weight parameter can be changed thus to either increase or decrease the influence of positive and negative words.

Call to action. At attempt to stimulate reader's desire for action. We measure call to action by simple in text frequencies of the words from the dictionary. The resulting feature has a continuous distribution of real numbers from zero to n.

Public response. The reaction of the public to a particular mass media publication. We measure public response with the two continuous variables, number of comments and number of views.

Part of speech features. In addition to the dictionary features, we extract parts of speech such as nouns, adjectives, adverbs, verbs and pronouns.

Additional features. To complete the set of features, we count the number of sentences in each text, along with a number of words and number of characters.

Data. We use the corpus of 35643 publications from five mass media sources with the following distribution: tengrinews.kz (29128), kazpravda.kz (59), vesti.kz (3693), vlast.kz (890), kp.kz (1873).

4. Machine learning algorithms

We apply three supervised learning algorithms support vector machine, random forest and XGboosting. However, before conducting the machine learning algorithms, we put resulting 36 variables through the procedures of feature extraction.

4.1 Feature extraction

We first start by conducting the procedure of feature extraction. We apply the lasso model, which iteratively fits the features along the regularization path³. The model dropped 3 variables with an overall coefficient of zero; a number of views, word count and terms. The rest 33 variables have the magnitude from -0.21 to approximately 0.06 ; these set of features is further used to train the machine to detect the sentiment polarity. The sentiment polarity, in turn, is divided into two categories where the neutral category is assigned either to positive or to negative polarity. We follow the strategy of one vs all and thus create two outputs, one with the neutral category assigned to positive documents and the other with the neutral category assigned to negative documents. To check the results of the lasso model, we apply the procedure of random forest with 10-fold cross-validation and check the resulting importance of the input variables. The picture below shows 10-fold cross-validated results of the random forest. Overall, 9 features account for almost 81.4% of accuracy, adding the rest features increases the accuracy only on 1.2 percent. The top five features include the topic of crime and incidents, manipulation, adjectives, the topic of emergency situations, characters count.

Though the random forest indicates that the subset of 9 variables is enough for training the machine, nevertheless we use the set of 33 features as both, lasso model and random forest does not indicate a reduction in accuracy with the increase in the number of features.

³ The python 3.6 with the package sklearn. linear_model and function LassoCV is used to fit the model

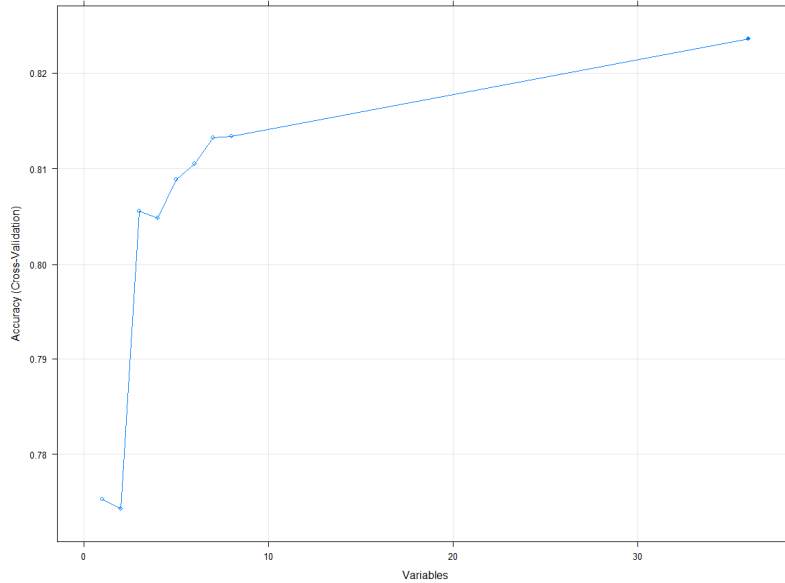


Figure 1. The feature extraction after random forest procedure

4.2 Supervised machine learning

We first start by running the support vector machine on the two output variables. We split the full data set of publications on training and testing following the scheme of 70/30 and normalize input variables. The SVM is particularly useful for classification tasks, and the only complication is that it has several tuning parameters. To avoid manual selection of parameters, we apply parameter grid search with the 5-fold cross-validation and try linear, polynomial and radial kernels⁴. The results for both output variables are presented below in tables from two to seven.

Table 2. (SVM with different kernels. Positive VS neutral+negative)

Kernel	Accuracy	F1	Precision	Recall	Sensitivity	Specificity	Cost	Degree or Sigma
linear	0.8155	0.7689	0.8349	0.7127	0.7127	0.8933	1	
radial	0.8186	0.7748	0.8327	0.7244	0.7244	0.8889	1	Sigma=0.056
polynomial	0.8186	0.7748	0.8327	0.7243	0.7243	0.8899	1	3, scale=0.01

From the table above it is clear that the SVM with the polynomial kernel of 3rd degree has better results of the specificity parameter, though the rest results are identical to the SVM with the radial kernel. At the same time, both kernels do not differ much from the simple linear classification. Next, we run the random forest procedure built on ensemble of trees with the 5-fold cross-validation. The results are presented in table 3 below.

⁴ All calculations of svm and random forest were done in R statistical programming software, package *caret*. For more information on *caret* please see, <http://topepo.github.io/caret/index.html>

Table 3. Random forest. Positive VS neutral+negative

Procedure	Number of trees	Accuracy	F1	Precision	Recall	Sensitivity	Specificity
Random forest	1000	0.8257	0.7910	0.8180	0.7656	0.7656	0.8712

The procedure shows better accuracy, F1 score, recall and sensitivity than the SVM model, however, fails to indicate better results of specificity and precision. Finally, we conclude with the XGboost algorithm. Table four shows the output of the procedure. Comparing to other models XGbosst indicates slightly increased results in all parameters except the F1 score, which is less than the output from the random forest.

Table 4. Performance of the XGboost. Positive VS neutral+negative

Procedure	Number of trees	Accuracy	F1	Precision	Recall	Sensitivity	Specificity	Parameters
XGboost	250	0.8260	0.7885	0.8194	0.7598	0.7598	0.8753	eta=0.3 gamma=0

Though the above models show only slight differences in accuracy, nevertheless it is clear that the XGboost does a better job of finding negative sentiments as well as detects both classes with slightly better accuracy. Next, we continue following the one vs all strategy and use the output of negative vs neutral + positive variable on the same set of models.

Table 5. SVM with different kernels. Negative VS neutral+positive

Kernel	Accuracy	F1	Precision	Recall	Sensitivity	Specificity	Cost	Degree or Sigma
linear	0.8278	0.7812	0.8345	0.7344	0.7344	0.9033	1	
radial	0.8283	0.7681	0.8332	0.7123	0.7123	0.9053	1	Sigma=0.064
polynomial	0.8359	0.7812	0.8345	0.7344	0.7344	0.9033	1	3, scale=0.01

Table 5 shows that the new output variable increases the accuracy of the SVM model. As in table one, the SVM with polynomial kernel produces better results than the radial and linear kernels, with the exception of specificity, where the radial has a slightly better number. At the same time, like in the case of positive vs neutral+negative output, the ensemble of trees improves detection of sentiments but indicates slightly less negative class than the polynomial SVM. Unlike the results above the XGboost output does not refine the accuracy, but increases detection of true positive sentiments almost on 7%.

Table 6. Random forest. Negative VS neutral+positive

Procedure	Number of trees	Accuracy	F1	Precision	Recall	Sensitivity	Specificity
Random forest	1000	0.8386	0.7880	0.8280	0.7517	0.7517	0.8963

Table 7. Performance of the XGboost. Negative VS neutral+negative

Procedure	Number of trees	Accuracy	F1	Precision	Recall	Sensitivity	Specificity	Parameters
XGboost	250	0.8235	0.82	0.82	0.82	0.82	0.83	eta=0.3 gamma=0

Thus, the comparative performance of three machine learning algorithms suggests that in our case the ensemble of random forest is overall better than XGboost and SVM models for detecting negative and positive sentiments. Here it is necessary to point out, that the purpose of the training procedures was not to find the most suitable model, but to check the proposed methodology of in-text measurement of sentiments. Subsequent research can pay more attention to model training and finding the best set of parameters or testing the methodology on a broader set of algorithms. For now, the evidence obtained from the test data is in favor of our dictionary-based approach. It is clear that the informative features of the content of mass media publications, the topic dictionaries and parts of speech features are enough to achieve an acceptable level of classification accuracy.

Discussion

Though we achieved relatively satisfying results, nevertheless much work must be done further to improve the sentiment detection using the proposed methodology. There are several potential ways of improving the accuracy of sentiment detection. First, it is necessary to improve the existing dictionaries and rather concentrate on the precision of every word in the dictionary rather than working on simply expanding the dictionaries. Alternatively, it is possible to measure the magnitude of sentiments, so far, we only assigned either positive or negative category to the words in the dictionaries. One can attempt to assign the magnitude of the sentiments to every word in the dictionary for example from -5 to 5, or from -3 to 3, where the end categories indicate weak and strong sentiments respectively. Second, formalize sentiment detection within sentences, that is work on formal constructions of sentiment negators, amplifiers and diminishers. We managed to formalize and detect constructions with a maximum of three words, but further attempts can aim at formalizing more than three words. Third, one can endeavor to compare the proposed algorithm with the existing online dictionaries as well as trained models of sentiment detection in the English language. Such a step perhaps would be the most valuable attempt to understand whether the proposed methodology can be compared to existing state-of-the-art systems in a multilingual framework. Finally, with the availability of software, another attempt can focus on the bigger data set of publications, we only used roughly 36 thousand texts. Though the number is enough, nevertheless the law of big numbers suggests repeating the attempt on a bigger corpus of publications.

This research is conducted within the framework of the grant num. BR05236839 "Development of information technologies and systems for stimulation of personality's sustainable development as one of the bases of development of digital Kazakhstan".

References:

1. Berry, W.M. (Ed.). Survey of Text Mining: Clustering, Classification and Retrieval / W.M. Berry. – New-York, Springer Verlag, 2004. – P. 244.
2. Aggarwal, C., & Zhai, Ch. (Eds.). Mining Text Data / C. Aggarwal, & Ch. Zhai. – New-York, Spronger Verlag, 2012. – P. 521.
3. Chetviorkin, I. & Loukachevitch N. Extraction of Russian Sentiment Lexicon for Product Meta-Domain / I. Chetviorkin & N. Loukachevitch // In Proceedings of COLING. – Mumbai, 2012. – P. 593-610.
4. Chetviorkin, I. & Loukachevitch N. Evaluating Sentiment Analysis Systems in Russian / I. Chetviorkin & N. Loukachevitch // In Proceedings of the 4th International Workshop on Balto-Slavic Natural Language Processing. – Sofia, 2013. – P. 12-17.
5. Pang, B., Lee, L., & Vaithyanathan, S. Thumbs up? Sentiment Classification using Machine Learning Techniques / B. Pang, L. Lee & S. Vaithyanathan // In Proceedings of the Conference on Empirical Methods in Natural Language Processing. – Philadelphia, 2002. – P. 79-86.
6. Pang, B. & Lee, L. A Sentimental Education: Sentiment Analysis Using Subjective Summarization Based on Minimum Cuts / B. Pang & L. Lee // In proceedings of the 42th Annual Meeting of the Association for Computational Linguistics. – Barcelona, 2004. – P. 271-278.
7. Pang, B. & Lee, L. Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Rating Scales / B. Pang & L. Lee // In Proceedings of the 43th Annual Meeting of the Association for Computational Linguistics. – Ann Arbor, Michigan, 2005. – P. 115-124.
8. McDonald, R., Hannan, K., Neylon, T., Wells, M., & Reynar J. Structured Models for Fine-to-Coarse Sentiment Analysis / R. McDonald, K. Hannan, T. Neylon, M. Wells, & J. Reynar // In proceedings of the 45th Annual Meeting of the Association of Computational Linguistics. – Prague, 2007. – P. 432-439.
9. Ikeda, D., Takamura, H., Ratinov, L., & Okumura M. (2010) (or 2008). Learning to Shift the Polarity of Words for Sentiment Classification. / D. Ikeda, H. Takamura, L. Ratinov, & M. Okumura // Transactions of the Japanese Society for Artificial Intelligence. – 2008 or 2010. – Vol. 25(1). – P. 50-57.
10. Maas, A., Daly, R., Pham, P., Huang, D., Ng, A., & Potts C. Learning Word Vectors for Sentiment Analysis / A. Maas, R. Daly, P. Pham, D. Huang, A. Ng, & C. Potts // In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. – Portland, 2011. – P. 142-150.
11. Xia, R., Xu, F., Yu, J., Qi, Y., & Cambria, E. Polarity Shift Detection, Elimination and Ensemble: A Three-Stage Model for Documents-Level Sentiment Analysis / R. Xia, F. Xu, J. Yu, Y. Qi, & E. Cambria // Information Processing and Management. – Vol. 52, 2015 or 2016. – P. 36-45.
12. Araque, O., Corcuera-Platas, I., Snachez-Rada, F., & Iglesias, C. Enhancing Deep Learning Sentiment Analysis with Ensemble Techniques in Social Applications / O. Araque, I. Corcuera-Platas, F. Snachez-Rada, & C. Iglesias // Expert System with Applications. – Vol. 77, 2017. – P. 236-246.
13. Chen, T., Xu, R., He, Y., & Wang, X. Improving Sentiment Analysis Via Sentence Type Classification Using BiLSTM-CRF and CNN / T. Chen, R. Xu, Y. He, & X. Wang // Expert Systems with Applications. – Vol. 72, 2016. – P. 221-230.
14. Jianqiang, Z., Xiaolin, G., & Xuejun, Z. Deep Convolution Neural Networks for Twitter Sentiment Analysis / Z. Jianqiang, G. Xiaolin & Z. Xuejun // IEEE Access. – Vol. 6, 2018. – P. 23253-23260.

15. Zhang, L., Wang, S., & Liu, B. Deep Learning for Sentiment Analysis: A Survey / L. Zhang, S. Wang, & B. Liu // Wiley Interdisciplinary Review: Data Mining and Knowledge Discovery. – 2018. – DOI: 10.1002/widm.1253
16. Turney, P. D., & Littman, M. L. Unsupervised Learning of Semantic Orientation from a Hundred-Billion-Word Corpus (Report No. ERB-1094) NRC 44929 / P.D. Turney & M.L. Littman. – 2002.
17. Denecke, K. Using SentiWordNet for Multilingual Sentiment Analysis / K. Denecke // In Proceedings of the 2008 IEEE 24th International Conference on Data Engineering Workshop. – Cancun, 2008. – P. 507-512.
18. Ding, X., Liu, B., & Yu, P. A Holistic Lexicon-Based Approach to Opinion Mining / X. Ding, B. Liu, & P. Yu // In Proceedings of the 2008 International Conference on Web Search and Data Mining. – Palo-Alto, 2008. – P. 231-240.
19. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. Lexicon Based Methods for Sentiment Analysis / M. Taboada, J. Brooke, M. Tofiloski, K. Voll, & M. Stede // Computational Linguistics. – Vol. 37(2), 2011. – P. 267-307.
20. Tan, S., Wang, Y., & Cheng X. Combining Learn-Based and Lexicon-Based Techniques for Sentiment Detection Without Using Labeled Examples / S. Tan, Y. Wang, & X. Cheng // In Proceedings of the 31st Annual International ACM SIGIR Conference of Research and Development in Information Retrieval. – Singapore, 2008. – P. 743-744.
21. Mudinas, A., Zhang, D., & Levene, M. Combining Lexicon and Learning Based Approaches for Concept-Level Sentiment Analysis / A. Mudinas, D. Zhang, & M. Levene // In Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining. – New-York, 2012. – P. 1-8.
22. Hutto C. & Gilbert E. VADER: a parsimonious rule based model for sentiment analysis of social media text / C. Hutto & E. Gilbert // In Proceedings of the Eight International Conference on Weblogs and Social Media. – California, 2014. – P. 216-225.

INITIAL EXPLORATION FOR TOPIC MODELING WITH WORD SENSE

Toleu A., Tolegen G., Mussabayev R.

e-mail: alymzhan.toleu@gmail.com, gulmira.tolegen.cs@gmail.com,
rmusab@gmail.com

Institute of Information and Computational Technologies, Almaty, Kazakhstan

Abstract. *This paper presents the clustering algorithm-based approaches for the tasks of word sense discovering and topic modeling. The former aims to find the number of word senses for the target words by clustering its context words. The latter is to solve the problem of topic modeling by clustering the words in a vocabulary which obtained from a large collection of corpora. The models were evaluated on English Wikipedia and 20 Newsgroups collection. The preliminary experimental results show that it is feasible to automatically discover the word senses by clustering the context words with their word representations. Result of clustering-based topic modeling shows that the clusters of unique words from corpora can be viewed as topics, but it has many “noise” words (non-*

topic words) should be identified or removed in advance. This work is a preliminary attempt to perform the topic modeling without a run through the large collection of corpora but only uses unique words of them.

Keywords: *topic modeling; word sense discovery; representation; LDA; PLSA*

1. Introduction

In this paper, we consider the problem of modeling large text corpora by using various approaches. The goal is to find short descriptions of documents that enable efficient processing of big data. The short descriptions can be in many forms, one of them is a brief summary, keyphrases, keywords and etc. from a document., which involves different task such as summarization [20] and key phrase extraction [1, 2].

When it comes to processing large collections of the document, the short descriptions become more important. An efficient way of describing large collections of corpora is to use topics [6, 8] which reveals both the meaningful and important information, and also preserving the essential statistical relationships useful for tasks such as classification, novelty detection, etc. Using a bunch of topics to describe large collections of documents, we may lose the detailed information like the word order, but for the tasks like clustering the documents, media-level document analysis, there is no need to know these detailed pieces of information, and using the short description which able us to speed up the analyzing process.

Uncovering the latent topics within the text corpora has been studied extensively in recent years. Mixture of unigram [7], probabilistic latent semantic analysis (PLSA) [11, 17], Latent Dirichlet allocation (LDA) [8] are the widely used approaches for topic modeling. In general, in those models, the documents are models as a mixture of topics where a topic is a probabilistic distribution over words. Then utilize statistical techniques [12] to learn the topic components and mixture coefficients of each document. LDA is a generative probabilistic model of a corpus. The generative process of LDA as follows: first choose a distribution over the topics, then for each word, choose a topic assignment and choose the word from the corresponding topic. In essence, those conventional topic models reveal the latent topics by estimating the distribution over documents to topics, topics to words. These two parameters are often in the form of the matrix, and the former need to be estimated by capturing the global information across corpora, which may slow down the process of analysis.

In this paper, we present two approaches for word sense discovering and topic modeling, which based on various clustering algorithms [3]. To automatically find the word senses, we describe an algorithm in which we extract the bigrams from a collection of corpora at first, then build the indexing for searching the neighboring words, and use clustering algorithms to classify these neighboring words. The number of clusters may indicate the number of word senses. Following this intuition, we run the algorithm on 500,000 English Wikipedia documents. For topic modeling, firstly, we use four different representations to represent words, namely the number of common context words, co-occurrences, point mutual information and pre-trained word embedding [9,10]. From the visualization results, we could observe that using those representations, the words can be projected into different vector space with a different distribution, which allows us to do the clustering with various setup and compared their results. Secondly, we utilize affinity propagation and k-means algorithms to do clustering over unique words which were extracted from the corpora. We report the coherence score with different settings and compare our clustering-based approach to LDA, the state-of-the-art for topic modeling. The preliminary experimental results show that LDA has the ability to identify the number

of topics contained in the corpora, but the clustering-based approach not, this issue should be solved in future work.

2. Notations

We process large size of text corpora D which consists of a collection of M documents, denoted by $D = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M\}$.

A document \mathbf{w} is a sequence of words, denoted by $\mathbf{w} = (w_1, w_2, \dots, w_N)$.

A word is basic unit of the text, its original form is discrete, denoted by w . We use several types of representation to use this unit.

Topics are denoted by $\mathbf{Z} = (z_1, z_2, \dots, z_{|Z|})$. Each topic $z = (w_1, w_2, \dots, w_{|z|})$ contains words belong to that topic.

3. Representations

A discrete word w is projected into a space vector by using the following different calculations:

- Common context words' number, represent a word by the number of the common words with its neighboring words:

for a given word w

use function $\mathbf{nw} = \mathit{neigh}(w)$, find all the neighboring (context) words of the w ;

use function $\mathbf{nw}_{\text{remaining}} = \mathit{neigh}(\neg w)$, find all the all the neighboring words of remaining words in the vocabulary;

for each set in $\mathbf{nw}_{\text{remaining}}$, do intersection operation with \mathbf{nw} ;

the number of intersection of all two sets be the representations of the word w .

- Co-occurrence, calculate how many time the two words occurred in the large corpora, the representation of the w is the co-occurrence between the w and the remaining words in the vocab.

- Point mutual information [13]:

$$PMI(w_1, w_2) = \frac{\#w_1 \#w_2}{\#(w_1, w_2)}$$

$\#w_1$ - is the frequency of the word w_1 .

$\#(w_1, w_2)$ - is the co-occurrence.

- Pre-trained word embedding [19]:

It is a simple feed-forward neural network that trains large corpora by predicting the target word given the context or vice versa.

4. Methodology

Word Sense Discovery

Word sense discovery is a task of automatically finding the word senses without using any dictionaries which have manually defined their senses. For instance, a given word “mouse” commonly has two senses: one refers to a live mouse; another could be the computer device mouse. These two senses can be identified by the word context, if the word has context words have related with computer device words such as “keyboard”, “computer” etc. Its sense can be identified to the second. This task is called word sense disambiguation (WSD) [5].

In this work, we focus on automatically finding word senses by analyzing its neighboring words which give us the most distinguishable information as described

above. More formally, for a word w , we want to find how many senses ns does the w have. The discovery process can be following:

Algorithm 1. Word sense discovery approach.

Input: $D = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M\}$ documents with words $\mathbf{w} = (w_1, w_2, \dots, w_N)$.

Output: $ns = \{ns_1, \dots, ns_{|V|}\}$, the number of senses for words in the vocabulary V .

building the vocabulary $V = \{w_1, \dots, w_{|V|}\}$ from document D .

for word $w_i \in V$, $1 < i < |V|$:

find the neighboring words $\mathbf{nw}_i = neigh(w_i)$;

for word $w_j \in \mathbf{nw}_i$, $1 < j < |\mathbf{nw}_i|$

find the neighboring words $\mathbf{nw}_j = neigh(w_j)$;

use the words' frequency in \mathbf{nw}_j to represent the word w_j ;

$ns_i = clustering(\mathbf{nw}_i)$

return ns

The clustering can be any different algorithms which can be K-means [14] and Affinity propagation [15]. The former need to specify the cluster number in advance, but for word senses discovery, the class number is we want to find out. It not practicable to use K-means, but it is interesting to see how the algorithms perform when we know how many senses the target word has. In contrast to K-means, Affinity propagation is a clustering algorithm that can identify the cluster number automatically, it fit the problem, to find the number of word senses. In the experiment, we use these two clustering approaches to make comparisons.

Clustering-based Topic Modeling

We attempt to use clustering approaches to perform topic modeling. The main idea behind is that the cluster can be treated as topics. Each cluster (from now on, the word cluster and topic are used interchangeably) contains a bunch of words to describe this topic. This approach is designed to solve the problem of processing the large collection of corpora efficiently and make the topic modeling algorithm as simple as possible. Below we describe our clustering-based topic modeling approach:

Algorithm 2. Clustering-based topic modeling approach.

Input: $D = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M\}$ documents with words $\mathbf{w} = (w_1, w_2, \dots, w_N)$.

Output: $Z = \{z_1, \dots, z_{|Z|}\}$, the topics.

building the vocabulary $V = \{w_1, \dots, w_{|V|}\}$ from document D .

for word $w_i \in V$, $1 < i < |V|$:

calculate the representations (described in Section 3) for w_i ;

$Z = apply\ a\ clustering\ algorithm\ to\ V$,

return Z

5. Experiments

5.1. Data Set

We use English Wikipedia as text corpora, the 500,000 documents were extracted from the whole English Wikipedia dump file. In order to find neighboring words for the target word efficiently, we first extracted all bigrams with their co-occurrences from those documents. Then we indexing each word in the vocabulary by using these bigrams, which enable us to return the neighboring words for a given word in $O(1)$. Figure 1 and Figure 2 show the distribution and statistics of the extracted bigrams.

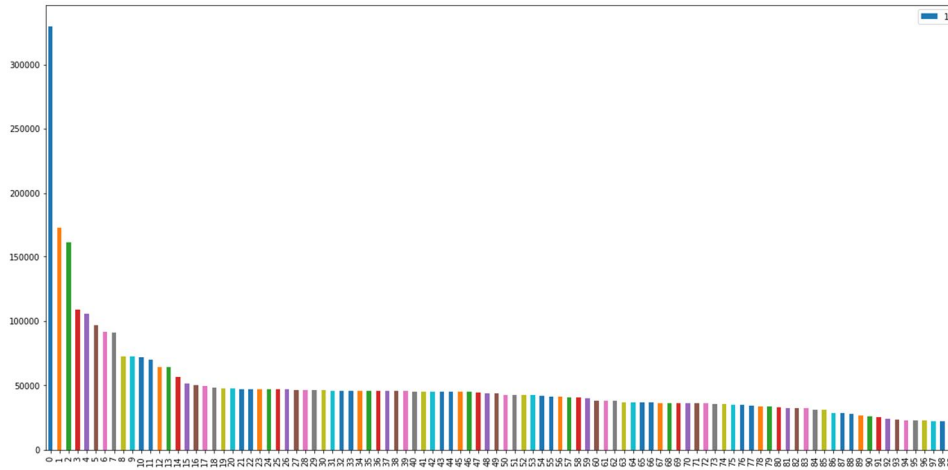


Fig 1. The distribution of bigrams frequency extracted from 500,000 English Wikipedia documents. X-axis denotes document ID, y-axis refers to frequency.

0	united_states	329419	8	per_square	64555
1	new_york	173211	9	square_mile	64138
2	median_income	108740	10	war_ii	56420
3	world_war	105956	11	population_density	51709
4	years_age	96437	12	housing_units	50046
5	age_older	91224	13	also_known	49763
6	high_school	71741	14	married_couples	48624
7	males_median	69776	15	median_age	47718

Fig 2. Top 16 extracted bigrams with its frequencies.

The experiments for topic modeling were carried out on the 20Newsgroup collection, a standard corpus including 18,828 messages harvested from 20 different newsgroups. Each newsgroup corresponding to a different topic. Table 1 and Table 2 lists the number of documents and the topic names contained in the corpus.

Table 1. Train and test set for topic modeling

	#documents	#unique words
Train	11314	14707
Test	7532	11164

Table 2. The topic names in the 20 News- groups collection

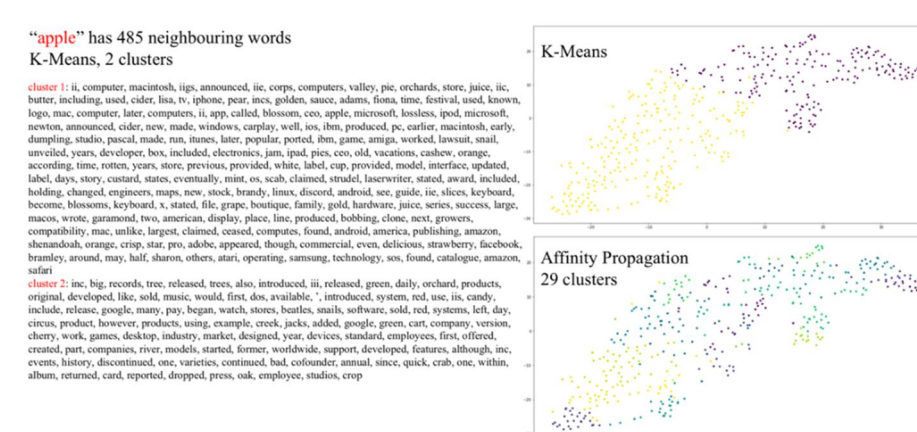
No.	Topic name	No.	Topic name
1	alt.atheism	1	rec.sport.hockey

2	comp.graphics	2	sci.crypt
3	comp.os.ms-windows.misc	3	sci.electronics
4	comp.sys.ibm.pc.hardware	4	sci.med
5	comp.sys.mac.hardware	5	sci.space
6	comp.windows.x	6	soc.religion.christian
7	misc.forsale	7	talk.politics.guns
8	rec.autos	8	talk.politics.mideast
9	rec.motorcycles	9	talk.politics.misc
10	rec.sport.baseball	10	talk.religion.misc

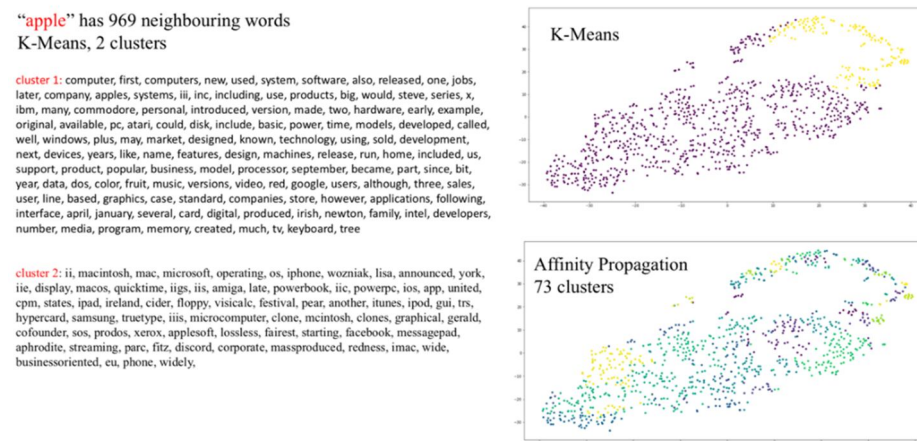
5.2. Results: Word Sense Discovery

Firstly, we test our algorithms on a simple configuration:

- i) give a single ambiguous word like “apple”;
- ii) use two clustering algorithms: K-means with specified word sense number (cluster number) and Affinity propagation automatically identify the word sense number;
- iii) use the various number of neighboring words (a hyper-parameter). The purpose is to see how the algorithm performs when we give a single ambiguous word.



a) 500 neighboring words



b) 1000 neighboring words

Fig 3. Example of word sense discovery for the target word “apple”.

Figure 3 presents the results for finding the senses of the word “apple”, figure 3(a) shows the algorithm uses 500 neighboring words. As we know, the word apple has two common senses, one for fruits, another for the company (we lowercased all the words). For K-means, it can be seen from the visualization, the words are grouped into two clusters (the cluster number was specified manually), the left side of figure 3(a) shows what kind of words are contained in those two clusters. Although there are many “noising” words are grouped in two clusters, the second cluster is more likely to be the sense of plant/fruits, and the first cluster is more likely to be the first sense, company.

Figure 3(b) shows the algorithm uses 1000 neighboring words, for K-means with 2 clusters set up, the results also show the case with 500 neighboring words. For affinity propagation, we can observe that the number of clusters is increased to 73 clusters compared to the algorithm uses 500 neighboring words (29 clusters). It means that the word “apple” has many senses, but it not practically meaningful. This is the problem should be solved in the future work, the number of the word senses increases with the number of neighboring words and it is unmeaningful. The simplest way to solve this problem is to reduce the number of the neighboring words, it is not the final solution (it leaves for the future work), here we just give it to try to see the results.

Secondly, we run the algorithm on 3000 words to automatically identify the number of senses of those words. As we found the 3000 words have 1267 neighboring words in average. Table 3 shows the results, at first, we reduce the neighboring words into 100 which gives 12.56 word senses in average for 3000 words. It means that each word has around 12 meanings, it still not practically meaningful. When we decrease the neighboring words into 15, we can see that the average word senses decreased into 3.667, it seems to be meaningful.

Table 3. Results for word sense discovering of 3000 words with different neighboring words

#Neighboring words	Avg. #word senses	Unambiguous words
100	12.56	4.3%
15	3.667	3%

5.3. Results: Topic Modeling

Quality of Topics

In order to evaluate the resulting topics, we utilize coherence score [4] as evaluation metric. The main idea of coherence is that a good topic should consist of words in cohesive semantic similarity. It is calculated as follows:

$$C = \frac{1}{K} \sum_{z=1}^K \sum_{m=2}^M \sum_{l=1}^{m-1} \log \frac{n_D(w_m^z, w_l^z) + \epsilon}{n_D(w_l^z)}$$

where $[w_1^z, w_2^z, \dots, w_M^z]$ denotes top-M representative words for topic z . $n_D(w_l^z)$ is the frequency of word w_l^z and $n_D(w_m^z, w_l^z)$ denotes co-occurrence counts in the corpus. C is negative number, the higher value indicates better results.

We run the algorithms on data sets with a different setting: the topic number $K=10$, $K=20$, $K=30$ and calculate coherence scores by choosing M from 5 to 20. Table 4 shows the results for LDA and the clustering-based (k-means) topic model. It can be seen that LDA still outperforms clustering-based approaches. As know that the data set has 20

topics, the results should be better at this point $K=20$, LDA has reached this. We can observe that LDA gives -37.09, -36.69, -41.23 coherence score for $K=10, 20, 30$ respectively. When $K=20$, the results of LDA increases and when $K=30$ the results decreased, which indicates LDA model has abilities to identify how many real topics are contained in a corpus. Contrast to this, clustering-based approach's results increasing when K is larger than the real topic number. This is a problem should be solved in future work.

Table 4. Comparison results of topic coherence scores

#Topics	Models	M=5	M=10	M=20
K=10	LDA	-7.85	-37.09	-85.49
	Clustering-based	-29.52	-54.54	-105.12
K=20	LDA	-15.26	-36.69	-86.57
	Clustering-based	-26.65	-49.67	-93.59
K=30	LDA	-17.80	-41.23	-90.10
	Clustering-based	-24.21	-42.68	-79.60

Clustering Documents with Topics

Clustering documents with their topics is another better way to evaluate the topic modeling approaches. In this experiment, we treat each topic as a cluster, and assign each document \mathbf{w} to the topic z with the highest value of conditional probability $P(z|\mathbf{w})$. We adopt two standard metrics to evaluate the clustering results. Let $\Omega = \{\omega_1, \dots, \omega_K\}$ be a set of output clusters and $\mathbb{C} = \{c_1, \dots, c_P\}$ be labeled classes of documents.

- Purity. It is an external evaluation criterion of cluster quality. It is the percent of the total number of documents that were classified correctly, in the unit range $[0..1]$.

$$Purity(\Omega, \mathbb{C}) = \frac{1}{n} \sum_{i=1}^K \max_j |\omega_i \cap c_j|$$

- Normalized Mutual Information (NMI). Note $I(\Omega; \mathbb{C})$ is the mutual information between two sets Ω and \mathbb{C} , and $H(\Omega), H(\mathbb{C})$ are the entropy.

$$NMI(\Omega, \mathbb{C}) = \frac{I(\Omega; \mathbb{C})}{[H(\Omega) + H(\mathbb{C})]/2}$$

Table 5 shows the results for documents clustering with a different setting: K from 5 to 30. For both metrics, the higher value indicates better results. It can be seen that the purity scores for LDA and clustering-based approach are comparable. As mentioned the data set 20news groups have 20 classes, and the number of unique words from the test set is 11164, which means the algorithms group the such number of words into K classes. If the class number K is small, the clustering-based approach cannot give better results as expected, it is also proven on metric NMI.

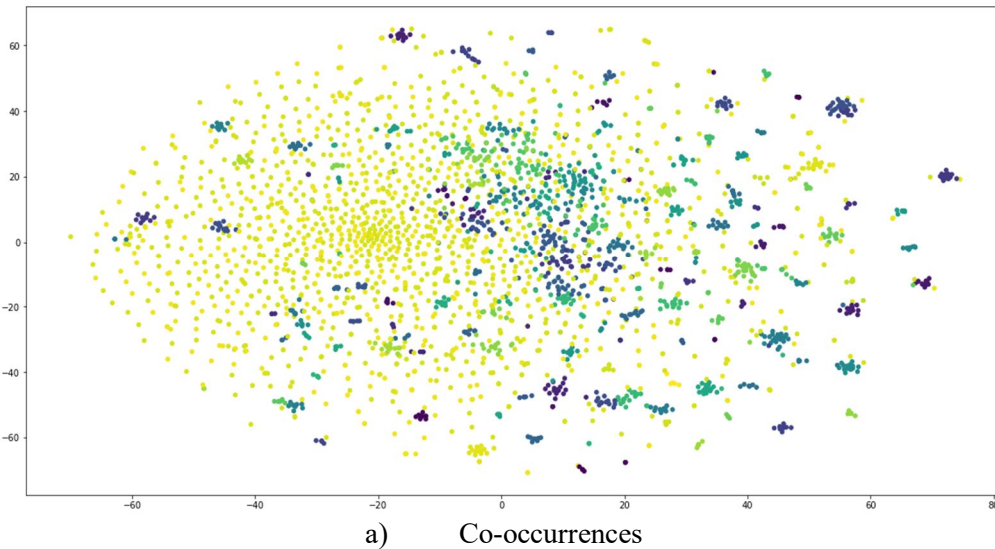
For metric NMI, LDA gives the highest results for $K = 5$ setting and the clustering-based approach yields the higher NMI value (0.263) at $K = 15$ among its results which is comparable with LDA (0.257). With class number K increases ($K > 15$), the NMI value of the clustering algorithms and LDA drops as expected, which fit the fact that the data set contains 20 topics.

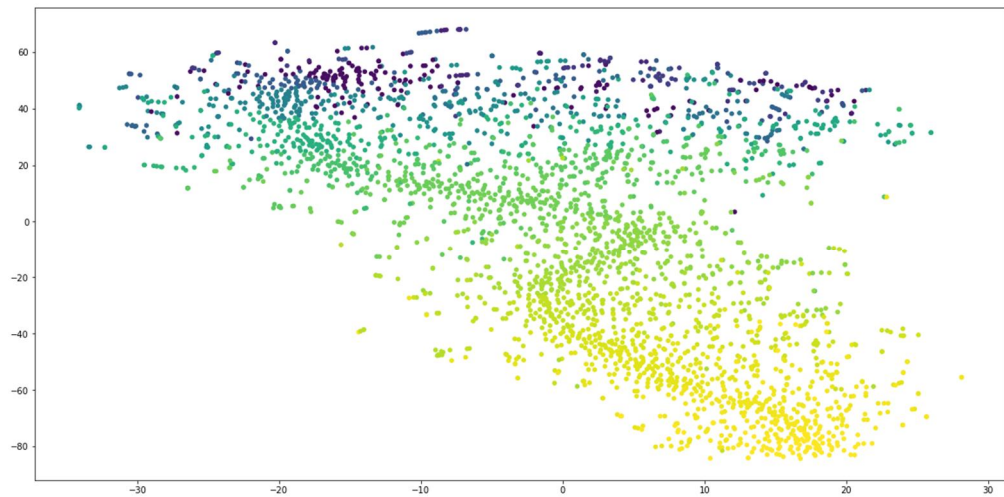
Table 5. Clustering performance of LDA and the clustering-based approach with different classes on 20newsgroups collection

Metrics	Models	K=5	K=10	K=15	K=20	K=25	K=30
Purity	LDA	0.216	0.197	0.232	0.230	0.212	0.250
	Clustering-based	0.098	0.205	0.262	0.230	0.238	0.246
NMI	LDA	0.315	0.249	0.257	0.226	0.238	0.263
	Clustering-based	0.074	0.225	0.263	0.218	0.225	0.227

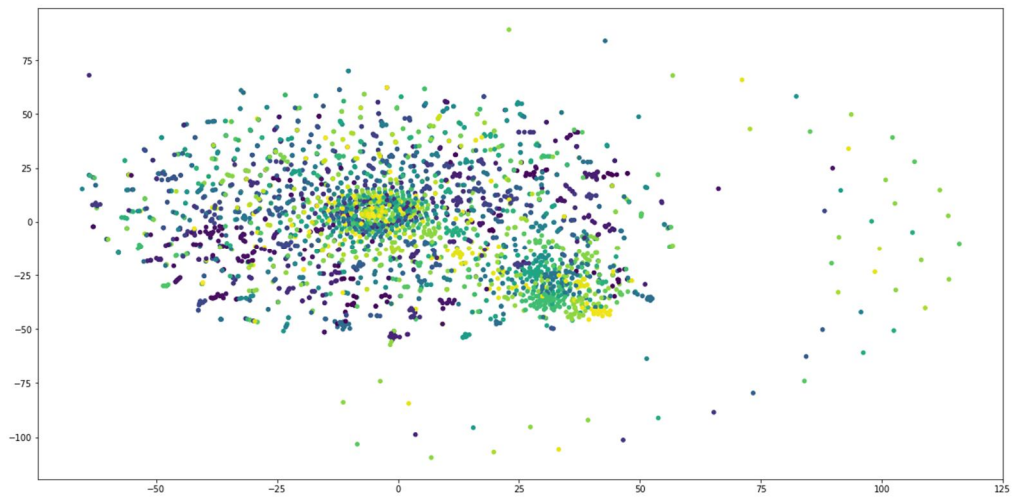
Visualization.

Figure 4 and Table 6 show the topic modeling results for representation co-occurrences, PMI and pre-trained word embedding. Obviously, using the different representations, the distribution of words in vocabulary have different distribution. We use affinity propagation as clustering algorithm which allow us to identify topic number automatically which can be a problem in model LDA. Table 6 shows the corresponding top 15-topics from the algorithm output, we apply the clustering algorithm on 3000 words. It can be seen that for co-occurrence, PMI and pre-trained embedding the clusters' numbers are 163, 285 and 307. From the real results, we observe that pre-trained word embedding as representation works well than others. Because the results from other representation contain several unrelated words are grouped together. Figure 4 shows the t-SNE [16] visualization of results from affinity propagation clustering algorithm using different representations.





b) Point mutual information (PMI)



c) Pre-trained word embedding

Fig 4. Results for topic modeling with co-occurrences, PMI, and word embedding
Table 6. Results of topic modeling with 15 topics

Affinity Propagation, 3000 words, 163 clusters, Co-occurrences

- 1: median, males, dollars
- 2: income, average, family, batting
- 3: older, units, gender sense
- 4: age, children, residents, stone, golden, bronze, iron, retirement, none
- 5: york, makeup, residing sense 6: also, may, see, could, must
- 7: every, later, following, refer, immediately
- 8: personal, rest, quality, everyday, aspects, loss
- 9: hall, roll, entrance, residence, concert, carnegie, albert, dining
- 10: park, office, station, airport, stations, centre
- 11: area, geography, winters
- 12: american, take, takes, censusdesignated, taking, taken

13: line, year, days, months, weeks, decades, month, hour, week, centuries, decade, minutes 14: sea, china, mediterranean, baltic, altitude, dead, timor, caribbean

15: population, city, town, township, cdp, village, cent, majority, estimated, distribution, citys, percent, boroughs, townships

Affinity Propagation, 3000 words, 285 clusters, PMI

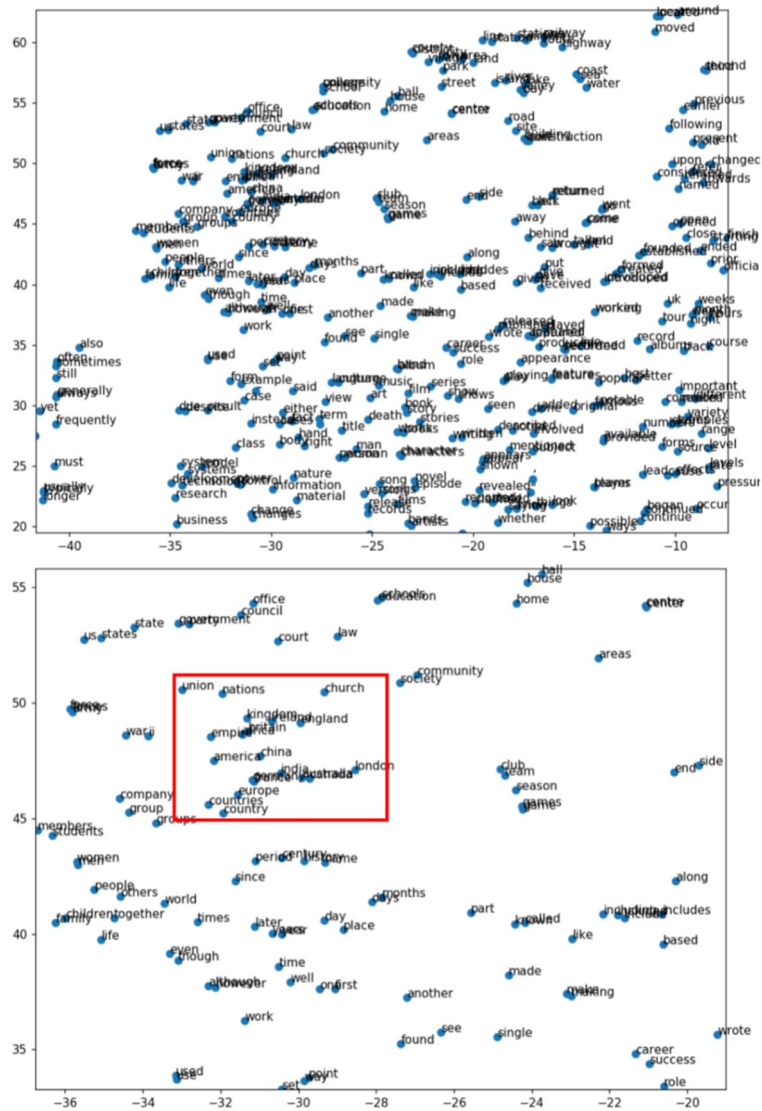
1: states; 2: new; 3: world; 4: years, time, year; 5: school sense^{[[SEP]]}
6: also; 7: known^{[[SEP]]}; 8: size, present, level, range; 9: family; 10: made
11: united, median, income, older, males, per, square, mile, density, housing, couples, racial, makeup, household, los, female, householder, households, alone, someone, native, residing, prime, races, hispanic, latino, total, capita, african, poverty, census, females, bureau, civil, spread, san, francisco, soviet, supreme, personal, hong, kong, demographics, de, la, pacific, islander, ago, roman, catholic, human, grand, prix, recent, representatives, fiction, geography, st, louis, diego, miles, orleans, middle, ages, least, takes, rico, coach, commons, royal, wide, male, labour, debut, box, formula, bowl, register, historic, primary, real, care, variety, regular, operating, baseball, bbc, kansas, harbor, directors, latin, short, listed, motion, picture, chief, executive, southeast, prize, peoples, jesus, gas, singles, divided, commercial, better, reached, formerly, red, computer, previous, performing, heritage, becoming, spent, directed, remain, almost, behind
12: American; 13: age, population, units, children, individuals, families, residents, women, americans, numbers; 14: first; 15: line, point

Affinity Propagation, 3000 words, 307 clusters, Pre-trained Embedding

1: median; 2: war, wars, occupation, invasion
3: years, months, weeks, century, seasons, decades, summers, decade
4: school, elementary, education, schools, academy, students, grade, student, grades, teacher, algebra
5: children, families, ages, daughters, fathers, child, childhood, boys, girls, childrens, sisters, adults
6: american, african, america, british, european, americans, indian, russian, canadian, german, japanese, chinese, australian
7: minister, ministers, ministry
8: city, town, district, airport, council, county, township, village, residents, metropolitan, borough, municipal, downtown, mayor, citys, boroughs, municipality
9: races, race, racing^{[[SEP]]}
10: males, female, females, male, gender, women^{[[SEP]]}
11: used, use, intended, derived, designed, using, employed, uses, applied, coined, interchangeably, distributed^{[[SEP]]}
12: two, three, four, five, several, six, seven, eight, nine, twin, consecutive, string, couple^{[[SEP]]}
13: north, south, northern, east, west, located, near, southeast, southern, western, eastern, unincorporated, northwest, southwest, northeast
14: best, great, better, greatest, highest, finest, fantastic, good, worst, fastest
15: even, though, times, become, make, much, like, eventually, still, ever, less, far, becoming, never, however, entirely, although, yet, things, always, little, slightly, easily, appear, enough, otherwise, none, lesser, get, afford,

perhaps, longer, unlike, exist, quite, makes, certainly, prove, expensive, naturally, thus, simply, else, without, readily, mean, something, anyone, probably, anything, bit, dickens, faster, particularly, suddenly, afterward

Figure 5 shows the visualization results for 3000 words using common words representations. Surprisingly, it can be seen that the space of common words representation, the semantically similar words are located together, such as numbers “one, two, three...”, and “country, countries, America, Europe...”. It indicates that the common words representation could capture the semantic information which can be used for other tasks.



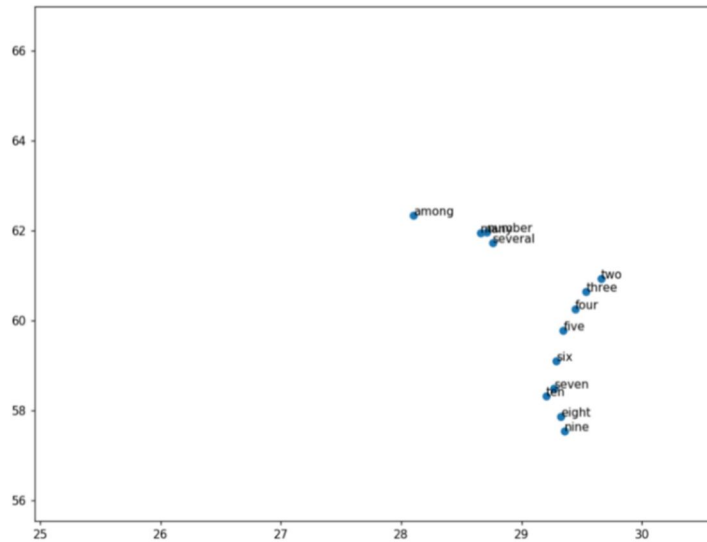


Fig 5. Visualization of the word representations by #common words

6. Conclusion

In this paper, we present two approaches to word sense discovery and topic modeling. This is a preliminary attempt to the latter by using clustering algorithms. For the problem of word sense discovery, we find that clustering the context word of the target word gives better results when manually specify the number of clusters. Because of “noise” words contained in the context words, the results are also being affected. The second problem we encountered in this experiment is that choosing the number of neighboring words shows great influences on the final performance of the model. For topic modeling, we conducted experiments for 20newsgroups data set and reported the coherence result for LDA and clustering-based approach. The results show that LDA outperforms the clustering-based approach in terms of coherence, and LDA also has the ability to identify how many topics are contained in the corpora. The problem for this task we found is that the clustering-based model should remove the noise words (non-topic words), then do the clustering. As future work, the problems found in this work could be solved to improve the model performance.

Acknowledgments

This research has been conducted within the framework of the grant num. BR05236839 “Development of information technologies and systems for stimulation of personality sustainable development as one of the bases of development of digital Kazakhstan”.

References

1. Bennani-Smires, K., Musat, C., Hossmann, A., Baeriswyl, M., Jaggi, M.: Simple unsupervised keyphrase extraction using sentence embeddings. In: CoNLL (2018)
2. Boudin, F., Mougard, H., Cram, D.: How document pre-processing affects keyphrase extraction performance. In: Proceedings of the 2nd Workshop on Noisy User-generated Text (WNUT). pp. 121–128. The COLING 2016 Organizing Committee, Osaka, Japan (December 2016)
3. Frey, B.J., Dueck, D.: Clustering by passing messages between data points. *Science* 315, 2007 (2007)
4. Mimno, D.; Wallach, H. M.; Talley, E.; Leenders, M.; and McCallum, A. 2011. Optimizing semantic coherence in topic models. In Proceedings of the Conference on

Empirical Methods in Natural Language Processing, 262–272. Association for Computational Linguistics.

5. Le, Minh and Postma, Marten and Urbani, Jacopo and Vossen, Piek. A Deep Dive into Word Sense Disambiguation with LSTM. Proceedings of the 27th International Conference on Computational Linguistics. ACL 2018. Santa Fe, New Mexico, USA.

6. Lu, Hengyang et al. “Don't Forget the Quantifiable Relationship between Words: Using Recurrent Neural Network for Short Text Topic Discovery.” AAAI (2017).

7. K. Nigam, J. Lafferty, and A. McCallum. Using maximum entropy for text classification. IJCAI-99 Workshop on Machine Learning for Information Filtering, pages 61–67, 1999. ^[1]_{SEP}

8. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. *J. Mach. Learn. Res.* **3**, 993–1022 (Mar 2003)

9. Toleu, Alymzhan and Tolegen, Gulmira and Makazhanov, Aibek. Character-Aware Neural Morphological Disambiguation. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017, Vancouver, Canada.

10. Yoshua Bengio, Rejean Ducharme, Pascal Vincent. A Neural probabilistic language model. *Journal of Machine Learning Research*. Volume 3, Pages 1137-1155.

11. T. Hofmann. Probabilistic latent semantic indexing. Proceedings of the Twenty-Second Annual International SIGIR Conference, 1999.

12. M. Jordan, Z. Ghahramani, T. Jaakkola, and L. Saul. Introduction to variational methods for graphical models. *Machine Learning*, 37:183–233, 1999.

13. Kenneth Ward Church and Patrick Hanks (March 1990). "Word association norms, mutual information, and lexicography". *Comput. Linguist.* 16 (1): 22–29.

14. Lloyd, S. P. (1982). "Least squares quantization in PCM". *IEEE Transactions on Information Theory*. 28 (2): 129–137. CiteSeerX 10.1.1.131.1338.

15. Brendan J. Frey; Delbert Dueck (2007). "Clustering by passing messages between data points". *Science*. 315 (5814): 972–976. CiteSeerX 10.1.1.121.3145. doi:10.1126/science.1136800. PMID 17218491.

16. L.J.P. van der Maaten and G.E. Hinton. Visualizing High-Dimensional Data Using t-SNE. *Journal of Machine Learning Research* 9(Nov):2579-2605, 2008.

17. S. Deerwester, S. Dumais, T. Landauer, G. Furnas, and R. Harshman. Indexing by latent semantic analysis. *Journal of the American Society of Information Science*, 41(6):391–407, 1990.

18. C. Papadimitriou, H. Tamaki, P. Raghavan, and S. Vempala. Latent semantic indexing: A probabilistic analysis. pages 159–168, 1998.

19. Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'13), C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger (Eds.), Vol. 2. Curran Associates Inc., USA, 3111-3119.

20. Ed Collins, Isabelle Augenstein, Sebastian Riedel. A Supervised Approach to Extractive Summarisation of Scientific Papers. Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017). Association for Computational Linguistics 2017, Vancouver, Canada.

ҚАЗАҚСТАННЫҢ ЭЛЕКТР ЭНЕРГЕТИКАЛЫҚ ЖҮЙЕСІ ҮШІН МАТЕМАТИКАЛЫҚ МОДЕЛЬ ҚҰРУ

Калимолдаев М.Н., Абдилдаева А.А., Жукабаева Т.К.,
Галиева Ф.М.

E-mail: mnk@ipic.kz, abass_81@mail.ru, tamara_kokenovna@mail.ru,
galyeva.faryza@gmail.com

ҚР БҒМ ҒК Ақпараттық және есептеуіш технологиялар институты
Алматы, Қазақстан

Аңдатпа. Соңғы жылдары энергияға сұраныстың артуы жағдайында интеллектуалдық электроэнергетикалық жүйелер әлемдік энергетиканың болашақ міндеттерін шешу үшін ортақ таңдау болды. Бұл мақалада Қазақстанның энергетикалық жүйесінің қазіргі жағдайы берілген. Қазақстанда электр энергиясын өндіру мен тұтыну зерттелді, электр энергиясын тұрақтандырудың математикалық моделі қарастырылды. Осы мәселенің сандық шешімі Адамас-Баифорд көп қадамды әдістерімен алынды және графикалық нәтижелер көрсетілді.

Кіріспе

Энергетика бүкіл әлемде бірнеше түрге бөлінеді: жылу, ядролық, гидроэнергетика және соңғы уақытта дамып келе жатқан жаңартылатын энергия көздері, жел энергетикасы және гелиоэнергетика. АӨҚО рейтинг агенттігінің деректері бойынша 2005-2015 жылдар кезеңінде әлем бойынша энергия тұтынудың үлесі: мұнай-32,9, табиғи газ - 23,8, көмір - 29,2, ядролық энергия - 4,4, гидроэнергия - 6,8, жаңартылатын көздер - 2,8 құрады. (1-сурет) [1].

Қазақстан энергетикалық ресурстардың ірі қорына (мұнай, газ, көмір, уран) ие және энергия тасығыштардың табиғи қорларын сату есебінен тұратын шикізат елі болып табылады (экспорттың 80% - ы-шикізат, ал өнеркәсіптік экспорттың үлесі жыл сайын қысқарады).

АӨҚО рейтинг агенттігінің деректері бойынша 2005-2015 жылдар кезеңінде Қазақстанда энергия тұтыну үлесі: мұнай - 23,2, табиғи газ - 14,1, көмір - 59,4, гидроэнергия - 3,3, жаңартылатын көздер - 0,02 құрады (2- сурет) [1].



1-сурет. 2005-2015 жж. энергия тасымалдаушы түрі бойынша әлемдегі энергия тұтынудың құрылымы



2-сурет. 2005-2015 жж. Қазақстан Республикасында энергия тасығыштың түрі бойынша энергия тұтынудың құрылымы

Қазақстан Республикасының Бірыңғай электр энергетикалық жүйесі (ҚР БЭЖ) республиканың тұтынушыларын сенімді және сапалы энергиямен жабдықтауды қамтамасыз ететін электр станцияларының, электр беру желілерінің және қосалқы станциялардың жиынтығын білдіреді. Қазақстанда электр энергиясын өндіруді әртүрлі меншік нысанындағы 138 электр станциясы жүзеге асырады.

Қазақстанның барлық электр станцияларының белгіленген жиынтық қуаты 20 мың МВт, ал нақты қуаты — 15 мың МВт құрайды (2013 жылғы дерек бойынша). Электр станцияларының түрі бойынша өндіру келесідей бөлінеді:

- ЖЭС (жылу электр станциялары) — 87,7 %, оның ішінде:
- КЭС (конденсациялық электр станциялары) — 48,9 %;
- ЖЭО (жылу электр орталықтары) — 36,6 %;
- ГТЭС (газтурбиналық электр станциялары) — 2,3 %;
- ГЭС (су электр станциялары) — 12,3 %.

Қазақстанда электр энергиясының 72% — ға жуығы көмірден, 12,3% — ы — гидроресурстардан, 10,6% - ы-газдан және 4,9% - ы-мұнайдан өндіріледі. Осылайша, электр станцияларының негізгі төрт түрі электр энергиясының 99,8% өндіріледі, ал баламалы көздерге 0,2% - дан кем келеді [2].

01.01.2019 ж.жағдай бойынша Қазақстан электр станцияларының жалпы белгіленген қуаты 21 901,9 МВт, қолда бар қуаты — 18 894,9 МВт құрайды[3].

Келесі 1-кестеде электр энергиясын тұтыну мен өндіру және Қазақстанда соңғы жылдардағы өндірудегі ЖЭК үлесі көрсетілген.

1-кесте. Қазақстандағы электроэнергияны тұтыну және өндіру бойынша статистика. 2014-2018 жылдар

Жыл	Өндіру млрд. кВтч	Тұтыну млрд. кВтч	ЖЭК млрд. кВтч
2018	106,797	103,228	1,352
2017	102,384	97,857	1,100
2016	93,9	92,1	0,928
2015	90,8	90,9	0,704
2014	93,9	90,4	0,580

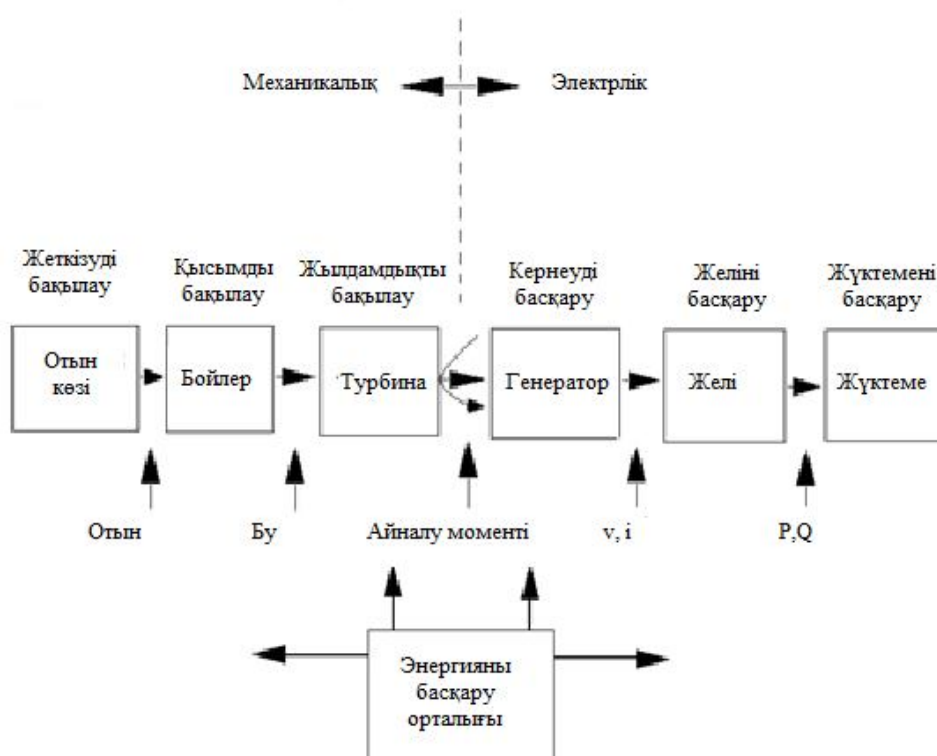
Деректер бойынша Қазақстанның электр энергиясын өндіруде ЖЭК пайдалану үлесі шамалы екенін көруге болады, өйткені бұл сала жақында дамып келеді.

Қазақстан Республикасының электр энергетикалық жүйесі бұрынғы КСРО-дан мұра болды және қазіргі уақытта жабдықтарды жаңғырту және электр энергиясын беру мен бөлудің жаңа объектілерін салу қажеттілігі туындайды. Қазақстанның электр энергетикасын дамытудың қазіргі заманғы жағдайында тозудың жоғары деңгейі және жабдықты жаңғыртудың жеткіліксіз жоғары қарқыны кезінде бұл мәселе одан да өткір тұр. Қазақстанның электр энергетикасында жаңғырту қажеттілігінен басқа, қазіргі заманғы жабдықты қолдану тиімділігін және оның қалыпты жұмыс жағдайын бағалау талап етіледі.

Басты бөлім

Ғылым мен техниканың әр түрлі салаларындағы тепе-тең жағдайларды тұрақтылығы мен орнықтылығы мәселесінің маңыздылығы жалпыға белгілі. Ол электр энергетикалық жүйелер үшін де үлкен мәнге ие. Бұл мәселені сенімді шешусіз барлық салалардың тұтынушыларын электр энергиясымен сенімді және сапалы жабдықтау мүмкін емес.

3-суретте электр станциясындағы электр энергиясын өндіруден бастап жүктемені тұтынуға дейінгі динамикалық электр энергетикалық жүйеге жалпы шолу көрсетілген.



3-сурет. Электр энергетикалық жүйенің динамикалық құрылымы [4]

Жүктемені басқару

Бұл блок-схема отын көзінен басталады, содан кейін бумен қамтамасыз ету үшін жағылады. Бұл бу генераторды іске қосатын турбинаның айналуына әкеледі. Электр станцияларында орнатылған генераторлар негізінен синхронды машиналар болып табылады. Синхронды машина турбинамен қамтамасыз етілетін

механикалық қуатты электр қуатына түрлендіреді, содан кейін жүктемемен желі бойынша бөлінеді. Жүктеме кезінде қуат тұтынылады және бұл қуат тарату блок-схемасын аяқтайды. Жүйенің барлық компоненттері сенімді жұмыс үшін энергияны басқару орталығымен бақыланады.

Сызықты емес дифференциалдық теңдеулер жүйесімен сипатталатын көп өлшемді күрделі электр энергетикалық жүйелердің жалпы математикалық моделін қарастырайық:

$$\begin{aligned} \frac{d\delta_i}{dt} &= S_i, \\ \frac{dS_i}{dt} &= w_i - D_i S_i - f_i(\delta_i) - \psi_i(\delta_i^*), \quad w_i = C_i^* x_i, \\ \frac{dx_i}{dt} &= A_i x_i + q_i S_i + b_i u_i, \quad i = \overline{1, l}, \end{aligned} \quad (1)$$

мұндағы

$$\psi_i(\delta_i^*) = \sum_{\substack{k=1, \\ k \neq i}}^l P_{ik}(\delta_{ik}), \quad \delta_{ik} = \delta_i - \delta_k \quad (2)$$

Функция $\psi_i(\delta_i^*)$ генераторлардың өзара әрекеттесуін сипаттайды және келесі түрде ашылады:

$$\psi_i(\delta_i^*) = \sum_{\substack{j=1, \\ j \neq i}}^l \frac{1}{T_i} [P_{ij} \sin(\delta_{j0} + \delta_{ji}) - P_{ij} \sin \delta_{j0}] \quad (3)$$

Периодтық үздіксіз дифференциалданатын функция $f_i(\delta_i)$ келесі формула бойынша анықталады:

$$f_i(\delta_i) = \frac{1}{T_i} [P_i \sin(\delta_{i0} + \delta_i) - P_i \sin \delta_{i0}], \quad i = \overline{1, l}, \quad (4)$$

(1) жүйеде δ_i - кейбір синхронды айналу осіне қатысты i -генератор роторының айналу бұрышы; S_i - i -генератордың сырғуы; $D_i > 0$ - демпфирлеу коэффициенті.

Мақалада сондай-ақ келесі дифференциалды теңдеудің жүйесімен сипатталатын синхронды жұмыс істейтін генераторлар n үшін электр энергетикалық жүйенің қозғалысын тұрақтандырудың математикалық моделі қарастырылған:

$$\begin{aligned} \frac{d\delta_i}{dt} &= S_i, \\ \frac{dS_i}{dt} &= c_i x_i - K_i S_i - f_i(\delta_i) - P, \\ \frac{dx_i}{dt} &= A_i x_i + u_i \end{aligned} \quad (5)$$

мұнда бірінші 2 теңдеу жүйенің жұмысын сипаттайды, ал үшінші теңдеу реттеуіштің күйін сипаттайды. Сондай-ақ x_i – фазалық айнымалылар; c_i – скалярлар; u_i – басқару; δ – ротордың айналу бұрышы; S – генератордың сырғуы.

$$f_i(\delta_i) = \frac{1}{T} [P_i \sin(\delta_{i0} + \delta_i) - P_i \sin \delta_{i0}]$$

$$P_{ij} = \sum_{i \neq j, j=1, i=1}^n \frac{1}{T_i} (P_{oj}(\sin(\delta_{ji} - \delta_{ii}) + \delta F_{ji} - \sin(\delta F_{ij}))$$

Электр энергетикалық жүйелерді тұрақтандыру міндетін сандық іске асыру үшін осы міндеттерді неғұрлым дәл шешуді беретін келесідей алгоритм пайдаланылатын болады.

2-кесте. Адамс-Башфорт әдісі үшін қателік коэффициенттері мен тұрақтылары [5]

k	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	C
1	1								$\frac{1}{2}$
2	$\frac{3}{2}$	$-\frac{1}{2}$							$\frac{5}{12}$
3	$\frac{23}{12}$	$-\frac{4}{3}$	$\frac{5}{12}$						$-\frac{3}{8}$
4	$\frac{55}{24}$	$-\frac{59}{24}$	$\frac{37}{24}$	$-\frac{3}{8}$					$\frac{251}{720}$
5	$\frac{1901}{720}$	$-\frac{1387}{360}$	$\frac{109}{30}$	$-\frac{637}{360}$	$\frac{251}{720}$				$-\frac{95}{288}$
6	$\frac{4277}{1440}$	$-\frac{2641}{480}$	$\frac{4991}{720}$	$-\frac{3649}{720}$	$\frac{959}{480}$	$-\frac{95}{288}$			$\frac{19087}{60480}$
7	$\frac{198721}{60480}$	$-\frac{18637}{2520}$	$\frac{235183}{20160}$	$-\frac{10754}{945}$	$\frac{135713}{20160}$	$-\frac{5603}{2520}$	$\frac{19087}{60480}$		$-\frac{5257}{17280}$
8	$\frac{16083}{4480}$	$-\frac{1152169}{120960}$	$\frac{242653}{13440}$	$-\frac{296053}{13440}$	$\frac{2102243}{120960}$	$-\frac{115747}{13440}$	$\frac{32863}{13440}$	$-\frac{5257}{17280}$	$\frac{1070017}{3628800}$

1-алгоритм. Адамс-Башфорттың көп қадамды әдісі.

1-қадам. $t = 0$ б олғанда δ, S, x бастапқы мәндер және c, K, P, A белгілі.

2-қадам. δ, S, x табу барысында 1,2,3,4 нүктелері үшін 4-ші ретті Рунге-Кутта әдісі қолданылады.

$$\delta_{i+1} = \delta_i + \frac{1}{6}(m_1 + 2m_2 + 2m_3 + m_4)$$

$$m_1 = \Delta t(S_i)$$

$$m_2 = \Delta t\left(S_i + \frac{1}{2}m_1\right)$$

$$m_3 = \Delta t\left(S_i + \frac{1}{2}m_2\right)$$

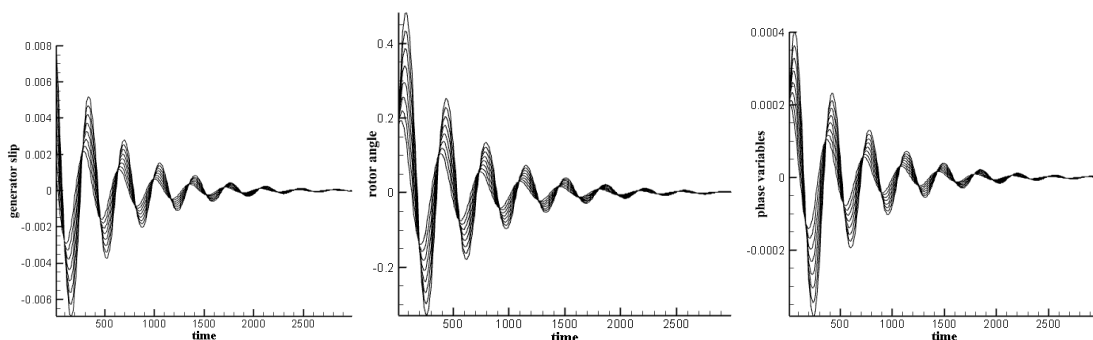
$$\begin{aligned}
m_4 &= \Delta t \left(S_i + \frac{1}{2} m_3 \right) \\
S_{i+1} &= S_i + \frac{1}{6} (m_1 + 2m_2 + 2m_3 + m_4) \\
m_1 &= \Delta t ((cx_i - KS_i - f_i(\delta_i) - P_{12})) \\
m_2 &= \Delta t \left((c(x_i + \frac{1}{2} m_1) - K(S_i + \frac{1}{2} m_1) - f_i(\delta_i + \frac{1}{2} m_1) - P_{12}) \right) \\
m_3 &= \Delta t \left((c(x_i + \frac{1}{2} m_2) - K(S_i + \frac{1}{2} m_2) - f_i(\delta_i + \frac{1}{2} m_2) - P_{12}) \right) \\
m_4 &= \Delta t \left((c(x_i + \frac{1}{2} m_3) - K(S_i + \frac{1}{2} m_3) - f_i(\delta_i + \frac{1}{2} m_3) - P_{12}) \right) \\
x_{i+1} &= x_i + \frac{1}{6} (m_1 + 2m_2 + 2m_3 + m_4) \\
m_1 &= \Delta t (Ax_i + u) \\
m_2 &= \Delta t \left(A(x_i + \frac{1}{2} m_1) + u \right) \\
m_3 &= \Delta t \left(A(x_i + \frac{1}{2} m_2) + u \right) \\
m_4 &= \Delta t \left(A(x_i + \frac{1}{2} m_3) + u \right)
\end{aligned}$$

3-қадам. $i=4$ нүктесінен бастап $i=n$ дейін δ, S, x Адамс-Башфорд көп қадамды әдісімен табамыз. Адамс-Башфорд әдісінің коэффициенті 2-кестеде көрсетілген.

$$\begin{aligned}
\delta_{i+1} &= \delta_i + \Delta t \left(\frac{1901}{720} S_i - \frac{1387}{360} S_{i-1} + \frac{109}{30} S_{i-2} - \frac{637}{360} S_{i-3} + \frac{251}{720} S_{i-4} \right) \\
S_{i+1} &= S_i + \Delta t \left(\frac{1901}{720} (cx_i - KS_i - f_i(\delta_i) - P_{12}) - \frac{1387}{360} (cx_{i-1} - KS_{i-1} - f_{i-1}(\delta_{i-1}) - P_{12}) + \right. \\
&\quad \left. + \frac{109}{30} (cx_{i-2} - KS_{i-2} - f_{i-2}(\delta_{i-2}) - P_{12}) - \frac{637}{360} (cx_{i-3} - KS_{i-3} - f_{i-3}(\delta_{i-3}) - P_{12}) + \right. \\
&\quad \left. + \frac{251}{720} (cx_{i-4} - KS_{i-4} - f_{i-4}(\delta_{i-4}) - P_{12}); \right) \\
x_{i+1} &= x_i + \Delta t \left(\frac{1901}{720} (Ax_i + u) - \frac{1387}{360} (Ax_{i-1} + u) + \frac{109}{30} (Ax_{i-2} + u) - \frac{637}{360} (Ax_{i-3} + u) + \frac{251}{720} (Ax_{i-4} + u); \right)
\end{aligned}$$

Қорытынды

Бұл мақалада Қазақстанның электр энергетикалық жүйесінің қазіргі жағдайы берілген. Қазақстанда электр энергиясын өндіру мен тұтыну зерттелді, сондай-ақ электр энергетикалық жүйені тұрақтандырудың математикалық моделі қарастырылды. Бұл есептің сандық шешімі Адамс-Башфорд көпқадамды әдісін пайдалана отырып алынған және графикалық нәтижелер көрсетілген.



4-сурет. Сандық нәтижелер

Әдебиеттер тізімі

- 1 Шейкин Д. Энергетика Республики Казахстан <http://www.rfcaratings.kz/>
АО «Рейтинговое Агентство РФЦА» 2016
- 2 Данные Агентства Республики Казахстан по статистике
- 3 Электроэнергетика Казахстана: ключевые факты. Электронный ресурс:
www.kegoc.kz/ru/elektroenergetika/elektroenergetika-kazahstana-klyuchevye-fakty
- 4 P. W. Sauer and M. Pai, Power System Dynamics and Stability. Stipes Publishing L.L.C., 2006.
- 5 J. C. Butcher. Numerical Methods for Ordinary Differential Equations / Second Edition. The University of Auckland, New Zealand., 2008.

УДК 004.934

МАШИНАЛЫҚ ОҚУДЫ ҚОЛДАНУДА ДАУЫСТЫҢ ГЕНДЕРЛІК ЖІКТЕЛІНУІ

**Қалимолдаев М.Н., Мамырбаев О.Ж., Мекебаев Н.О.,
Тұрдалыұлы М.**

e-mail: nurbapa@mail.ru

*ҚР БҒМ ҒК Ақпараттық және есептеуіш технологиялар институты,
Алматы қ., Қазақстан*

Аңдатпа. Гендерлік анықтау бүгіндері сөйлеуді сараптау ісінің негізгі проблемеларының бірі болып табылды. Мысалы, акустикалық берілгендер бойынша яғни жоғарылақ, медианалық, жиілік т.б. – бойынша жынысты айқындау. Машиналық оқудың жақсы нәтижелер беруі, зерттеудің барлық салаларында жіктеу есептері үшін нәтижелерді, PCA – ны бағалау алгоритм үшін біздің салыстырмалы моделіміз. Кез-келген алгоритмді бағалауда негізгі параметрлер болып, оның өнімділігі саналады. Жіктеу деңгейі жіктеу есептерінде төмен аз болуы тиіс, бұның өзі дәлірек деңгейінің жоғары болуы керектігін білдіреді. Берілген мақалада сөйлеушіні тану жүйесінде танымал құрал ұсынылған, ол сараптаудың (PCA) негізгі мүшесі болып табылады.

Кілттік сөздер: Сөйлеуді тану, дауыс, гендерлік анықтау, PCA, SVM.

I. Кіріспе

Диморфизм – дауыс белгісі, оның адамдарда байқалуы жоғары. Дауыс екпіні, сөйлеу жылдамдығы және ұзақтығы – бұлар адам дауысын, негізінен ер және әйел дауыстарын ажырататын белгілі бір сипаттар [1]. Қабылданатын диморфизм барлық дауыстың 98,8 % құрайды, ол сөйлеушінің жынысынан және тиісті жиіліктен тұрады. Алайда әуезді сөйлеу белгілері жыныстың өзгеруін алдын ала анықтай алмайды. Ер адам мен әйел адам дауысының арасындағы айырмашылықтың жоғарылығы өзгермелі болуы мүмкін, сондықтан сөйлеушінің ер адам ба әлде әйел адам екенінен болжау қиын. R тілі көмегімен ақиқат уақыт ішінде сөйлеуді өңдеу үшін қолданылатын әдістер арқылы керекті сөйлеушінің жынысын анықтай аламыз. Әуездік дауыс қабатының қалыңдығы жыныстық айырманың болу себебі болып табылады және түрлі жыныстар үшін өзара өлшенген болуы ықтимал. Қазіргі осы шақтағы физикалық жағдаяттар да сөйлеуші сөйлеп тұрған себепші стиль тағы бір себеп болып табылады. Бұл ауауытқуларды сөйлеушінің ер немесе әйел адам дауыстарының бір-бірінен айырмасын зерттеуге арналған болатын. Бұларға әр түрлі параметрлер жатады. Сөйлеуді сараптау үшін жиілік пен қадам негізгі параметр болып есептеледі, ол ол жіктеуді де тануға әкеледі. Сөйлеуді тану жыныс туралы, оның жасы туралы және сөйлеп тұрған диалектісі туралы ақпарат алуға көмектеседі. Бұл сала бойынша үлкен жұмыс атқарады. Белгілі бір сөйлеу статистикасы қолданылады. Ол уақытында қолданылады және жынысты анықтау үшін ең шекті мәндерді де қолданады.

Дауыстың мәліметтер жиыны дауыс күші, дыбыс жоғарылығы, жиілігі, q21, q25 т.б сияқты түрлі параметрлерге айналады, содан соң олар оқылады және тестіленеді, мұнда жоғарыда көрсетілгендей алгоритмдер бойынша жынысты болжап тану үшін әр түрлі алгоритмдер бірге қолданылады.

Берілген мақалада біз болжаудың түрлі алгоритмдерінің негізінде жынысты жіктейтін салыстырмалы модельді ұсынамыз. Болжау негізінен мағына-мәндері сөйлеуден өңделіп алынатын мәліметтер салыстырылады ол белгілі бір параметрлер бойынша жынысты белгілеріне қарай жіктеуде ең жақсы өнімділікке жеткізетін алгоритм туралы қорытынды жасауға мүмкіндік береді. Жынысты анықтауда салыстырмалы модельдің алгоритмі қаншалықты қаншалықты дәл болжамдайтыны осы алгоритмдер негізінде табылады.

II. Алгоритм

Сөйлеуді жіктеу мен өңдеу және гендерлік тану мен жіктеу ұзақ уақыт бойында пайдаланып келеді. Уақыт өте өңделген кейбір тұжырымдамаларды гендерлік мойындауда жүзеге асыруда пайдаландық. Жынысты анықтауға негізделген таяудағы зерттеулер әр түрлі параметрлерге айналатын көрсетеді. Негізгі параметрлер жоғарылық пен жиілікті қамтиды. Жіктеу ер, әйел адамдарды анықтау үшін жүргізіледі. Ол үшін алдымен жүйе оқу мәліметтерімен және тестілеумен бірге оқытылады, сөйтіп бұл берілгендер үшін жүйенің өнімділігі үшін мәліметтер енгізіледі және бағаланады. Әр түрлі алгоритмдер үшін әр түрлі алынған жетістіктер мен нәтижелер әр түрлі уақытта әр түрлі нәтиже беретін сияқты.

Ересек адамдардың сөйлеуін тануда олар қолма-қол түзету және вокалдык ұзақтығы түзетуге қабілетті екенін көрсетеді, сөйтіп олар ер және әйел болып дыбысталады. Осылайша ер және әйел дауыстарын жіктеу қиындық келтіреді. Кейбір әйел дауыстарын тон жоғарлығы [4] негізінде сараптау қиындығына соғады. Мұнда әйел дауыстарының бір қырын ғана зерттеу біздің талаптарға сәйкес емес. Ер және әйел арасындағы мұндай қадамдар әйел дауысын параметрлердің басқа

әдістері арқылы анықталынуы керектігін аңғартады. Бұл әртүрлі параметрлермен іске асырылады. Бұларда әйелдер пайдаланушы болып саналады. Ал ол әйелден әйелге беріліп отыруға тәуелді, демек мәліметтер жиыны осы негізде ер мен әйелді жіктеуге дейін өңделіп қоюы керек.

Қадамды қабылдауға сәйкес [5] негізгі жиілік [F0] сөйлеушінің тілдік, паралингвистикалық және лингвистикалық емес ақпараттарының комбинациясын түзеді. Бұл үшеуі ер мен әйел дауыстарына сәйкес келеді, сондай-ақ олар жоғары тонмен сөйлеуші тонына тәуелді. Бұл жиілікті [F0] қашықтық жұмыс тәжірибесінсіз, буындық сыртқа ақпаратсыз-ақ ораналастыру сәті түсті. Оның өзі сөйлеушінің даусы жоғары және төмен жиіліктер мен динамикалар арасында құбылып отыртынын көрсетеді.

SVM [5] көмегімен жынысты анықтау жинақтап сөйлеу, телефонмен сөйлеу, сөйлеу тілдердегі айырмашылықтар т.б. сияқты әр түрлі сөйлеу тетіктерін сараптау мәселе гендер туралы екенін көрсетеді. Бұның өзі ер дауысы тон жоғарлығынан, кезеңнен Мел – жиілік аралығы шамамен 100-146 Гц диапазонда, ал әйелдердікі 188-221 Гц диапазонда орналасатынын білдіреді. Мұнда дауыс жиілік негізінде бөлініп шығады және ол зерттеліп алынған әрі сарапталынған болып табылады.

Тірек векторлар машинасы (SVM). Бақыланатын оқыту – маркерленген оқытудағы функцияларды анықтау үшін машиналық оқыту есебінің берілгендері, оны жіктеу үшін де, регрессиялау үшін де қолдануға болады. Тірек векторлық алгоритмі болып есептеледі. Тірек векторлар гипер жазықтықтар бар кіріккен аралас нүктелердің берілгендері. Мұнда егер берілгендер жиыны қашықтаса, онда ол бөлінген гипер жазықтықтың қалпын өзгертеді. SVM – да мысалдар жиынының әрбір қосағының кіретін нысаны болады және оның артықшылықтары бар мәні де болады. Бақыланатын оқыту алгоритмі берілгендерді сараптайды және нәтижесі функцияны шығарады, ол жаңа нәтижелерді бейнелеуге қол жеткізеді.

SVM – жіктеу үшін, аппроксимация үшін т.б. қолданылатын машиналық оқыту құралы. Ол жалпыланған жинақты түрде де қолданылуы мүмкін, бұның өзі көптеген салаларда жетістіктерге жеткізеді. SVM метрикасы минимумға дейін шектеледі және гипер жазықтық пен берілгендер жиынын бөліп тұратын шекараларды максимумдауда жоғарғы шектерегі кемістіктерді жинақтап жалпылайды. Таңдаудың артықшылықтары ұтымды оңтайлы саны бойынша да, функциялардың орналасуы бойынша да оқыту барысында автоматты жүзеге асады.

III. PCA-ны қолана отырып функцияларды шығару

PCA-ның басты компоненттерін сараптау берілгендер кестесін сараптайтын көпқырлы әдіс болып табылады. Мұнда өзара байланысқан бірнеше өзгермелі сандық тәуелділіктер сипатталады. Оның мақсаты – кестеден маңызды ақпарат алу, оны негізгі компоненттер деп аталатын жаңа ортгондық өзгермелілер жиыны ретінде беріледі.

PCA – бұл өзгермелілердің қысқару рәсімі. Бұл жаңадан алынған өзгермелілер қатары бойынша мәліметтерден көп өзгермелілер болған жағдайда тиімді. Сөйтіп өзгермелілердік көп саны мүмкіндікті болып табылады және мұндай өзгермелілерде қайсы бір артықшылықар болады деп саналады. Мұнда артықшылық кейбір өзгермелілер бір-бірімен корреляцияға түсетінін білдіреді, бұл олардың бірдей құрылғыларды өлшеуіне байланысты болуы мүмкін. Осындай артықшылықтан байқалатын өзгермелерді негізгі компоненттердің ең аз санына кеміту мүмкін, олар бақыланатын өзгермелердің дисперсиясының басым бөлігін ескеретін болды [4].

1-ші суретте көрсетілген PCA-ны табу алгоритмі сигналды орталықтандыруды білдіреді. Ең басты және алдын ала қажетті өңдеу – сигналды орталықтандыру, яғни оның $m = \bar{x}$, орта векторын есептеп шығару, ол x -тік нөлдік орташасы бар өзгермеліні алу үшін керек. Мысалы $\{x_1, x_2, x_3, \dots, x_N\}$ мәнді берілгендерді N вектордың белгілері деп алғанда, мұнда дауыс жиілік негізінде анықталады, бұл алынған және сарапталынған болады. Белгілер векторларының орташа мәні \bar{x} берілген және енгізілген сигналмен есептеледі.

$$\hat{x} = x_i - \bar{x}$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$



Сурет1. PCA алгоритмінің блок схемасы

IV. Нәтижелер

Қолданылатын берілгендер жиыны – жынысына қарай дауыстық берілгендердің жиыны 4120 жазу қамтылған, оның 2547-ерлер және 1550 әйелдер дауысы, ыңғайлы болу үшін ер және әйел жыныстық лейбтер 0 және 1 аралығында конверттеленеді.

Мұнда 1-суретте төрт түрлі динамик сигналдары деректер қоры және бір сынақ сигналы ретінде пайдаланылады. Графиктер негізгі компоненттерді талдау кезінде сигналдарды көрсетеді.

Әрбір жазу орта жиілік, және т.б.с.с. Әрбір динамика үшін максималдық сәйкестікті таңдау мақсатында біз корреляция коэффициентін пайдаланадық. Корреляция тәуелділікпен байланысты статистикалық қатынастардың мейлінше

кең класының ішінен кез-келгеніне қатысты болып табылады. Корреляция дәрежесінің өлшем шамасы «корреляция коэффициенті» деп аталады. 1 кестеде әрбір динамикада берілгендер базасы бар енетін белгісіз дауыстың корреляциялық коэффициенті көрсетілген, мұнда енетін дауыс сигналын тек бір ғана динамикпен төрт дүркін беріледі.

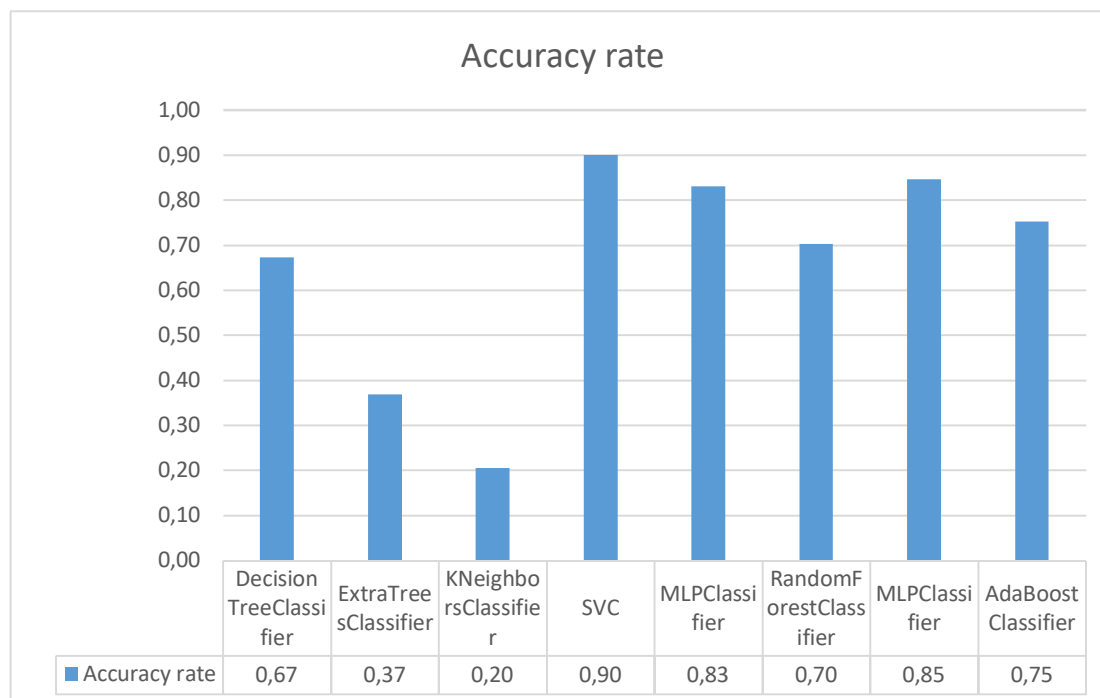
Жүргізілген эксперименттер нәтижесінде егер белгісіз енетін дауыс пен берілгендер базасында алдын ала жүктелген кез-келген сөйлеу сигналының арасында сәйкестік болатын болса, онда нәтижелер корреляция коэффициентінің жоғарғы мәнін бейнелейтіні және бұл нақты сөйлеу сигналы өз кезегінде, белгілі бір динамикті тексере алатыны анықталады.

Эксперименттер PCA алгоритмдерінің өнімділігін өлшеу үшін жүйеге нақты уақытта ендіру арқылы шуылды ортада жүргізіледі. Ол Python-ды қолдану арқылы жүзеге асты. 1 кестеде PCA көмегімен динамикті тану өнімділігінің артқанын көре аламыз. Бұл PCA мейлінше сенімді екенін көрсеттік.

Кесте 1. PCA көмегімен алынған нәтиже

Әдіс	Сөйлеуші	Дәлдігі (%)	Бірлік уақыты
PCA	1	75,00	3,63
	2	76,05	3,89
	3	78,02	4,15
	4	79,00	3,77

2-ші суретте зерттеу барысында қолданылған алгоритмдер нәтижелер дәлдігі және деректер жиынындағы жіктеу дәлдігі.

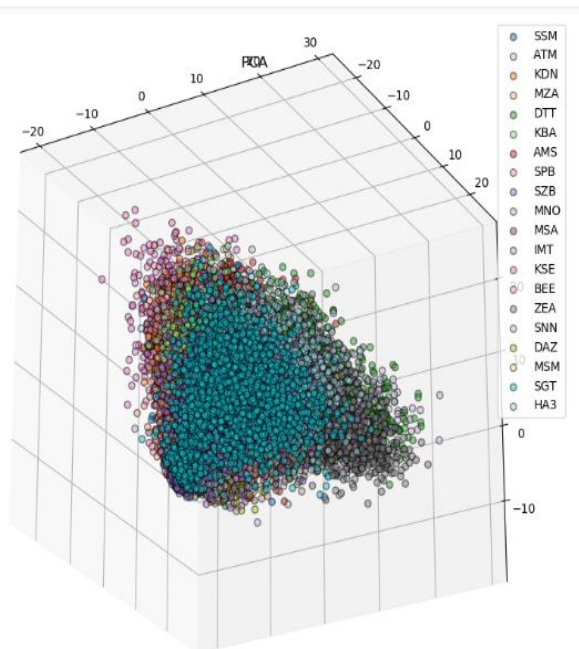


Сурет 2. Алгоритмдер және деректер жиынындағы жіктеу дәлдігі

Диаграммадан көрініп тұрғандай, тірек векторлар әдісі және көп қабатты перцептрон тиісінше 0,90 және 0,83 ең жақсы нәтижелер көрсетті.

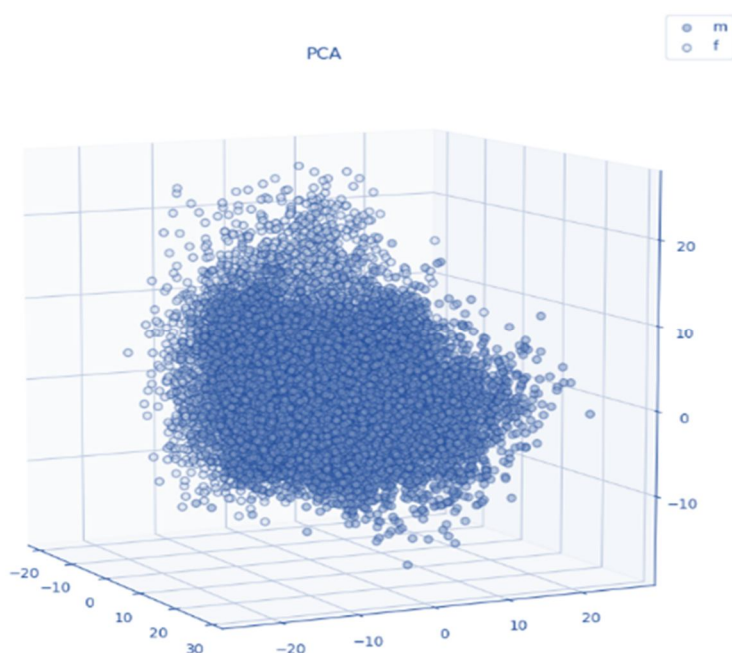
Зерттеу барысында бас компонент әдісімен 1 479 белгіге дейін өлшемдігін төмендетумен деректерде жіктеудің дәлдігі алынды.

3-ші суретте жоғарыдағы алгоритмдерді қолданғандағы 20 адамның даусын анықтауды көрсеткен нәтиже.



Сурет 3. 20 адам дауысын анықтаудағы нәтиже.

4-ші суретте алынған нәтижеде ер мен әйелдің дауыстарын анықтау көрсетілген. Ақшыл түстегі әйел дауысы, ал қою көк түстегі ер дауысы.



Сурет 4. PCA-ны пайдаланғандағы ер,әйел дауысын анықтау.

Қорытынды

Берілген мақалада микрофон көмегімен нақы бір уақытта жиналған сапасы төмен әр түрлі сөйлеу сигналдарынан сапалы сөйлеу сигналдарын ажырату үшін PCA негізіндегі алгоритм пайдаланды. PCA-ны қайта құруды пайдалан отырып нақты бір уақытта тану үшін программалық алгоритм әзірленді.

Мақала ЖТН – AP05131207 «Терең нейрондық желілерді пайдаланатын мультитілдік автоматты сөйлеуді тану технологиясын құру» жоба негізінде дайындалады.

Әдебиеттер

1. Ali, Md Sadek, Md Shariful Islam, and Md Alamgir Hossain. "Gender recognition system using speech signal." *International Journal of Computer Science, Engineering and Information Technology (IJCEIT)*, Vol. 2, No. 1, 2012, pp. 1-9.
2. Alías, Francesc, Joan Claudi Socoró, and Xavier Sevillano. "A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds." *Applied Sciences* 6, No. 5, 2016, pp. 143.
3. Subramanian, Hariharan, P. Rao, and S. D. Roy. "Audio signal classification." *EE Dept, IIT Bombay*, 2004, pp. 1-5.
4. Bach, Jörg-Hendrik, Jörn Anemüller, and Birger Kollmeier. "Robust speech detection in real acoustic backgrounds with perceptually motivated features." *Speech Communication*, Vol. 53, No. 5, 2011, pp. 690-706.
5. Campbell, J. P., "Speaker recognition: a tutorial" *Proceedings of IEEE*, vol. 85, no. 9, pp. 1437-1462, Sept. 1997.
6. Sadaoki Furui, "Digital Speech Processing, Synthesis and Recognition", 2nd edition.
7. Jian Yang; Zhang, D.; Frangi, A.F.; Jing-yu Yang; "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol.26, no.1, pp.131-137, Jan. 2004
8. Jolliffe. I.T.; *Principal Component Analysis*", Second Edition
9. Aapo Hyvarinen, Juha Karhunen and Erkki Oja. "Independent Component Analysis".

АҚПАРАТТЫҚ ҚАУІПСІЗДІК ҚАУІП-ҚАТЕРЛЕРІНІҢ ЖІКТЕЛУІ

Капалова Н.А., Самрат С.М.

e-mail: nkapalova@mail.ru, sanzhar.samrat@gmail.com

ҚР БҒМ ҒК «Ақпараттық және есептеуіш технологиялар институты»,
Қазақстан

Аңдатпа. Бұл мақалада ақпараттық қауіпсіздіктің негізгі қауіп-қатерлері мен осалдықтары қарастырылып, әлемдік статистикаға талдау жасалынды. Қауіп-қатерлерді бағалау әдістері және қауіптерге қарсы тұру жолдары көрсетілген. Мониторинг жасайтын бағдарламалардың қызметі мен түрлері қарастырылды.

Ақпараттық қауіпсіздік мәселелері көптеген елдерде жетекші орындарды иеленеді. «Перспективный мониторинг» компаниясының зерттеулеріне сәйкес, 2017 жылы шабуылшылардың ең көп бөлігі жеке тұлғаларға бағытталған [1]: олар барлық шабуылдардың төрттен бір бөлігін құрайды (26%). Ұйымдарға келетін болсақ, кибершабуылдардан көп зардап шеккендер – мемлекеттік құрылымдар (13%), банктер және онлайн-сервистер (8%).

Microsoft корпорациясы ақпараттық қауіпсіздік қауіпі туралы статистикалық есебі бойынша бұлттық технологиялардың танымалдылығының артуымен оларға жасалатын шабуылдардың саны өсті. 2017 жылы өткен жылмен салыстырғанда әлемде 4 есе көп ақпараттық қауіпсіздік қауіп-қатерлері болды. Microsoft тіркеулік жазбасына зиянды IP-адресстерден кіріп шығу әрекеттері 44%-ға артып, бұлттық сервистердің инфекциясының негізгі себебіне (51%) айналды. Сондай-ақ, ең көп тараған шабуылдар [2]: қашықтан қатынау хаттамалары (23%), спам (19%), портты сканерлеу (3,7%), SSH хаттама (1,7%) және басқалар арқылы жүзеге асырылады. Трояндар ең көп таралған санатқа айналды: 2017 жылдың бірінші тоқсанының аяғында қолданушылардың 10.26%-ында табылды. Екінші және үшінші орынды вирустар (1,59%) және троян жүктегіштер (0,64%) алды.

Kaspersky Security Network мәліметтері бойынша 2018 жылдың үшінші тоқсанындағы ақпараттық қауіптер статистикасы [3]:

- «Kaspersky Lab» шешімдері әлемнің 203 елінде орналасқан интернет-ресурстардан 947027577 шабуылдарға тойтарыс берді.
- Веб-антивирустар зиянды 246695333 URL мекен-жайларды тіркеді.
- Банктік шоттарға онлайн қол жеткізу арқылы ақша ұрлайтын зиянды бағдарламаларды іске қосу әрекеттері 305315 пайдаланушылардың компьютерлерінде тіркелді.
- Шифрлаушы бағдарламалардың шабуылдары 259867 пайдаланушылардың компьютерлерінде тіркелді.
- Антивирус 239177356 зиянды және жағымсыз нысандарды анықтады.

Жыл сайын ақпараттық жүйелерге жасалатын шабуылдар және қауіп-қатерлердің саны мен түрлері көбеюде. Ақпараттық қауіпсіздік қауіп-қатерлер – ақпараттық қауіпсіздіктің бұзылуына әкелуі мүмкін әр түрлі әрекеттер. Басқаша айтқанда, бұл ақпараттық және компьютерлік жүйелерге зиян келтіруі мүмкін процестер, ықтимал оқиғалар немесе әрекеттер.

2017 жылы «Kaspersky Lab» мамандары автоматтандырылған жүйелердің әр түрлі компоненттерінде 63 осалдықты анықтады. Қауіпсіздік осалдықтарының 52%-ы DoS-шабуылына, 8% - осал жүйелерде кодты қашықтан іске қосуға мүмкіндік берді [2].

Ақпараттық жүйелердің қауіп-қатерлеріне және шабуылдарына қарсы тұру, қорғаныс әдістерін, құралдарын және қауіпсіздікті қамтамасыз ету саясатын таңдау үшін ақпараттық қауіпсіздік қауіп-қатерлерін жіктеу қажет.

Ақпараттық қауіпсіздік қауіп-қатерлерін екі түрге бөлуге болады: табиғи және жасанды. Табиғи құбылыстар адамға тәуелді емес, мысалы, дауыл, су тасқыны, өрт және т.б. табиғи құбылыстар жатады. Жасанды қауіп-қатерлер тікелей адамға тәуелді, олар байқаусызда және қасақана жасалуы мүмкін. Байқаусызда туындайтын қауіп-қатерлер немқұрайлылықтан, қауіп-қатерлердің салдарын білмегендіктен туындайды. Қасақана жасалатын қауіп-қатерлерге сырттан да, компанияның ішінен де жасалатын шабуылдарды жатқызуға болады[4].

Әр түрлі жіктеу әдістеріне байланысты ақпараттық қауіпсіздіктің барлық ықтимал қауіптерін келесі негізгі топтарға бөлуге болады [5]:

- Жағымсыз контент;
- Рұқсатсыз қол жеткізу;
- Ақпараттың жайылып кетуі;
- Деректердің жоғалуы;
- Алаяқтық;
- Кибертерроризм;

Жағымсыз контентке ақпаратты жою немесе ұрлау үшін тікелей жасалған зиянды бағдарламалар, сонымен қатар заңмен тыйым салынған сайттар немесе тұтынушы жасына сәйкес келмейтін ақпаратты қамтитын қажет емес сайттар, ықтимал қауіпі бар бағдарламалар мен спамдар жатады.

Рұқсатсыз қол жеткізу. Рұқсатсыз қол жеткізу – белгілі бір ақпаратты пайдалануға рұқсаты жоқ қызметкердің ақпаратты қарауы. Рұқсатсыз кіру ақпараттың жайылып кетуіне әкеледі. Ақпараттың түріне және сақталуына байланысты, жайылып кетулер түрлі жолдармен ұйымдастырылуы мүмкін, атап айтқанда, веб-сайттарға шабуыл жасау, хакерлік бағдарламалар, желідегі деректерді ұстап қалу, рұқсат етілмеген бағдарламаларды пайдалану.

Ақпараттың жайылып кетуі. Туындаған себептерге байланысты, қасақана және кездейсоқ болып бөлінеді. Кездейсоқ жайылып кетулер аппараттық, бағдарламалық қамтама және адамдық қателіктерден туындайды. Қасақана жайылып кетулер деректерге қол жеткізу, зақым келтіру үшін әдейі ұйымдастырылады.

Деректердің жоғалуы. Деректердің жоғалуы ақпараттық қауіпсіздіктің негізгі қауіп-қатерлердің бірі деп санауға болады. Ақпараттың тұтастығының бұзылуы жабдықтың дұрыс жұмыс істемеуінен немесе пайдаланушылардың қасақана әрекеттерінен туындауы мүмкін.

Алаяқтық. Фрод – ақпараттық технологияларды қолданумен жасалатын алаяқтық. Алаяқтық тек қана несие картасын манипуляциялауды және онлайн-банктің бұзылуын ғана емес, ішкі фродты да қамтиды. Бұл экономикалық қылмыстың мақсаты – компанияның нормативтік актілерін бұзу, мүлікті иемдену.

Кибертерроризм. Жыл сайын бүкіл әлемде террористік қауіп біртіндеп виртуалды кеңістікке ауысады. Бірақ мұндай шабуылдар алдын-ала барлаусыз жүзеге асырылмайды, сондықтан қажетті деректерді жинауға көмектесетін кибершпионаж қажет.

Ақпараттық қауіпсіздік қатерлерін талдау кезінде үш негізгі әдіс пайдаланылады:

- Тікелей сараптама. Сараптамалық бағалау әдісі сарапшылардың қауіп-қатер параметрлерін белгілеуіне негізделеді. Сарапшылар ақпараттық қауіпсіздік қауіп-қатерлерін сипаттайтын параметрлердің тізімін анықтайды және әрбір параметрдің маңыздылығын субъективті коэффициенттермен қамтамасыз етеді.

- Статистикалық талдау. Статистикалық талдау ақпараттық қауіпсіздік оқиғалары туралы жинақталған деректерге, атап айтқанда, белгілі бір типтегі қауіптердің пайда болу жиілігіне, олардың көздері мен орындалмау себептерін талдауға негізделген. Мысалы, қауіптің пайда болу жиілігі белгілі болса, сол қауіптің белгілі бір уақыт аралығында пайда болу ықтималдығын анықтауға мүмкіндік береді. Статистикалық әдісті тиімді пайдалану үшін үлкен инциденттердің дерекқоры қажет.

- Факторлық талдау белгілі бір ықтималды қауіптерді іске асыруға және бір немесе басқа жағымсыз салдарларға әкелетін факторларды анықтауға негізделген. Мұндай факторлар киберқылмыскерлер үшін қызықты ақпараттың болуы, ақпараттық жүйелердің осалдығы, сыртқы ортадағы вирустық белсенділіктің

жоғары деңгейі және т.б. болуы мүмкін. Қазіргі заманғы ақпараттық жүйелерге көптеген факторлар әсер еткендіктен, әдетте көп өлшемді талдау пайдаланылады.

Ақпараттық қауіпсіздіктің қауіп-қатерлерін талдай отырып, әр түрлі аналитикалық әдістерді қолдану тиімді. Бұл бағалаудың дұрыстығын айтарлықтай жақсартады.

Ақпараттық қауіпсіздіктің қауіп-қатерлеріне және ақпараттық жүйелерге жасалатын шабуылдарға қарсы тұрудың ең тиімді шешімдерінің бірі – компьютерлік жүйелердің ақпараттық қауіпсіздігіне әсер ететін оқиғаларды тіркеу және жедел талдау құралдарын пайдалану. Компьютерлік желідегі қауіпсіздік оқиғасын жедел талдауды қамтамасыз ету үшін қауіпсіздік оқиғаларының мониторинг жүйесін әзірлеу және енгізу керек. Қауіпсіздік оқиғаларының мониторинг жүйесі жаңа қауіпсіздік оқиғалары туралы деректерді автоматтандырылған жинауды және жылдам талдауды қамтамасыз етеді.

Мониторинг орталығы ақпараттық технологиялар инфрақұрылымының қауіпсіздігі, осалдықтары және пайда болған ақпараттық қауіпсіздік оқиғалары туралы ақпаратты жинау, сақтау және өңдеу үшін арналған алаң болып табылады.

Мониторинг орталығының көмегімен компания келесі міндеттерді шеше алады [6]:

- АТ-инфрақұрылымын инвентаризациялау және қызмет көрсетуін қамтамасыз ету;
- Жүйенің барлық ішкі және сыртқы көздерінен ақпараттық қауіпсіздік оқиғалары туралы жүйелі ақпарат жинау;
- АҚ оқиғаларына тез және тиімді әрекет ету;
- АҚ оқиғалары туралы ақпаратты жинау және талдау, сондай-ақ АҚ оқиғаларын болдырмау үшін қорғау жүйелерінің жұмысын талдау;

Қазіргі кезде ақпараттық қауіпсіздікті қамтамасыз ету саласында мониторинг жасайтын жүйелерінің түрлері өте көп. Көптеген үлкенді-кішілі компаниялар мониторинг жасайтын бағдарламалар мен жүйелердің өз шешімдерін әзірлеген. Мониторинг жасайтын бағдарламалардың функциялары алуан түрлі және бұл бағдарламалар көптеген ақпараттық қауіпсіздік мәселелерін шешуге көмектеседі. Жасалып жатқан жұмыстың негізгі мақсаты – ақпараттық қауіпсіздіктің негізгі қауіп-қатерлері мен осалдықтарын талдау, оларды бағалау әдістерін қарастыру, ақпараттық қауіпсіздіктің қауіп-қатеріне мониторинг жасайтын бағдарламаларды жетілдіру мақсатында тиімді әдістерді қарастыру, оларды зерттеу және бағдарлама әзірлеу.

Пайдаланылған әдебиеттер тізімі:

1. Отчет Центра мониторинга информационной безопасности за I полугодие 2018г. – <https://amonitoring.ru> (07.01.2019).
2. Security Intelligence Report 2017. – <https://www.securitylab.ru> (12.01.2019).
3. Развитие информационных угроз в третьем квартале 2018 года. Статистика. – <https://securelist.ru> (20.01.2019).
4. Плетнёв П.В. Оценивание вероятности угроз в общей схеме определения рисков информационной безопасности / В.М. Белов, О.А. Крыжановская, П.В. Плетнёв // 2-я Международная научно-практическая конференция. - Симферополь-Гурзуф, КФУ им. В.И. Вернадского, 2016.
5. Системы мониторинга событий безопасности. – <https://www.anti-malware.ru> (24.01.2019).
6. Центр мониторинга компьютерных атак. – <https://amonitoring.ru> (27.01.2019).

7. Баранова Е.К. Методики анализа и оценки рисков информационной безопасности / Вестник Московского университета им. С.Ю. Витте. Серия 3: Образовательные ресурсы и технологии, 2015. – № 1(9). - С.73-79.
 8. Баранова Е.К., Бабаш А.В. Информационная безопасность и защита – М.: Инфра-М, 2017. – 324 с.
 9. Нестеров С.А. Основы информационной безопасности – Санкт-Петербург: Лань, 2016. – 324 с.
 10. Ревнивых А.В., Федотов А.М. Мониторинг информационной инфраструктуры организации / Вестник НГУ. Серия: Информационные технологии, 2013. – Т. 11, вып. 4. С. 84–91.
 11. Прузан А.Н., Николаенко В.Л., Сечко Г.В. Мониторинг инцидентов информационной безопасности в облачных вычислениях на малом предприятии / Доклады БГУИР, 2015. – № 7 (93). – С. 126 - 128.
 12. Галицкий А. В., Рябко С. Д., Шаньгин В. Ф. Защита информации в сети. – М.: ДМК Пресс, 2014. – 616 с.
 13. Петренко С.А., Курбатов В.А. Политики информационной безопасности. – М.: Компания АйТи, 2014. – 400 с.
 14. Шаньгин В.Ф. Защита компьютерной информации. Эффективные методы и средства. – М.: ДМК Пресс, 2013. – 544 с.
 15. Воронов А.А. Обеспечение системы управления рисками при возникновении угроз информационной безопасности / Информация и безопасность, 2006. – Т. 9. № 2. С. 8–11.
 16. Мухамадиева З.Б. Защита информации в информационных системах / Молодой ученый. – 2018. – №9. – С. 34-36.
- Партыка Т.Л., Попов И.И. Информационная безопасность. – 2-е изд., М.: ФОРУМ:ИНФРА-М, 2007. – 368 с.

ЗИЯНДЫ БАҒДАРЛАМАЛАРДЫҢ ЖҰМЫСЫНА ТАЛДАУ ЖАСАУ

Капалова Н.А., Хаумен А., Сулейменов О.Т.

e-mail: kapalova@ipic.kz, haumen.armanbek@gmail.com,
suleimenov97@gmail.com

*ҚР Білім және ғылым министрлігі Ғылым комитетінің
«Ақпараттық және есептеуіш технологиялар институты»*

***Аңдатпа.** Бұл мақалада қазіргі таңдағы ақпараттық технологияларда кеңінен таралған зиянды бағдарламалық қамтамаларға (Malicious Software) зерттеу жасалынды. Олардың даму тенденциялары, қоғамға әкелетін кері әсерлері қарастырылған. Заманауи зиянды бағдарламалардың таралу функцияларына байланысты классификация көрсетілген. Зиянды бағдарламаларға қарсы бағытталған бағдарламалық кешендердің негізгі кемшіліктері сипаттала отырып, оларды талдаудың статикалық және динамикалық әдістерінің негізгі маңыздылықтары көрсетілген.*

Қазіргі таңдағы ақпараттық коммуникациялық технологиялар адамның және қоғамның ажырамас, қажетті нағыз көмекшілері және оларсыз заманауи әлемді көзге елестету мүмкін емес. Мемлекеттік ұйымдар, коммерциялық фирмалар, сонымен қатар қоғамның әрбір мүшесі ақпараттық технологияларды күнделікті өмірде қолданатынын ескертпесек те болатын шығар. Бүгінде компьютерлік бағдарламалар мен қосымшалар үлкен жылдамдықпен әзірленуде. Компьютерлік техника мен интернеттің қарқынды дамуына орай ақпараттық әлеуметтік қоғамда адамдар мен жеке компанияларды өздерінің жеке деректерінің қауіпсіздігі, тұтастығы алаңдатуына байланысты ақпаратты қорғау мәселелері пайда болуда. Жыл сайын индустриалды және қаржылық кәсіпорындар зиянды бағдарламалардың әсерінен, бағдарламалық қызметтерге зиян келетін шабуылдарға ұшырайды. Шабуылдар басқарудың жоғарылауы, ақпараттардың ұрлану, бағдарламалық қамтамалардың істен шығуы және тағы басқа да ақпараттық қауіпсіздіктің негізгі мәселелеріне жағымсыз әсерлерін тигізеді. Бүкіл әлем бойынша миллиондаған компьютерлерді, ірі кәсіпорындардың жұмысын істен шығару үшін бірнеше байт код жеткілікті. Осыған дейін әлемдегі ең қауіпті Mydoom, LoveYou, Melissa секілді он вирустың өзі 100 миллиард АҚШ долларынан астам шығын әкелген [1].

Зиянды бағдарламалардың әсері компьютерлік жүйелер үшін негізгі қауіптердің бірі болып табылады. Жаңа қауіп-қатерлердің пайда болуы мен таралуы технологиялық прогресс заманында өте тез жүреді. Осының салдарынан жаңа зиянды бағдарламаларды жылдам табу және оларға дер кезінде идентификация жасау антивирустық бағдарламалық қамтамаларды әзірлеушілердің негізгі міндеті. Бұл міндетті қолмен немесе автоматтандырылған сигнатуралық талдау сияқты тек статикалық әдістерді пайдалана отырып тез және тиімді шешу бүгінгі күні осы талдауды қиындататын әр түрлі әдістердің белсенді қолданылуына байланысты мүмкін емес. Оның негізгі себебі зиянды бағдарламалардың қарапайым бағдарламалардан айырмашылығын жасыру үшін обфускация, полиморфизм, өзін-өзі шифрлеу, метаморфизм және сол секілді тағы да басқа көптеген әдіс-тәсілдердің пайда болуы [2]. Соған қарамастан көптеген жаңа зиянды бағдарламалардың мақсаттық типтері бұрынғыдай болып қалуды. Осыған байланысты талдаудың динамикалық әдістеріне қызығушылық артуда. Оларды статикалық талдаумен бірге пайдалану зиянды бағдарламалардың үлгілерін толық және терең зерттеуге мүмкіндік береді. Талдаудың заманауи тәсілдерін тиімді пайдалану арқасында антивирустық бағдарламалық қамтамаларды әзірлеушілер жаңа қауіп-қатерлерді алып келетін зиянды бағдарламаларды дер кезінде тауып, олардың одан әрі таралуын болдырмай, зиянкестерге қарсы базалар мен зиянды бағдарламалар каталогтарына тиісті өзгерістер енгізе алады.

Зиянды бағдарламалық қамтамалардың (зиянды БҚ) дамуы және олардың кеңінен таралуы қорғаудың неғұрлым заманауи механизмдерінің дамуына алып келеді. Киберқауіпсіздікпен айналысатын McAfee зертханасының деректері бойынша 2018 жылдың 4 тоқсанында 350 миллионнан астам бірегей зиянды бағдарламалар пайда болды (2018 жылдың 3 тоқсанына қарағанда 17% өсімді көрсетуде). Symantec зерттеуі бойынша мамыр айында 44,5 миллионнан астам зиянды БҚ пайда болған. 2017 жылы әлемдік қоғамда кеңінен тарау алған WannaCry, Petya, NotPetya атты зиянды бағдарламалармен таныстық [3]. «Лаборатория Касперского» зертханасының 2018-жылғы көрсеткіші бойынша қайталанбас, бір-бірінен өзгеше 192 053 604 зиянды бағдарлама табылған [4].

Зиянды БҚ – бұл белгілі бір компьютерге, серверге немесе компьютер желісіне зиянын келтіретін кез келген бағдарламаға сілтеме жасау үшін пайдаланылатын жалпы термин [5]. Зиянды бағдарлама - бағдарламалық қамтамалардың кең санатынан тұрады. Олар рұқсатсыз орнатылып, компьютердің жұмысына кері әсер етеді, деректердің тұтастығы мен сақталу қауіпсіздігін бұзады сонымен қатар логин мен пароль, банктік деректер сияқты жеке ақпаратты ұрлау үшін әзірленуі мүмкін, сондай-ақ олар компьютердегі құнды деректерді шифрлауға және иесіне дешифрлеу үшін кілтін сатып алуды мәжбүрлеуге тырысуы мүмкін [6].

Антивирустық бағдарламалық қамтамаларды әзірлеуші компанияларда зиянды бағдарламалардың жеке жіктемелері мен номенклатуралары болады. Қазіргі таңдағы таралған зиянды бағдарламалық қамтамаларды олардың таралу тәсілдері, зиян көрсету әсері, жұмыс істеу принциптері және т.б. қасиеттері бойынша жіктеуге болады [7].

Таралу тәсілдері бойынша зиянды бағдарламаларды келесідей бес түрге жіктесек болады [7]:

- компьютерлік вирус;
- экслойт;
- желілік құрт;
- трояндық бағдарлама;
- логикалық бомба және т.б.

Компьютерлік вирус – өзі көшірмелерін жасай алатын және басқа бағдарламалардың кодына, жүйелік жады аймақтарына, жүктеу секторларына ену қабілеті бар болатын және әр түрлі байланыс арналары арқылы өзінің көшірмелерін тарататын зиянды бағдарламалардың бір түрі [8]. Өзінің зиянын дербес компьютерге, корпаративтік немесе жеке желілерге, сонымен қатар басқа да ақпараттық технологияларға алып келетін вирустардың негізгі мақсаттарының бірі сол аймақтармен арналарда кеңінен таралу болып табылады. Олардың функционалды қызметі бағдарламалық және аппараттық кешендердің жұмысын бұзады - файлдарды жояды, тіпті операциялық жүйені жою қауіпі бар, деректерді орналастыру құрылымдарын жарамсыз ету, пайдаланушылардың жұмысын блоктау және т.б қабілеттерге ие.

Эксплоит деп – бағдарламалық қамтамалардың осалдықтарын пайдалану арқылы компьютерлік жүйеге шабуыл жасау үшін пайдаланылатын компьютерлік зиянды бағдарламаны айтады [9]. Шабуылдың мақсаты жүйені бақылауды (жеңілдіктерді көтеру) және оның дұрыс жұмыс істеуін бұзуды (DoS-шабуыл) алып тастауға бағытталған.

Желілік құрттардың компьютерлік вирустардан айырмашылығы, олар адамдардың араласуын қажет етпейді: бір компьютерге зиян келтіреді, ал компьютерлік желілер арқылы олардың иелерінің қатысуынсыз басқа машиналарға таратылады [10]. Электрондық почтадағы кемшіліктер сияқты желілік осалдықтарды қолдану арқылы құрттар өздерінің мындаған көшірмелерін жіберіп, барлық жаңа жүйелерді жұқтыра алады, содан кейін бұл үрдіс қайтадан басталады. Көптеген құрттардың жүйелік ресурстардың көп бөлігін пайдалану, компьютерлік өнімділікті азайтатынымен қатар, олардың көпшілігінде файлдарды ұрлауға немесе жоюға арналған зиянды компоненттер болады.

Логикалық бомба – зиянды әрекеттерді орындау үшін белгілі бір уақытта немесе белгілі бір ақпараттық жағдайларда жұмыс істейтін бағдарлама (әдетте ақпаратқа рұқсатсыз қол жеткізу, деректерді өзгерту немесе жою) [11]. Вирустар немесе құрт сияқты көптеген зиянды бағдарламаларда көбінесе белгілі бір уақыт

ішінде (сағаттық механизмдермен логикалық бомбалар) немесе кейбір шарттар орындалғанда өз жұмысын атқарады.

Зиянды БҚ-ға қарсы бағытталған бағдарламалық қамтамаларды әзірлеуші бірлестіктер өз өнімдерінің жұмыс істеу принциптерін ашық етуге тырыспайды және оның маңызды себептері бар. Біріншіден, жұмыс механизмдерінің ашылуы зиянды бағдарламалардың жаңа түрлерінің пайда болуына әкеп соғады. Сонымен қатар ондай бағдарламалардың осал тұстарын анықтауға, қорғаныс механизмдерінің әлсіздігін анықтауға жол ашады [12]. Екіншіден, маркетингтік саясат зиянды бағдарламаға қарсы өнімге мүмкіндігінше ашық емес, бірақ өз жұмысы мен функцияларын ұқсас өнімдерден тиімдірек орындайтын өнім ұсыну қажет. Осындай себептерге байланысты кез-келген деректерді қорғау механизмдері өздерінің жұмыс істеу принциптерін және функцияларын құпия түрде ұстауға тырысады. Осы өнімдердің жұмысын ұйымдастыруға қойылған негізгі принциптер келесідей [13]:

- сигнатуралық талдау;
- эвристикалық талдау;
- бағдарламаны орындаудың эмуляциясы;
- операциялардың үзіліссіз мониторингі;

Зиянды БҚ-ларға қарсы бағытталған бағдарламалық кешендерде жоғарғы тізімдегі қорғаныс механизмдері болады және үлкен көлемді бағдарламалық кешен құру зиянды бағдарламалардың жұмысына талдау жасауды талап етеді. Қазіргі таңда бағдарламалардың жұмыс істеу принциптерін зерттеу екі ірі бағытпен айқындалған [14]:

1. Статикалық талдау (бинарлық файлдың құрылымын, оның атрибуттарын, логикалық құрылымдарын, орындалу ағынын және деректерін талдау).

2. Динамикалық талдау (орындау кезінде бағдарламаның іс-қимылдарын қадағалау, оның профилін құру).

Кез-келген бағдарламаны талдаудың бұл түрлерін бастапқы және ілгері деп аталатын екі деңгейге бөліп қарастырса болады және осы екі талдау әдістерін пайдалана отырып зиянды бағдарламалардың жұмыс істеу принциптерін анықтаса болады [15].

Бастапқы деңгейдегі статикалық талдау кезінде орындалатын файл әрекеттері мен функцияларының қатысуынсыз зерттеледі. Талдау хэш-мәнін анықтау, мәтіндік жолдардан, оның функцияларынан, файлдың тақырыпшасынан ақпарат жинау арқылы жүзеге асырылады. Ілгері деңгейдегі статикалық талдауда зиянды бағдарламаны кері жобалау әдісі арқылы оны зерттеу жұмысы жүргізіледі. IDA Pro, WinMD5, Strings утилитасы, PEiD, Dependency Walker, PEview, Resource Hacker, PEBrowse Professional, File Alyzer секілді бағдарламалар арқылы әр түрлі деңгейлердегі статикалық талдау жасауға болады [16].

Бастапқы деңгейдегі динамикалық талдау кезінде зиянды бағдарламаның орындалуы арқылы оның іс-әрекеттерін бақылау жүзеге асырылады. Ілгері деңгейдегі динамикалық талдауда бағдарламаны ретке келтіруші (debugger) қолданылады. Ол орындалып жатқан зиянды бағдарламаның ішкі құрылысын зерттеуге бағытталған. Бастапқы деңгейде Process Monitor, Process Explorer, Regshot, Wireshark, INetSim бағдарламаларын, ілгері деңгейдегі талдауда OllyDbg, WinDbg бағдарламаларын қолданса болады [17].

Әр әдістің өз артықшылықтары мен кемшіліктері бар. Әдетте, зиянды бағдарламаларды анықтаудың мәнін жоғарылату үшін олар бір уақытта пайдаланылады. Осы әдістердің әрқайсысында шын мәнінде файл таза болған кезде зиянды файл ретінде қабылдау секілді қате анықтау мүмкіндіктері бар. Мұндай

жағдайларда болжамды файл емдеу файлы бұлдіруі мүмкін, бұл ақпараттың жоғалуына әкеледі.

Өкінішке орай, зиянды БҚ-ларға қарсы бағытталған бағдарламалық кешендердің барлық әзірлеушілері өз туындыларының ішкі құрылымын мұқият жасырады. Әдетте, ашық деректерде «сигнатуралық іздеу», «эвристика», «эмуляция», «проактивті қорғау», «HIPS» және т. б. жалпы атауларды пайдалана отырып қана өз өнімдерін сипаттайды. Бұл бізге қазіргі таңдағы антивирустық бағдарламалардың деректердің тұтастығын, қолжетімділігін қамтамасыз етуде олардың жұмыс істеу принциптерімен танысуға, алгоритмдеріне мұқият зерттеу жүргізуге мүмкіндіктер бермейді [18].

Қазіргі таңда зиянды бағдарламалық қамтамалардың үлкен компаниялардан бастап дербес компьютерді қолдаланатын қарапайым пайдаланушыларға дейін өз зияндарын тигізетінін ескерсек, оларға қарсы ақпараттық қауіпсіздікті қамтамасыз ету технологиялық қоғамдағы басты мәселелердің бірі болып есептелінеді. Біздің негізгі мақсат зиянды бағдарламалардың жұмысының алгоритмдеріне талдау жасай отырып, олардың жүйеге, әр-түрлі масштабтағы бағдарламалық кешендерге әсерін зерттеу сонымен қатар белгілі бір зиянды бағдарламалық қамтамаларды зерттеуге арналған тәсіл ұсыну. Зиянды БҚ-ға, олардың дамуының заманауи тенденцияларына жүргізілген шолу нәтижелері бойынша зиянды БҚ-лардың белгілі бір санаты таңдалып, оларды зерттеуге арналған компьютерлік бағдарлама әзірлену жоспарлануда.

Пайдаланылған әдебиеттер тізімі:

1. Radziwill Y. Cyber-Attacks and the Exploitable Imperfections of International Law – BRILL, 2015. – P. 66-68.
2. Jean-Marie Borello, Ludovic Me. Code obfuscation techniques for metamorphic viruses / Journal in Computer Virology, 2008. Vol. 4. P. 211-220.
3. Symantec Intelligence Report May 2015. – <https://www.symantec.com/content/dam/symantec/docs/security-center/archives/intelligence-report-may-15-en.pdf> (08.12.2018).
4. Чебышев В., Сеницын Ф., Паринов Д., Лискин А., Купреев О. Развитие информационных угроз во втором квартале 2018 года. – <https://securelist.ru/it-threat-evolution-q2-2018-statistics/90919/> (11.12.2018).
5. Что такое вредоносное программное обеспечение? – <http://kz-cert.kz/ru/malicious-software> (16.11.2018).
6. Собейкис В.Г. Азбука хакера 3. Компьютерная вирусология / Собейкис В.Г – Москва: Майор, 2006. – 512 с.
7. Лаборатория Касперского. Классификация вредоносных программ – <https://www.kaspersky.ru/blog/klassifikaciya-vredonosnyx-programm/2200/> (25.01.2019).
8. Холмогоров В. PRO Вирусы. – Санкт-Петербург: Страта, 2017. – 162 с.
9. Эриксон Дж. Хакинг. Искусство эксплойта. – Санкт-Петербург: Питер, 2008. – 496 с.
10. О вирусах: Основные источники проникновения угроз на компьютер. – <http://support.kaspersky.ru/789> (03.02.2019).
11. Таненбаум Э. Логические бомбы / Современные операционные системы. Modern operating systems. – Питер, 2011. – 1120 с.
12. Гошко С.В. Технологии борьбы с компьютерными вирусами. Практическое пособие / С.В Гошко. – Москва: СОЛОН-ПРЕСС, 2009. – 351 с.

13. Гульев И.А. Создаем вирус и антивирус / Игорь Гульев – Москва: ДМК-Пресс, 2011. – 250 с.
14. Elisan C.C. Advanced Malware Analysis. – McGraw-Hill Education, 2015.
15. Hahn K. Robust Static Analysis of Portable Executable Malware / HTWK Leipzig, 2014. – P. 13-17.
16. Пискунов И. Реверсинг малвари для начинающих. <https://xaker.ru/2016/12/22/reversing-malware-tutorial-part2/> (12.01.2019).
17. Dang B., Gazet A., Bachaalany E. Practical Reverse Engineering: x86, x64, ARM, Windows Kernel, Reversing Tools, and Obfuscation. – Wiley, 2014. – P 384.
18. Дроботун Е.Б. Изучаем антивирус. / Хакер, 2011. – 88-93 с.

ҚАЗАҚ ТІЛІНІҢ АВТОМАТТАНДЫРЫЛҒАН МАРКЕРЛІК КОРПУСЫН ӘЗІРЛЕУ

Сейтжаппар М.А.¹, Рахимова Д.Р.²
e-mail: mseitzhappar@mail.ru, di.diva@mail.ru

¹Әл-Фараби атындағы Қазақ Ұлттық Университеті, Алматы, Қазақстан.

²Ақпараттық және Есептеуіш Технологиялар Институты, Алматы, Қазақстан.

Аннотация. Қазіргі уақытта корпуслық лингвистика қолданбалы шешімдер саласы құралдардың кең спектрімен өңделуде. Алайда, ТТӨ құралдарының кез – келген заманауи іске асырылуы ішінара шешім және толық, әмбебап шешім бірнеше бағыттарда дамуы шығармашылық іздестіруді ынталандыратын осындай құралдардың болашақ әзірлемелерінің қайнар көзі болып табылатынын ескеру қажет. Корпус — белгілі бір ережелер бойынша таңдалған және өңделген, тілді зерттеу үшін база ретінде қолданылатын мәтіндер жиынтығы. Олар статистикалық талдау және статистикалық гипотезаларды тексеру, осы тілдегі лингвистикалық ережелерді растау үшін пайдаланылады. Әр түрлі лингвистикалық қолданбалы есептерде электронды мәтін корпус қажет етеді. Осы мақалада маркерлік корпус түрлері мен әзірлеу технологиялары көрсетілген. Қазақ тілі үшін автоматтандырылған маркерлік корпус әзірлеу жұмысы мен мысалдары келтірілген.

1. Маркерлеу мәселесі

Корпус – корпуслық лингвистиканың негізгі түсінігі мен деректер базасы. Корпустардың әртүрлі типтерін талдау және өңдеу компьютерлік лингвистика, сөйлеу және машиналық аударма саласындағы жұмыстардың көпшілігінің мәні болып табылады, онда корпустар сөйлеудің бөліктерін және басқа да міндеттерді таңбалау үшін жасырын маркалық модельдерді жасау кезінде жиі қолданылады. Корпустар мен жиілік сөздіктер шет тілдерін оқытуда пайдалы болуы мүмкін.

Мәтіндік корпустарды құрудың маңыздылығы:

- нақты контексте лингвистикалық деректерді ұсыну;
- деректердің үлкен өкілеттілігі (корпус үлкен көлемде);

мысалы, мәтінді графикалық және лексикалық-грамматикалық талдауды жүзеге асыру және т. б. сияқты әртүрлі лингвистикалық міндеттерді шешу үшін бір рет құрылған корпуссты бірнеше рет қолдану мүмкіндігі.

Мәтіндік корпусстарды құрудың мақсаттылығы түсіндіріледі:

- нақты контексте лингвистикалық деректерді ұсыну;
- деректердің үлкен өкілеттілігі (корпус үлкен көлемде);

-мысалы, мәтінді графикалық және лексикалық-грамматикалық талдауды жүзеге асыру және т. б. сияқты әртүрлі лингвистикалық міндеттерді шешу үшін бір рет құрылған корпуссты бірнеше рет қолдану мүмкіндігі.

Мәтіндердің бастапқы маркерлеуде әрбір корпус үшін міндетті кезеңдер жатады:

- токенизация (орфографиялық сөздерді бөлу)
- лемматизация (сөз қорын сөздік формасына келтіру)
- морфологиялық талдау

Нәтижелерді ұсыну мәселесі

Үлкен корпусстарда бұрын өзекті емес мәселе пайда болады: сұрау бойынша іздеу шектелген уақытта физикалық тұрғыдан қарауға мүмкін емес жүздеген және тіпті мыңдаған нәтижелер (қолдану контекстері) бере алады. Бұл проблеманы шешу үшін іздеу нәтижелерін топтастыруға және оларды ішкі жинақта автоматты түрде бөлуге мүмкіндік беретін (іздеу нәтижелерін кластерлеу) не олардың маңыздылығын статистикалық бағалаумен неғұрлым тұрақты сөз тіркестерін (коллокация) беретін жүйелер әзірленеді.

Кейбір корпусстарда талдаудың одан әрі құрылымдық деңгейлері қолданылады. Атап айтқанда, кейбір шағын корпусстар толығымен синтаксистік таңбалануы мүмкін. Мұндай корпусстар әдетте терең аннотацияланған немесе синтаксистік деп аталады, ал синтаксистік құрылым-тәуелділік ағашы. Тұтас корпуссты белгілеумен қамтамасыз ету күрделілігі мұндай корпусстар көбінесе аз және шамамен бір миллионнан үш миллионға дейін сөздерді қамтиды. Морфологияның, Семантиканың және прагматиканың аннотациясын қоса алғанда, лингвистикалық құрылымдық талдаудың басқа да деңгейлері де мүмкін.

2. Корпустық маркерлеу түрлері

Қазіргі уақытта мәтіндерде лингвистикалық және басқа да ақпарат түрлерін ұсынудың жалпыға танылған стандарттары жоқ. Text Encoding Initiative (TEI) арнайы халықаралық жобасы стандартталған таңбалар құралдарын әзірлеуге арналған. Ол үшін SGML құжаттарын белгілеудің жалпыға танылған халықаралық тілі немесе XML ішкі жиыны қолданылады. Корпус болуы мүмкін белгілер типтерін лингвистикалық және сыртқы лингвистикалық бөлуге болады. Соңғысы жатады:

- Мәтінді пішімдеу ерекшеліктерін көрсететін белгілеу (тақырыптар, абзацтар, шегіністер және т.б.);

- автор және мәтін туралы мәліметтерге қатысты белгі. Автор туралы мәліметте оның аты ғана емес, Жасы, жынысы, өмір жасы және т. б. болуы мүмкін, ал мәтін туралы мәліметтер әдетте атаудан басқа, сондай-ақ ол жазылған тілді, шыққан жылы мен орнын және т. б. қамтиды.

Мұндай ақпараттың болуы мәтіндік деректер базасында іздеуді айтарлықтай нақтылауға мүмкіндік береді және бұдан басқа, тиісті құжатты сәйкестендіру құралдарын ұсынады.

Маркерлеу мәтіннен тәуелсіз болуы тиіс: маркерді алып тастау және мәтінді онсыз қарау мүмкіндігі болуы және керісінше, тек таңбаны алып тастау мүмкіндігі

болуы тиіс. Белгілеу принциптері және оларды әзірлеушілер соңғы пайдаланушыға белгілі болуы тиіс. Пайдаланушы белгі қатесіз емес, тек әлеуетті пайдалы құрал болып табылатыны туралы хабардар болуы тиіс. Таңбалардың негізіне жалпы қабылданған және мүмкіндігінше теориялық бейтарап лингвистикалық қағидаттар жатқызылуы тиіс. Ақырында, бірде-бір таңба стандарт болып саналмайды. Мұндай ақпараттың жо мәтіндік деректер базасында іздеуді айтарлықтай нақтылауға мүмкіндік береді және бұдан басқа, тиісті құжатты сәйкестендіру құралдарын ұсынады.

Корпустарды қолдану арқылы зерттеудің статистикалық әдістерін қолдана отырып, тілдік құбылыстар туралы болжамдарды растауға немесе жоққа шығаруға болады. Зерттеушілердің алдына қойған міндеттерді шешу үшін корпустың болуы жеткіліксіз. Мәтінде лингвистикалық ақпараттың болуы қажет. Осылайша белгіленген корпус идеясы пайда болды. Таңбалар сөздердің жиілігін және әр түрлі тіл бөліктері өкілдерінің жиілігін есептеуге көмектеседі. Лингвистикалық таңба сөзге кодты (тэг) беру үшін пайдаланылады, ол сөзді сипаттайтын грамматикалық белгілер жиынтығын білдіреді. Белгілеулер шартты түрде бөліседі арналған " лингвистика және сыртқы лингвистика. Сыртқы лингвистикалық:

- мәтінді пішімдеу ерекшеліктерін көрсететін таңбалар (тақырыптар, абзацтар, шегіністер және т. б.);

- автор және мәтін туралы мәліметтерге қатысты белгілер. Авторда оның аты, жасы, жынысы, өмір сүру жылдары және т.б. көрсетілуі мүмкін, ал мәтінде атауы, тілі, шыққан жылы және орны және т. б. көрсетіледі.

3. Корпус прототипінің бағдарламалық жүзеге асырылуы

Корпус құрудың технологиясы келесі қадамдардан тұрады:

1. Мәтіндердің дереккөз талаптарына сай жинақталуы.

2. Машинаға ыңғайлы форматқа келтіру-құрылымдау. Корпусқа қажетті мәтіндер әртүрлі тәсілмен алынуы мүмкін: сканерлеу, қолмен теру, авторлық көшірме, интернет, түпнұсқалық макет т.с.с.

3. Талдау және мәтінді алдын ала өңдеу. Бұл қадамда әртүрлі дереккөздерден алынған мәтіндер филологиялық түзетулерден өтеді. Корпустың техникалық сипаттамасы мәтіннің библиографиялық және экстралингвистикалық сипаттамасын қамтиды.

4. Конверттеу және графематикалық талдау. Кейбір мәтіндер алдын ала машиналық өңдеудің бірнеше сатысынан өтеді. Бұл орайда (қажет болса) қайта кодтау, мәтіндік емес элементтерді (кесте, сурет) жою не өзгерт, жаңа азат жолдарды өшіру, сызықшаның бірегей қойылуын қадағалау (мысалы: <<->, <<->>) т.б. әрекеттер орындалады. Графикалық талдауда болса: кіріс мәтінін элементтерге (сөйлем, сөз т.б.) болу, лексикалық емес элементтерді табу және рәсімдеу, арнайы символдарды (атаулар (аты-жөн инициалдары)) өңдеу, өзге тіл лексемдерін бақылау, сурет атауларын өңдеу жүзеге асырылады.

5. Мәтінді таңбалау (разметка). Таңбалау кезінде мәтін сөздеріне қосымша мәліметтер (метмәлімет) тағайындалады. Метамәліметтерді 3 типке бөлуге болады: экстралингвистикалық-барлық мәтінге қатысты, мәтін құрылымы бойынша мәлімет, лингвистикалық метамәлімет мәтін элементтерін сипаттайды. Метамәліметтер элементтің библиографиялық, жанрлық сипаттамасына қатысты, стиліне және авторына қатысты мәліметтермен қоса файл аты, код тәсілі, таңбалау нұсқасы сынды формалды ақпараттарды да қамтиды. Әдетте айтылған мәліметтерді қолмен теріп енгізеді. Құжатты құрылымдық талдау (абзацтарға бөлу, сөйлемдер ажырату т.б.) автоматты түрде жүзеге асырылады.

6. Автоматты таңбалау нәтижелерін түзету: қателерді жөндеу, бірегей емес бөліктерді теңгеру.

7. Көпәспектiлi iздеу мен статистикалық талдауға жағдай жасайтын арнайы лингвистикалық ақпараттық iздеу жүйесiне конвертациялау (қорытқы этап).

8. Корпусқа ену мүмкiндiгiн жарату. Корпус жеке есептеуiш құрылғы жадынан бөлек басқа да тасымалдағыштар арқылы тiптi, глобальды желi арқылы төралуы мүмкiн.

9. Сұраныстар тiлiне, таңбалау тәсiлiне және корпуссты қолдану туралы нұсқалық бойынша құжаттамалар даярлау.

3.1 Кiрiс мәлiметтерi. Сөздi қор

Негiзгi мақсат маркерленге корпусстың репрезентативтiлiгi болып табылады. Жазбаша және ауызша тiл көптеген тiркеулерде (жанрларда және т. б. - әр түрлi дәстүрлердегi терминдер әртүрлi), дiни әдебиеттен техникалық әдебиетке дейiн, көпшiлiк саяси сөйлеуден бейресми әңгiмеге дейiн жұмыс iстейдi. Көп ұзамай бұл регистрлер көптеген тiлдiк деңгейлерде ерекшеленедi, онда корпуста олардың барлығы немесе көпшiлiгi ұсынылуы тиiс. Қазiргi тұста айтылғандай, Интернетте тек сол жерге қойғанын табуға болады; сондай-ақ корпусқа да қатысты. Бұл мiндеттi шешу, сөзсiз, мәтiндердi қамтуға байланысты. "Неғұрлым көп болса, соғұрлым жақсы" принципi мұнда жұмыс iстейдi; кез келген лингвистикалық құбылыстың жиiлiлiгi мен релевантылығы жүз мыңға қарағанда жүз миллион сөз қолданылатын көлемде бақылау оңай. Мұндай мiндет алғашқы корпусстарды, соның iшiнде көркем шығармалармен қатар журналдар мен газеттердi жариялаумен қатар енгiзiлген.

Корпус жұмыс ұсынатын қазақ тiлiне арналған құралдар келесi сатылар арқылы түрленедi:

- Токенизация;
- Лемматизация;
- POS-t;
- Форматтау;
- Сақтау.

Сөздiк қордың өлшеми шамамен 20 мыңнан аса лемма бар. Және олардың әрқайсысы iүмкiндiгiнше сөз таптары бойынша белгiленген.

Сөздiктен сөз iзделгенде оның соңғы кезiккен позициясы мен бүткiл мүмкiн морфологиялық қатыстылығы анықталады. Өйткенi сөз леммалары әртүрлi сөз табы болғандағана қарамастан сәйкес келедi. Себебi морфология және синтаксис сөз жасамда бiр бiрiн бұғауламайды.

Search класындағы search() әдiсi кiрiс параметрi келген сөз индексiн қайтарады.

```
int search(String s){
morph.clear();
int position=-1;
for(int i=0;i<listWord.size();i++)
if (listWord.get(i).equals(s)){
morph.add(listMorph.get(i));
position=i;}
return position;
}
```

Мәтiндi толықтай соңғы құрауыштарға, яғни токендерге бөлу үшiн TEXTM классы жұмыс жасайды. Stanford CoreNLP және OpenNLP пайдалана отырып мәтiндердi сөйлемге және сөздерге ыдыратады.

DocumentPreprocessor класы кіріс ағынының кіріс деректерін ыдыратады. Сонымен қатар, ол iterable интерфейсін іске асырады, ол кезектілігі бойынша жылжытуды оңай етеді. Токенизатор қарапайым мәтіндік және XML деректерді токенизациялауды қолдайды.

```

Reader reader = new StringReader(inputText);
DocumentPreprocessor dp = new DocumentPreprocessor(reader);
SimpleTokenizer simpleTokenizer = SimpleTokenizer.INSTANCE;
for (List<HasWord> sentence : dp) {
    List<Word> tmp = new ArrayList<Word>();
    String sentenceString = SentenceUtils.listToString(sentence);
    for (String token : simpleTokenizer.tokenize(sentenceString.toLowerCase()))
        if(token.length()>1)
        {
            try {
                stemmize=new Stemmize(token);
                stemmize.getStem();
                tmp.add(stemmize.word);
            }catch (Exception e){
                //System.out.println(e);
            }
        }
    text.add(tmp);
}

```

Ал StemmingMatrix класы стемминг жүзеге асыру үшін қажетті тілдік ерекшеліктерді және ережелерді ұсынады. Үзінді келтірер болсақ, төмендегі параметр есімшенің жалғауларын алып тастауда қолданылады:

```

final static String[][] regEsimshe={
{"ған", "ген", "жзаоұыяәөіерлмнңй"},
{"қан", "кен", "пкқтсш"},
{"мақ", "мек", "аоұыяәөіелруй"},
{"бақ", "бек", "мнңбғғзд"},
{"пак", "пек", "пкқтсш"},
{"атын", "етін", "жзрлмнңшйу"},
{"йтын", "йтін", "аоұыяәөіе"},
{"ушы", "уші", "жзаоұыяәөіерлмнңй"},
{"а", "е", "жзрлмшнңй"}
};

```

Бұл матрицаның алғашқы екі бағаны сәйкесінше қосымшаларды жазса, соңғы бағаны сол қосымшалармен үндесетін сөздің соңғы дыбысын қамтиды.

Ал Stemmize класы стемминг жүзеге асыру үшін қажетті және мүмкін болғанда сөз леммасын алу үшін қажет. Мұнда cutter() әдісі бар ол сәйкесінше StemmingMatrix класындағы параметрлерді пайдалана отырып сөз соңын қырқып тастайды және анықталған аффикстерді тіркеп отырады.

```

void cutter(String [][]regMatrix, String s, String type){
    int len=s.length();
    for(String []af:regMatrix){
        if(af[0].length()<len){
            String aff=s.substring(len-af[0].length());
            if(aff.equals(af[0])||aff.equals(af[1])){
                char c=s.charAt(len-af[0].length()-1);
                if(af[2].contains(""+c)){
                    word.addAFF(type, aff);
                    word.setStem(s.substring(0,len-aff.length()));
                }
            }
        }
    }
}
}
}
}

```

Нәтижелерді алу және сақтауда қолданылатын негізгі модель келесідей: модель әр токен үшін word, token, lemma, morph, affix параметрлеі бойынша инициализацияланады:

```

Stringtoken;
Stringstem;
Stringlemma;
Stringmorph;
Map<String,String>aff=new HashMap<String, String>();

```

Ал енді, маркерленген мәтінді JSON форматына түрлендіру және сақтау мәселесін шешу қажет. JSONProcessorклассы құрылды.TEXTM класының объектісін қабылдап алып, оны файлға жазуша writeJson(TEXTM textm)әдісіне ие:

```

public static void writeJson(TEXTM textm){
    File file = new File(path);
    int i= file.listFiles().length;
    ObjectMapper mapper = new ObjectMapper();
    try {
        mapper.setVisibility(PropertyAccessor.FIELD, JsonAutoDetect.Visibility.ANY);
        mapper.writeValue(new File( pathname: path+i+".json"), textm);
    } catch (IOException e) {
        e.printStackTrace();
    }
}
}

```

Ал getGaon(TEXTM textm) throws JsonProcessingException әдісі болса TEXTM textm объектісін json объектісіне түрлендіріп String түрінде қайтарады:

```

String getGaon(TEXTM textm) throws JsonProcessingException {
    ObjectMapper mapper = new ObjectMapper();
    mapper.setVisibility(PropertyAccessor.FIELD, JsonAutoDetect.Visibility.ANY);
    return mapper.writerWithDefaultPrettyPrinter().writeValueAsString(textm);
}

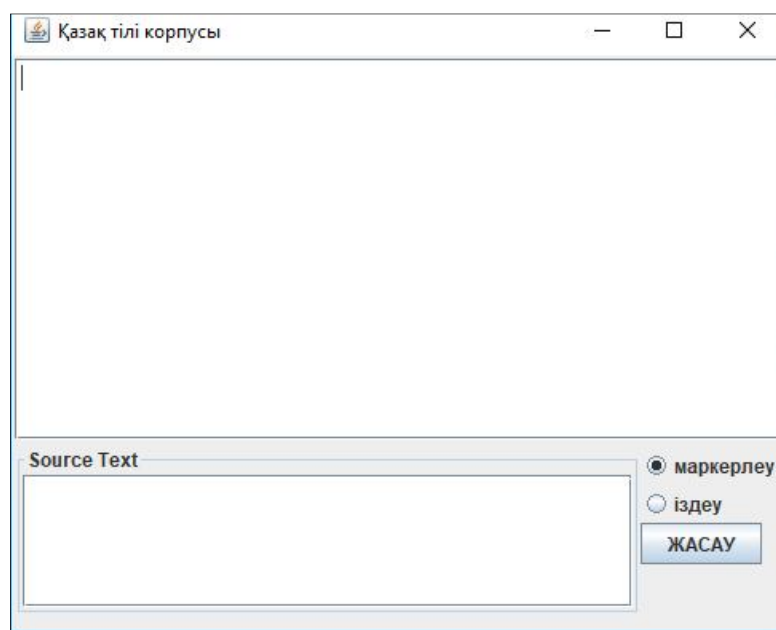
```

Енді жүйеміздің графикалық құралына тоқталар болсақ. Жүйе интерфейсі бастапқы көрінісі 1–суретте көрсетілген және өте қарапайым. Екі мәтіндік терезе де өзгертіле алынатын терезе болып табылады. Ал іс жүзіндегі кіріс терезесі төменгі терезе. Біз талдағымыз келген мәтінді сонда ұсынамыз.

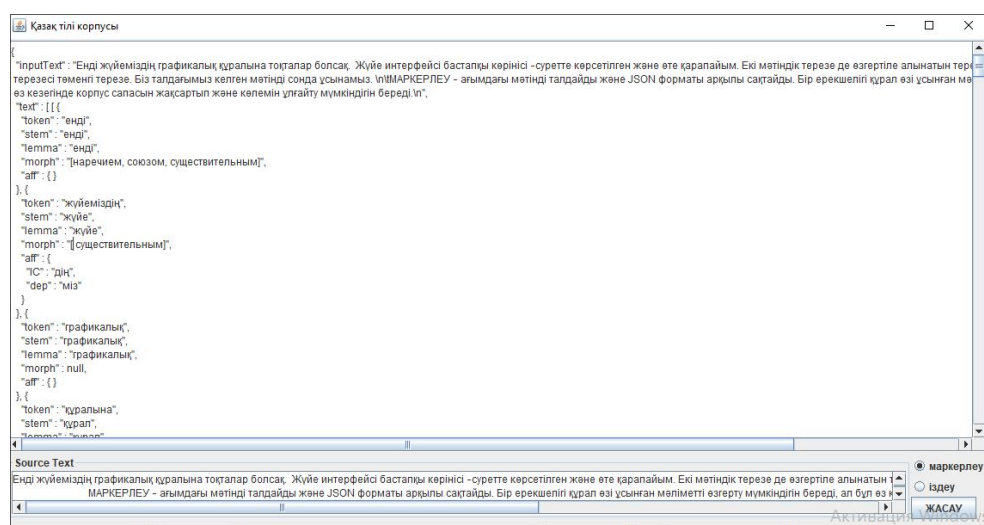
МАРКЕРЛЕУ – ағымдағы мәтінді талдайды және JSON форматы арқылы сақтайды. Бір ерекшелігі құрал өзі ұсынған мәліметті өзгерту мүмкіндігін береді, ал бұл өз кезегінде корпус сапасын жақсартып және көлемін ұлғайту мүмкіндігін береді.

ІЗДЕУ – корпус ішінін сөздерді іздеуге қызмет етеді.

ЖАСАУ – таңдалынған функция бойынша жұмысты жүзеге асырады.



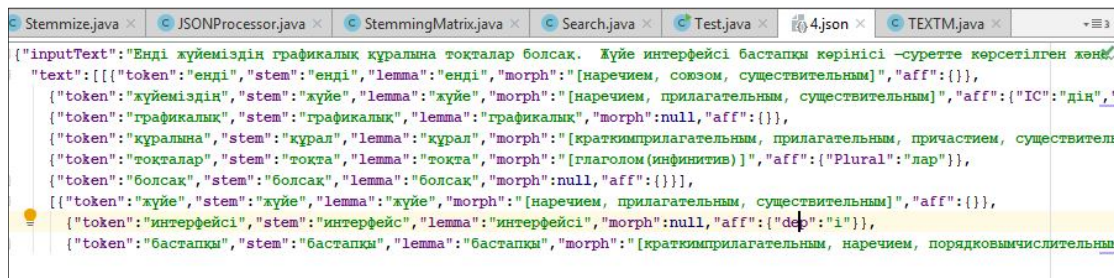
Сурет 1. АЖ-нің бастапқы беті



Сурет 2. Талдау мысалы

Ендігі кезекте біз өңделінген құжат қатесін толықтырып корпуста сақтай аламыз, және жасалынған .json файлы мазмұны 3–суретте көрсетілген.

Әрине жүйе жұмысының қателіктері бар, себебі қазірге сөздік қордың аздығы және тілдің математикалық моделін дұрыс құрастырып, алгоритмдерді жүзеге асыру қажет.



```
{
  "inputText": "Енді жүйеміздің графикалық құралына тоқталар болсақ. Жүйе интерфейсі бастапқы көрінісі –суретте көрсетілген және...",
  "text": [
    {
      "token": "енді",
      "stem": "енді",
      "lemma": "енді",
      "morph": "[наречием, союзом, существительным]",
      "aff": {}
    },
    {
      "token": "жүйеміздің",
      "stem": "жүйе",
      "lemma": "жүйе",
      "morph": "[наречием, прилагательным, существительным]",
      "aff": {"IC": "дің"}
    },
    {
      "token": "графикалық",
      "stem": "графикалық",
      "lemma": "графикалық",
      "morph": null,
      "aff": {}
    },
    {
      "token": "құралына",
      "stem": "құрал",
      "lemma": "құрал",
      "morph": "[краткимприлагательным, прилагательным, причастием, существительным]",
      "aff": {}
    },
    {
      "token": "тоқталар",
      "stem": "тоқта",
      "lemma": "тоқта",
      "morph": "[глаголом (инфинитив)]",
      "aff": {"Plural": "лар"}
    },
    {
      "token": "болсақ",
      "stem": "болсақ",
      "lemma": "болсақ",
      "morph": null,
      "aff": {}
    },
    {
      "token": "жүйе",
      "stem": "жүйе",
      "lemma": "жүйе",
      "morph": "[наречием, прилагательным, существительным]",
      "aff": {}
    },
    {
      "token": "интерфейсі",
      "stem": "интерфейс",
      "lemma": "интерфейсі",
      "morph": null,
      "aff": {"def": "і"}
    },
    {
      "token": "бастапқы",
      "stem": "бастапқы",
      "lemma": "бастапқы",
      "morph": "[краткимприлагательным, наречием, порядковымчислительным]"
    }
  ]
}
```

Сурет 3. Сақталған файл кескіні

Қорытынды

Корпустарды қолдану арқылы зерттеудің статистикалық әдістерін қолдана отырып, тілдік құбылыстар туралы болжамдарды растауға немесе жоққа шығаруға болады. Зерттеушілердің алдына қойған міндеттерді шешу үшін корпустың болуы жеткіліксіз. Мәтінде лингвистикалық ақпараттың болуы қажет. Осылайша маркерленген корпус идеясы пайда болды. Таңбалар сөздердің жиілігін және әр түрлі тіл бөліктері өкілдерінің жиілігін есептеуге көмектеседі. Қазақ тілі үшін әзірленген маркерлік корпус әрі қарай қолданбалы зияткерлік жүйелерде қолданыс табады.

Бұл зерттеу жұмысы АР05132950 «Қазақ тіліндегі деректерге арналған ақпараттық-аналитикалық іздеу жүйесін құру» жобасы бойынша жүзеге асырылды және қаржыландырылды, Ақпараттық және есептеуіш технологиялар институты, Алматы қ.

Әдебиеттер

1. Архипов А.В. Разметка лингвистическая. – http://www.lomonosov-fund.ru/enc/ru/encyclopedia:0127221#r1_1. (15.03.2019)
2. Азарова И.В., Алексеева К.Л., Захарова Л.А. Разметка текстовых фрагментов в корпусе агиографических текстов СКАТ / Труды международной конференции «Корпусная лингвистика – 2006». – СПб: изд-во С.-Петербур. ун-та, Изд-во РХГА, 2006. – С. 16-24.
3. Нурхан А.К., Рахимова Д.Р. Исследование и создание размеченного корпуса текстов для казахского языка /Сборник материалов Шестой Международной конференции по компьютерной обработке тюркских языков «TurkLang-2018» Ташкент, Узбекистан, 18–20 октября 2018 г. С. 127-133.
4. Bowker, L., Pearson, J. Working with Specialized Language: A Practical Guide to Using Corpora. /London, N.Y.: Routledge, 2002, 242 p.
5. М.П. Концевой, Концевая, Г.М. Компьютерный семантический анализ XML-реализации гипертекста /материалы Всероссийской научно-практической конференции, 15 марта 2010 г., Самара, РФ. – Самара: ПГСГА, 2010. – С. 61 – 66.
6. Беляева Л.Н. Лексикографический потенциал параллельного корпуса текстов / Труды международной конференции «Корпусная лингвистика – 2004». – СПб., 2004. – С. 55-64.

АВТОМАТТАНДЫРЫЛҒАН ЖҮЙЕЛЕРДЕГІ АҚПАРАТТЫ ҚОРҒАУ ӘДІСТЕРІНЕ ШОЛУ

Сейтқали Ғ.Т.

e-mail: g.seitkali@gmail.com

ҚР БҒМ ҒК «Ақпараттық және есептеуіш технологиялар институты»

***Аңдатпа:** Бұл мақалада автоматтандырылған жүйелерде ақпаратты қорғау барысында ескерілетін факторлар мен белгілі бір формалды ақпаратты қорғау құралына тоқталу жасалынған. Криптографиялық құралдарға шолу жасалынған. Электронды сандық (цифрлық) қолтаңбаны қолдану принциптері мен Эль-Гамаль алгоритмі, және оңтайландыру жолына қысқаша шолу жасалынған.*

Компьютерлік технологиялар біздің өмірімізде толыққанды өз орнын алғаны анық. Қазіргі таңда дербес компьютерсіз оңай жұмыс істейтін кәсіпорын немесе компанияны елестету қиын. Барлық ақпаратты бастың жадында немесе қағаз жүзінде ұстау мүмкін емес, сондықтан компьютерлік технологиялар біздің сандық әлемімізде өте құнды. Компьютерлік технология адамзатқа көмектесуге бағытталған, бірақ шексіз мүмкіндіктермен бірге инновациялық технологиялар жаңа мәселелерді тудырады. Олардың ішіндегі ең бастысы ақпаратты осы мәліметтерді пайдалануға қандай да бір құқығы жоқ адамдардан қорғау мәселесі болды. Сондықтан технологияларды жетілдірумен қатар ақпаратты қорғау әдістері тез дами бастады, бұл жаңа ақпараттық технологияларды әзірлеуден гөрі аса маңызды процесс болып табылады. Өйткені, қорғаныс жүйелерінің жақсартуларымен бір мезгілде бұзу алгоритмдері де жетілдіріледі. Ал бұл өз кезегінде дербес деректерді қорғау сенімділігін арттыру және жедел жетілдіруді талап етеді.

Қазіргі заманғы компьютерлік және ақпараттық технологиялардың даму деңгейі, сондай-ақ адам өмірінің барлық салаларында ақпаратты сақтау және өңдеудің электрондық әдістерін енгізу ауқымды деректерді сақтау, өңдеу, беру және пайдалану процестерін жетілдіруді талап етеді. Сонымен қатар, автоматтандырылған ақпараттық алмасу жүйелеріндегі ақпараттық қауіпсіздікті қамтамасыз ету мен ақпараттық қауіпсіздіктің маңыздылығы айтарлықтай артады.

Ақпаратпен алмасу автоматтандырылған жүйелердің ақпараттық қауіпсіздігін қамтамасыз ету негізін электронды сандық қолтаңба жүйелері қамтиды. Электрондық сандық қолтаңба (ЭСҚ) қандай да бір құжаттың авторын бір мәнді анықтау үшін қажет. ЭСҚ қандай да бір құжаттың немесе шарттың түпнұсқалығын белгілейтін әдеттегі қолдың аналогы болып табылады. Электрондық сандық қолтаңба жүзеге асыруға мүмкіндік береді:

- тұтастықты бақылау;
- құжатты өзгертуден (қолдан жасаудан) қорғау;
- авторлықтан бас тартудың мүмкін еместігі;
- құжаттың авторлығын дәлелдей отырып растау.

Автоматтандырылған ақпарат алмасу жүйелерінде электрондық ақпараттарды қорғау мәселелерін шешу кезінде орталық орындарды біреуі қол қойылған электрондық құжаттың авторлықты, тұтастығын және бас тартпауын қамтамасыз ете отырып, ЭСҚ алады. Жүйені қорғаудың қажетті деңгейін және

тиімділігін таңдау мәселесі әрқашан болған [1]. Сонымен қатар, көптеген қолданбалы салаларда ақпараттық қауіпсіздікті қамтамасыз ету және олардың орындалуын қамтамасыз ету әдістерін таңдау бойынша нақты талаптар бар. Сондықтан ЭСҚ алгоритмдерін оңтайландыру жолдарын таңдап алу олардың жұмысының сапасы мен тиімділігінің әр түрлі көрсеткіштерін ескере отырып, барынша маңызды болып табылады.

RSA (Rivest, Shamir және Adleman тектерінің аббревиатурасы) – әлемде бірінші және ең танымал ЭСҚ жүйесі болды, бұл жүйенің математикалық схемасы 1977 жылы АҚШ-тың Массачусетс технологиялық институтында әзірленген. Алгоритмнің сенімділігі үлкен сандарды факторизациялау қиындығына негізделеді [2].

RSA сандық қолтаңба алгоритмінің кемшіліктері:

–RSA сандық қолтаңба жүйесі үшін n модулін, e және d кілттерін есептеу кезінде біршама қосымша шарттарды тексеру қажет, бұл іс жүзінде қиын. Осы шарттардың кез келгенінің орындалмауы, осындай орындалмауды анықтаған адамға сандық қолтаңбаны бұрмалауға мүмкіндік береді;

– ақпаратты шифрлеуге арналған АҚШ ұлттық стандарты (DES алгоритмінің), яғни 1018 деңгейінде RSA сандық қолтаңбасының бұрмалау әрекеттеріне қатысты криптотұрақтылығын қамтамасыз ету үшін, n , d және e есептеулері кезінде әр қайсысы 2512-ден кем емес бүтін сандарды пайдалану қажет, бұл сол деңгейдегі криптотұрақтылықты қамтамасыз ететін басқа алгоритмдерден 20-30% артық есептеу шығындарын талап етеді [3].

– RSA сандық қолтаңбасы мультипликативтік шабуылдарға ұшырайды. Басқаша айтқанда, RSA сандық қолтаңба алгоритмі зиянкестерге d құпия кілтті білмей-ақ, қол қойылған құжаттардың хэширлеу нәтижелерінің туындысы ретінде есептеуге болатын құжаттардың астында қолтаңбаны қалыптастыруға мүмкіндік береді.

Эль-Гамаль сұлбасы

Сұлба 1984 жылы Тахер Эль-Гамальмен ұсынылған болатын. Эль-Гамаль Диффи—Хеллман алгоритмінің бір нұсқасын ойлап тапты [4]. Ол Диффи—Хеллман жүйесін жетілдірді, сонымен қатар шифрлау үшін қолданылатын және аутентификацияны қамтамасыз ету үшін екі алгоритм ойлап тапты. RSA-ға қарағанда Эль-Гамаль алгоритмі патенттелмеген болды, және лицензияға жарналардың төлеуі қажет болмайды, сондықтан арзанырақ альтернативаға айналды.

EGSA идеясы ЭСҚ – ны бұрмалаудың практикалық мүмкін еместігін негіздеу үшін үлкен бүтін санды көбейткіштерге ыдыратудан гөрі күрделірек есептеу есебі-дискретті логарифмдеу есебі пайдаланылуы мүмкін. Бұдан басқа, Эль-Гамальға құпия кілтті анықтаусыз кейбір хабарламалармен ЭСҚ-ны бұрмалау мүмкіндігімен байланысты RSA ЭСҚ алгоритмінің айқын әлсіздігін болдырмауға мүмкіндік берді [2].

Эль-Гамаль (Elgamal) сұлбасы - түпкі өрістегі дискретті логарифмдерді есептеудің қиындығына негізделген ашық кілтті криптожүйе. Криптожүйе өзіне шифрлеу алгоритмін және сандық қолтаңба алгоритмін енгізеді. Эль-Гамаль сұлбасы АҚШ-тың (DSA) және Ресейдің (ГОСТ Р 34.10-94) электронды сандық қолтаңба стандарттары негізінде жатыр.

Эль-Гамаль жүйесінің артықшылықтары:

1. Жүйенің әрбір пайдаланушысы кездейсоқ санды таңдап, жүйенің әрбір пайдаланушысы күрделі есептеу міндеті болып табылатын екі үлкен қарапайым санды табуды талап ететін күрделі емес есептеуді орындауы тиіс.

2. Әр жолы шифрлау кезінде қысқа мерзімді кілтті қолданамыз.

3. Алгоритмнің орнықтылығының берілген деңгейінде қатысатын бүтін сандар жазбасы есептеулерде 25% қысқа болады, бұл есептеулердің күрделілігін екі есе азайтады және қолданылатын жад көлемін айтарлықтай азайтуға мүмкіндік береді [5].

DSA (Digital Signature Algorithm) - цифрлық қолтаңба алгоритмі (DSA) 1991 жылы АҚШ-та DSS (Digital Signature Standard) цифрлық қолтаңба стандартында пайдалану үшін ұсынылған. DSA алгоритмі RSA ЭЦҚ алгоритмінің дамуы болып табылады. RSA ЭЦҚ алгоритмімен салыстырғанда DSA алгоритмі бірқатар артықшылықтарға ие: жад көлемі мен қолтаңбаны есептеу уақыты қысқартылды. Қол қою және тексеру кезінде үлкен санның модулі бойынша күрделі бөлу операцияларын орындау қажеттігі кемшілік болып табылады.

Төменде осы жоғарыда аталған электронды-сандық қолтаңбаларға қысқаша шолу жасалынып, 1-ші кесте түрінде бейнеленген [3].

Кесте 1. ЭСҚ алгоритмдерін салыстыру

Алгоритм	Кілт ұзындығы	Міндеті	Ескертпе
RSA	4096 битке дейін	Шифрлеу және қолтаңба қою	Үлкен сандарды факторизациялау есебінің қиындығына негізделген; алғашқы асимметриялы алгоритмдердің бірі. Көптеген стандарттарға қосылған.
ElGamal	4096 битке дейін	Шифрлеу және қолтаңба қою	Ақырғы өрістегі дискретті логарифмдерді есептеудің қиын есебіне негізделген; тұрақтылықты төмендетпей аз уақыт аралығында кілттерді құру мүмкіндігін береді. DSS стандартының – DSA сандық қолтаңба алгоритмінде қолданылады.
DSA	1024 битке дейін	Тек қана қолтаңба қою	Ақырғы өрістегі дискретті логарифмдеудің есептеу қиындығына негізделген; АҚШ ұлттық стандарты ретінде қабылданған; құпия және құпия емес байланыстар үшін қолданылады; АҚШ-ның ҰҚА әзірлеушісі болып табылады.

Жоғарыда аталған ЭСҚ алгоритмдерімен, олардың жұмыс істеу принциптерімен танысып, талдау жасағаннан кейін, Эль-Гамаль ЭСҚ алгоритмі арқылы автоматтандырылған ақпараттық жүйелерде ақпаратты қорғауды жүзеге асыру туралы шешім қабылданды. Себебі, RSA ЭСҚ алгоритміне қарағанда біршама артықшылықтары бар. Сонымен қатар өзіндік кемшіліктері де бар екендігі анық. Осы артықшылықтары мен кемшіліктерін ескере отырып, осы Эль-Гамаль ЭСҚ алгоритмін оңтайландыру арқылы алға қойылған мақсатты жүзеге асыру көзделіп отыр.

ЭСҚ алгоритмдерін оңтайландыру жолдары қарастырылып, талдау жасалынған кейін Эль-Гамаль алгоритмінің жұмыс жылдамдығын арттыру арқылы оңтайландыру жолы болашақта жүзеге асырылатын болып шешілді. Осы орайда бұл оңтайландыру алгоритміне қысқаша шолу жасайтын болсақ; Эль-Гамаль тектес алгоритмдерді іс жүзінде пайдалану кезінде туындайтын мәселелердің бірі болып уақытша асимметрия, яғни қолтаңба құру рәсімінің есептеу қиындығына қарағанда қолтаңбаны тексеру рәсімінің есептеу қиындығы ұзақ болуы. Осы уақытша асимметрияны шешу мақсатында аталмыш Эль-Гамаль алгоритмін оңтайландыру көзделіп отыр.

Қолданылған әдебиеттер тізімі

1. Романец Ю.В., Тимофеев П.А., Шаньгин В.Ф. Защита информации в компьютерных системах и сетях. – М.: Радио и связь, 2001. – 376 с.
2. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. – М., 2002 – 816 с.
3. Жантасова Ж.З., Попова Г.В. Абаскалов А.А. Вестник КАСУ №6 – 2011, стр. 96-101 - <http://www.vestnik-kafu.info/journal/32/> (11.02.2019).
4. Оков И.Н. Криптографические системы защиты информации. – СПб.: ВУС, 2001. – 236с.
5. Панасенко С.П. Алгоритмы шифрования. Специальный справочник. – СПб.: БХВ-Петербург, 2009 – 576 с.

РАЗРАБОТКА ИНТЕРРОГАТОРОВ ДЛЯ ДАТЧИКОВ НА ОСНОВЕ ВОЛОКОННОЙ РЕШЕТКИ БРЭГГА

**^{1,2}Wojcik Waldemar, ^{1,3}Калижанова А.У., ^{1,4}Кашаганова Г.Б.,
^{1,5}Оразбеков Ж.Т., ¹Арсланова С., ¹Искакова М.Т.
e-mail: kalizhanova_aliya@mail.ru, gulzhan_k70@mail.ru**

¹*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

²*Люблинский технический университет, Люблин, Польша*

³*Алматинский университет энергетики и связи, Казахстан*

⁴*Казахско-Американский университет, Казахстан*

⁵*Казахский национальный педагогический университет имени Абая,
Казахстан*

Аннотация. В данной статье рассмотрена структура интеррогатора для волоконной решетки Брэгга и различные соединения в системе. Разработанная система может быть применена в качестве устройства для проведения анализов, таких как идентификация химических соединений и определения степени загрязнения, контроль качества сырья и продуктов в фармацевтической, пищевой, косметической промышленности, в медицинской диагностике (например, при измерении уровня сахара в крови), а также для мониторинга телекоммуникационных сетей.

Введение. Потенциальное применение волоконно-оптических датчиков (ВОД) расширяется, возникает большой интерес в поиске систем измерения, которые просты, рентабельны и показывают высокую мощность эффективности. Последнее чрезвычайно полезно при работе с мультиплексными датчиками распределяется на большие расстояния, что приводит к сильному затуханию сигнала и ограничивает количество датчиков, которые могут быть опрошены минимальным количеством единиц измерения.

Повышением производительности сенсорных систем является возможность измерения нескольких параметров, которые чрезвычайно полезны при работе с датчиками на основе волоконных решетках Брэгга (ВРБ), испытывающими

аналогичные изменения спектральных характеристик в ответ на изменения температуры или деформации.

Важные характеристики и внутренние свойства ВОД:

- устойчивость к электромагнитным помехам, что означает, что они могут быть использованы в местах, где происходит высоковольтное электричество;
- легкий вес и относительно небольшой размер, что позволяет сделать оптоволоконный датчик компактным и портативным;
- гибкость, позволяющая размещать датчик в самых узких местах;
- высокие возможности мультиплексирования, что облегчает развертывание больших сенсорных сетей.

Благодаря этим функциям ВОД в настоящее время широко используются в гражданском строительстве, аэрокосмической отрасли, нефтегазовой отрасли, судостроении, интеллектуальных конструкциях, биомедицинских устройствах, электроэнергетике и многих других. Было разработано много разных типов датчиков, основанных на различных технологиях, включая распределенные датчики, основанные на рассеянии Рамана и Бриллюэна, датчики на основе полостей Фабри-Перо и, конечно, ВРБ – является основным предметом наших исследований.

Известные системы оптического рефрактометра Аббе [1], использующие явление полного внутреннего отражения. Рефрактометр – прибор, измеряющий показатель преломления света в среде. Такие системы характеризуются тем, что для измерения коэффициента рефракции, необходимо применять: источник света LED, который имеет эмиссию волн широкого спектра, призму с высоким коэффициентом рефракции матрицы CCD и системы приобретения, основанной на микроконтроллере в сети ARM.

Применение системы датчика давления с наклонной решеткой Брэгга и методы демодуляции сигналов с этого датчика описаны в работе [2-3]. Система демодуляции сигнала состоит из перестраиваемого лазера, фотодетектора, устройства приобретения данных с фотодетектора и устройство преобразующее данные. Лазер требует перестраивание, что усложняет систему демодуляции и увеличивает значительно его стоимость. В настоящее время продолжается такие работы над системами для измерений физических величин, которые складываются с много соединенных в ряд однородных решеток Брэгга [4-8].

Применение системы рефрактора с наклонной решеткой Брэгга для измерения коэффициента преломления газов и жидкости представлено в работах [9-11]. В состав таких систем входят оптоволокно с наклонной решеткой Брэгга, а его спектральный отклик зависит как от коэффициента преломления, так и от параметров источника света. Изменения спектра, которые возникают из-за изменения коэффициента преломления меди анализируется с использованием спектрометра, а этот прибор дорогой и сложный. Добавочно, спектральные характеристики требуют преобразования, так чтобы возможно было определение на их основе коэффициента преломление света.

Система измерения длины волны сигнала, отраженного от нескольких однородных измерительных решеток Брэгга [12], использует второе оптоволокно, на котором записано ряд референционных решеток, которые позволяют точно определить длины волн, отраженных через измерительную решетку Брэгга.

Известен метод и система интеррогации сигнала с однородной решеткой Брэгга [13-15]. В этом методе используется широкополосный источник света, соединенный с головкой датчика. В работе [16] описан метод измерения индекса рефракции, который использует измерение интенсивности изменений спектра

наклонной решетки Брэгга. В работе [17] представлен метод назначения индекса рефракции на основе измерения интенсивности трансмигрированной мощности через наклонную решетку Брэгга. В этом методе используется измерение меняющейся мощности с помощью анализатора оптического спектра. Система состоит из широкополосного источника излучения наклонной решеткой Брэгга, которая работает как преобразователь и анализатор оптического спектра. В этой системе не используется ни одного оптического элемента, который фильтрует меняющийся спектр наклонной решетки Брэгга.

Однако, в этих работах рассматриваются только отдельные виды волоконных решеток Брэгга, датчики на основе решеток, а вопросы измерения изменения показателя преломления на основе наклонных решеток Брэгга и влияние на них различных факторов не исследованы.

Интеррогаторы – приборы, с помощью которых опрашиваются ВРБ.

Опрашивающий обычно измеряет сдвиг длины волны Брэгга, который затем преобразуется в измеряемые данные (например, деформация, температура, давление так далее.). Параметры интеррогатора, такие как разрешение, скорость, точность и линейность могут существенно повлиять производительность датчика, например, для разрешения температуры и деформации изменение на $\sim 0,1^\circ\text{C}$ и 1мкм требует разрешения по длине волны 1нм .

Наиболее часто используемое решение в коммерчески доступных интеррогаторах для датчиков с волоконно-оптическими решетками Брэгга, состоит в определении сдвига длины волны Брэгга, определяемого по отражению обычного спектра пропускания (в форме спектра кривой Гаусса) с высоким коэффициентом отражения.

Интеррогатор, созданный как специализированная система, может характеризоваться гораздо большей универсальностью - особенно при использовании нетипичных Брэгговских структур (кроме обычных решеток со спектром, подобным кривой Гаусса). Прежде всего, специальная система может работать, анализируя смещение спектра Брэгга и изменяя общую часть спектров двух фотонных элементов, то есть создает амплитудную систему. Благодаря этому можно использовать решетки, отличные от обычных, и при правильной конфигурации системы можно анализировать мантийные моды наклонных решеток. В более сложной форме также возможно получить способность системы считывать изменения измеренной величины для обоих типов датчиков одновременно.

На рисунке 1 показано расположение интеррогатора, который мы создали в Люблинском техническом университете на схеме всей измерительной системы.

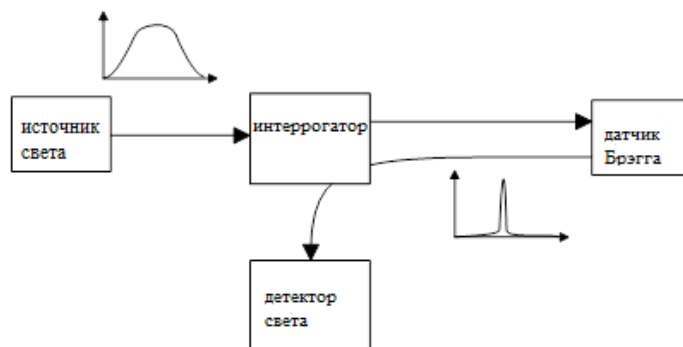


Рис. 1. Общая схема системы интеррогатора для оптоволоконных датчиков

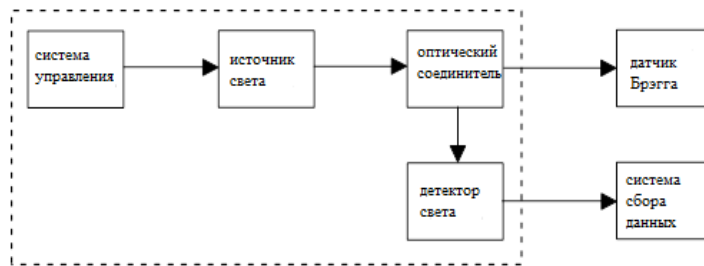


Рис. 2. Принципиальная блок-схема всей системы интеррогатора

Можно предложить различные реализации систем интеррогации, например, те, которые приведены на рисунке 3.

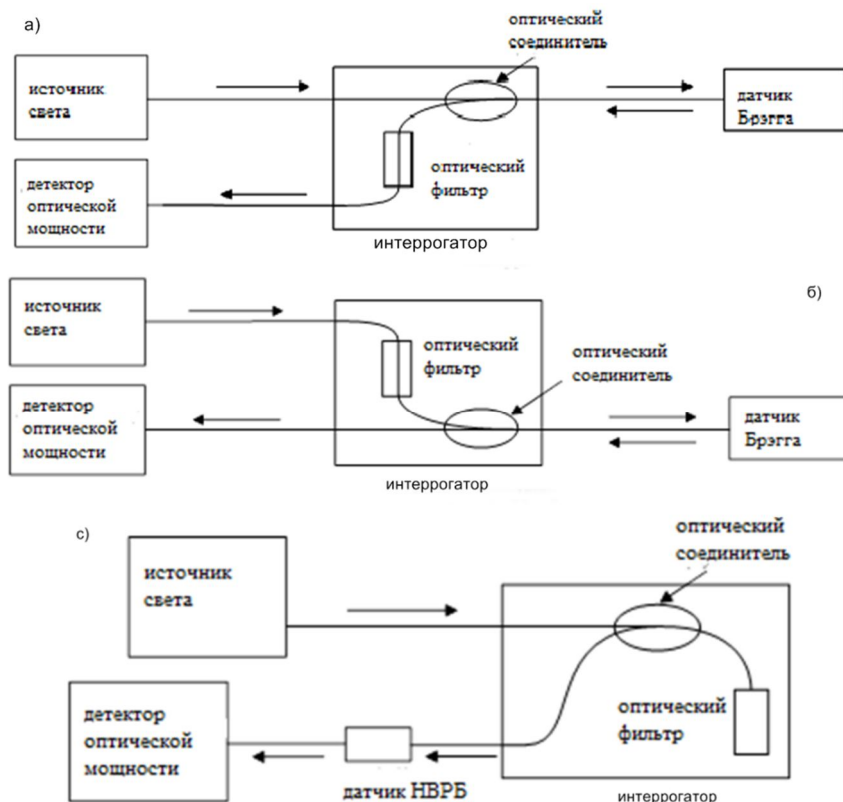


Рис. 3. Концептуальные конфигурации систем интеррогатора с использованием пассивных волоконно-оптических элементов
 а) и б) с использованием обычных датчиков ВРБ
 в) с наклонной волоконной решеткой Брэгга (НВРБ)

Нами разработанная система интеррогатора создана по макету 3а рисунка 3.

Интеррогатору также необходим источник света, детектор оптической мощности и решетка Брэгга, которая измеряет заданную величину (напряжение, температура, показатель преломления, деформация, удлинение, давление и т. д.). Благодаря сдвигу спектральных характеристик решетки Брэгга можно измерить с помощью обычного фотоприемника, который обходится не так дорого, по сравнению с неподвижного и относительно медленного анализатора спектра.

На рисунке 4 представлена схема системы интеррогатора для волоконно-оптического датчика на основе НВРБ.

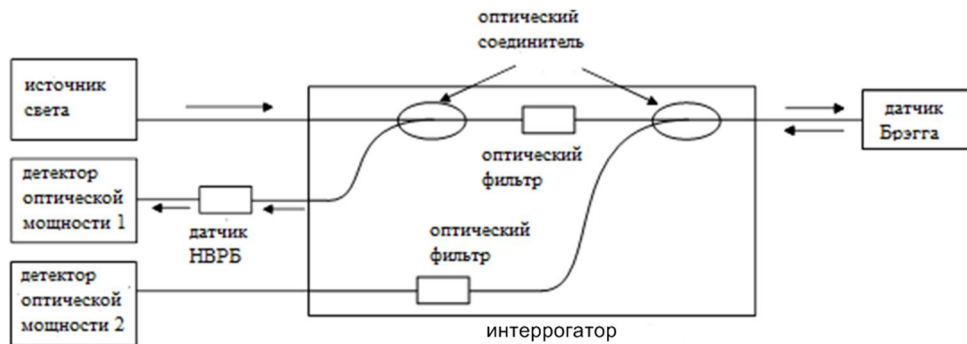


Рис. 4. Концепция системы интеррогатора для волоконно-оптического датчика на основе НВРБ

Систему интеррогатора для волоконно-оптического датчика на основе НВРБ мы будем рассматривать в дальнейших исследованиях.

Преимуществом применения разработанной системы интеррогации оптоволоконного рефрактометра с наклонной решеткой Брэгга является значительно упрощенная система для измерения коэффициента преломления среды, которая не требует применения спектрофотометров и анализаторов оптического спектра, а также не требует использования алгоритмов анализа оптического спектра. Существенная особенность разработанного метода заключается в том, что можно сделать измерения независимо от влияния внешней температуры и электромагнитного поля на месте измерения, что достигается за счет того, что решетки записаны на однородном оптоволокне. Применение этой системы устранил проблему флуктуации мощности источника света, что достигается с помощью показателя коэффициента преломления среды.

Заключение

Проведенный анализ и исследования определяют актуальность научно-технической проблемы улучшения функциональных возможностей систем интеррогации сигналов с оптоволоконным рефрактометром.

Разработанная система может быть применена в качестве устройства для проведения анализов, таких как идентификация химических соединений и определения степени загрязнения, контроль качества сырья и продуктов в фармацевтической, пищевой, косметической промышленности, в медицинской диагностике (например, при измерении уровня сахара в крови), а также для мониторинга телекоммуникационных сетей. Рефрактометр на основе наклонных решеток Брэгга будет использоваться для измерения медицинских величин, таких как удельный вес мочи, показатель преломления и содержание белка в сыворотке, для количественного анализа, например, концентрация сахара в вине, этиловый спирт в пиве, воды в меде, соли в морской воде.

Метод измерения коэффициента показателя преломления с использованием наклонных решеток Брэгга, также будет полезен при исследованиях физико-химических процессов, проведенных для объяснения эффектов влияния электромагнитных волн на электронную микроскопию атомов и молекул (например, молярное преломление и его компоненты).

Экспериментальные исследования проводились в лабораториях Оптоэлектроники факультета электротехники и компьютерных наук Люблинского технического университета в рамках проекта ГФ №АР05132778 «Исследование и разработка системы

Литература

1. Patent. Chiny, nr CN203772726. Optical system of multi-wavelength Abbe refractometer. Shanghai Instr Physical Optics INSTR CO LTD. Opubl.13.08.2014.
2. Применение оптического волокна в качестве датчиков деформации в полимерных композиционных материалах / Е.К. Каблов, Д.В. Сиваков, И.Н. Гуляев, К.В. Сорокин, М.Ю. Федотов, Е.М. Дианов, С.А. Васильев, О.И. Медведков // Все материалы. Энциклопедический справочник. – 2010. – № 3. – С. 10–15.
3. Udd E. Fiber optic sensors: an introduction for engineers and scientists / E. Udd, W. B. Spillman – John Wiley & Sons, 2011. – 498с.
4. Cieszczyk S, Harasim D, Kisała P. Novel twist measurement method based on TFBG and fully optical ratiometric interrogation. Journal Sensors and Actuators A: Physical 2018, pp. 18-22
5. Ciężczyk Sławomir, Kisała Piotr. Układ interogacji sygnałów z refraktometru światłowodowego [Patent nr (11) 230198 (21) 420259] // Wiadomości Urzędu Patentowego - 2018, 10, s. 3702-3702 [MNiSW: 30].
6. S.A. Wade, D.P. Attard, P.R. Stoddart, Analysis of transmission mode of a matched fiber grating interrogation scheme, Appl. Opt. 49 (24) (2010) 4498–4505.
7. Z. Zhang, L. Yan, W. Pan, B. Luo, P. Wang, L. Guo, W. Zhou, Sensitivity enhancement of strain sensing utilizing a differential pair of fiber Bragg gratings, Sensors 12 (4) (2012) 3891–3900.
8. Y. Zhan, M. Yu, J. Pei, X. Yang, S. Xiang, A linearity interrogation technique with enlarged dynamic range for fiber Bragg grating sensing, Opt. Commun. 283 (18) (2010) 3428–3433.
9. J. Pei, X. Yang, Y. Zhan, R. Zhu, S. Xiang, On a fiber sensor system with the capacity of cross-sensitivity discrimination, Optik 119 (12) (2008) 565–570.
10. C.-F. Chan, C. Chen, A. Jafari, A. Laronche, D. J. Thomson, and J. Albert, “Optical fiber refractometer using narrowband cladding-mode resonance shifts,” Appl. Opt., vol. 46, no. 7, pp. 1142–1149, Feb. 2007.
11. F. Liu et al., “High-sensitive and temperature-self-calibrated tilted fiber grating biological sensing probe,” Chin. Sci. Bull., vol. 58, no. 21, pp. 2611–2615, 2013.
12. G. Laffont and P. Ferdinand, “Tilted short-period fibre-Bragg-grating-induced coupling to cladding modes for accurate refractometry,” Meas. Sci. Technol., vol. 12, no. 7, p. 765, 2001.
13. Y. C. Lu et al., “Polarization effects in tilted fiber Bragg grating refractometers,” J. Lightw. Technol., vol. 28, no. 11, pp. 1677–1684, Jun. 1, 2010.
14. S. Cieszczyk, D. Harasim, and P. Kisała A Novel Simple TFBG Spectrum Demodulation Method for RI Quantification IEEE Photonics technology letters, vol. 29, no. 24, December 15, 2017 pp 2264-2267
15. Zheng J., Dong X., Ji J., Su H., Shum P.P., Power-referenced refractometer with tilted fiber Bragg grating cascaded by chirped grating. Optical Communication 312, 106-109 (2014)
16. Miao, Y.P., Liu B., Refractive index sensor based on measuring the transmission power of tilted fiber Bragg grating, OpticalFiber Technology 15(3), 233-236 (2009).
17. Cuadrado-Laborde C. Current trends in short- and long-period fiber gratings // InTech. – 2013. – URL: intechopen.com/books/current-trends-in-short-and-long-period-fiber-gratings.

БИОМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ЧЕЛОВЕКА ПО ОТПЕЧАТКАМ ПАЛЬЦЕВ

Алиаскар¹ М.С., Айпанов¹ Ш.А., Тусупова¹ С.А.,
Карымсакова¹ Н.Т., Амирханов¹ Б.С.
e-mail: 8701993011@mail.ru

¹Институт информационных и вычислительных технологий КН МОН РК,
Алматы, Казахстан

Аннотация. Для идентификации личности отпечатки пальцев были золотым стандартом в криминалистике на протяжении более ста лет. Следы от пальцев и отпечатки пальцев объединяются, чтобы предоставить наиболее мощные средства идентификации личности, доступные для полиции и судов. Основными образцами являются петли, завитки и арки, которые можно найти в отпечатках пальцев. Стандартная методология, используемая экспертами по отпечаткам пальцев для проведения исследований на гребень трения, называется ACE-V (*analysis, comparison, evaluation and verification*) для анализа, сравнения, оценки и проверки, которые являются четырьмя основными этапами, используемыми в этом процессе. Восстановленные отпечатки можно вручную сравнить с известными записями до смерти или выполнить поиск с помощью, автоматизированной системы отпечатков пальцев (*AFIS - automated fingerprint system*) для проверки или установления личности.

Ключевые слова: криминалистика, отпечатков пальцев, идентификация личности.

Введение

Отпечатки пальцев были золотым стандартом для идентификации личности в криминалистике на протяжении более ста лет. Наука идентификации отпечатков пальцев с течением времени развивалась от раннего использования для обозначения деловых операций в древнем Вавилоне, до их использования сегодня в качестве основной технологии в биометрических устройствах безопасности и в качестве научных доказательств в судах общей юрисдикции во всем мире. Отпечатки пальцев — это узоры, образованные приподнятыми папиллярными выступами на кончиках пальцев, которые содержат ряды пор, которые соединяются с потовыми железами. После многих лет исследований и разработок было установлено, что папиллярные гребни на пальцах и ладонях рук, а также на подошвах и пальцах ног остаются на протяжении всей жизни в соответствии с формой, в которой они были созданы природой, если только они не симметричны, постоянно страдают от глубокой травмы [1]. Эти гребни, которые образуются в течение первых нескольких месяцев жизни плода, не только остаются неизменными в течение жизни, но и после смерти, похоже, обладают своеобразным свойством переживать все другие узнаваемые признаки тела.

Отпечатки пальцев и следы от пальцев объединяются, чтобы предоставить наиболее мощные средства идентификации личности, доступные для полиции и судов. То, как это стало таким мощным инструментом, вытекает из осознания того, что рисунки кожи гребня и их детали уникальны, неизменны, универсальны, легко классифицируются и оставляют следы на любом объекте, обработанном голыми руками. Понимание ценности отпечатков пальцев привело к исследованию методов

обнаружения, оперативного и стратегического использования отпечатков пальцев. Отпечатки пальцев образуют определенные узоры, которые, похоже, имеют общее сходство по форме и дизайну. Их сходство позволило отпечатку пальца первопроходца создать устройство систем классификации по отпечаткам пальцев; теперь они позволяют на практике использовать большие коллекции отпечатков, которые можно подавать по определенной схеме [2].

Основы отпечатков пальцев

Использование отпечатков пальцев в криминалистике основано на нескольких фундаментальных принципах. Во-первых, вероятность найти двух людей с одинаковыми отпечатками пальцев очень мала. На самом деле, никогда не было найдено двух одинаковых отпечатков пальцев. Гальтон подсчитал, что вероятность нахождения идентичных отпечатков была 1 на 64 миллиона. Вторым принципом является то, что отпечатки пальцев человека не меняются со временем. Структура гребней на кончиках пальцев, ладонях и подошвах человека при рождении остается неизменной до самой смерти. Следовательно, детектив может быть уверен, что отпечатки пальцев преступника останутся без изменений до самой смерти. Наконец, существует достаточно сходства в узорах гребней на пальцах людей, которые можно классифицировать. Основными образцами отпечатков пальцев являются петли, завитки и арки, которые можно найти в отпечатках пальцев. Приблизительно от 60 до 65 процентов населения имеют петлевые структуры, от 30 до 35 процентов имеют завитки, и только около 5 процентов имеют арки. Арки могут быть либо простыми, либо тентовыми, а обороты могут быть классифицированы как центральный карман, боковой карман, двойняшки и случайные (Рис. 1).

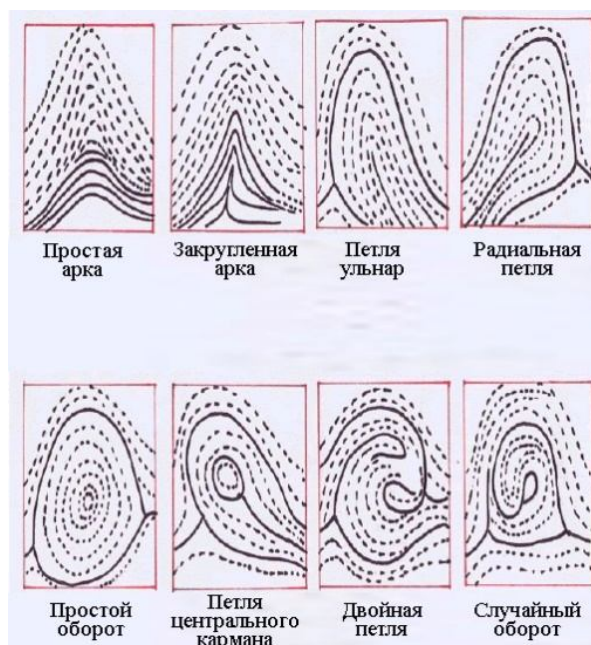


Рис. 1 Типы узоров пальцев: систематические рисунки с жирным шрифтом.

Детальное исследование кожи гребня трения также показывает, что путь гребня, в большинстве случаев, не является непрерывным по всей поверхности пальца. Некоторые гребни, называемые конечными гребнями, будут течь и внезапно заканчиваться, в то время как другие гребни, называемые

разделительными гребнями или раздвоениями, будут течь и разделяться на два отдельных и отдельных гребня. Кроме того, некоторые гряды имеют большую длину и называются точками [3]. Эти события гребня обычно называют характеристиками или мелочами, и их пространственное отношение друг к другу в отпечатке трения представляет собой основу для сравнения и идентификации отпечатков пальцев. Арка имеет выступы, которые входят с одной стороны узора, делают волну посередине и проходят через противоположную сторону, с которой они вошли. В круче, гребни трения имеют тенденцию иметь круговое или спиральное течение гребня.

Для эксперта важно отметить гребень отпечатка для ориентационных целей и распознавание основных областей, которые в конечном итоге помогут в процессе идентификации. Хотя одна конфигурация шаблона не может использоваться для индивидуализации, она может использоваться для исключительных решений, принятых экзаменатором [3].

Отпечатки пальцев делятся на три категории: скрытые, известные и пластиковые отпечатки. Скрытый отпечаток пальца - это двумерное воспроизведение выступов трения пальца на объекте с помощью пота, масел или других загрязнений, которые покрывают поверхность выступов, когда палец касается предмета. Эти типы отпечатков обычно должны быть видны с помощью судебно-медицинской технологии, такой как альтернативные источники света, химические методы или порошки отпечатков пальцев. В некоторых случаях скрытые отпечатки могут быть визуализированы без использования каких-либо методов обработки и называются явными отпечатками. Скрытые впечатления откладываются случайно и обычно носят фрагментарный характер с различной степенью качества. Альтернативно, известный отпечаток пальца - это намеренное воспроизведение выступов трения пальца на карте отпечатка пальца или соответствующей контрастной поверхности [4]. Известный пример может быть записан с использованием ряда стандартных методов, включая чернила черного принтера, без чернил/химические методы, и сканирование в режиме реального времени, которое представляет собой компьютерную систему, которая создает цифровые изображения отпечатков пальцев путем сканирования пальцев. Наконец, пластиковый отпечаток - это отпечаток, оставленный на ковкой подложке, такой как воск или замазка, которая сохраняет изображение расположения гребня трения [4].

Сравнение и идентификация моделей трения

Хотя терминология, используемая для описания процесса идентификации отпечатков пальцев, менялась на протяжении многих лет, базовая методология, используемая судебно-медицинскими экспертами, оставалась относительно неизменной. Однако один аспект, который изменился - это повышение осведомленности об основополагающей научной основе для идентификации отпечатков пальцев. Это изменение привело к стандартизации процесса идентификации, основанного на обширном исследовании бывшего сержанта Королевской канадской конной полиции Дэвида Р. Ашбо [5], которое сосредоточено вокруг количественной и качественной философии для исследования отпечатков пальцев, называемого риджологией. Риджеология - это целостный подход, который фокусируется на биологической уникальности гребней трения и включает в себя последовательное изучение особенностей и пространственных отношений гребней, отмечая качество и количество оцениваемой информации для целей идентификации. Этот экзамен проводится с

использованием методологии, которая включает в себя надежные научные протоколы и практики, позволяющие делать точные и повторяемые выводы, которые соответствуют строгим научным стандартам. Стандартная методология, используемая экспертами по отпечаткам пальцев для проведения исследований на гребне трения, называется ACE-V для анализа, сравнения, оценки и проверки, которые являются четырьмя основными этапами, используемыми в этом процессе.

Анализ фокусируется на изучении количества и качества информации, представленной в печатном виде, которая может быть разбита на три уровня детализации [6]. Уровень 1 подробно относится к общему течению гребня и типу рисунка отпечатка. Детали уровня 2 относятся к траектории гребня, которая соответствует пространственной взаимосвязи гребней и их характеристик в отпечатке. Уровень 3 относится к отдельным атрибутам гребня, которые включают в себя формы гребня и структуру/расположение пор в отпечатке. Эксперт по отпечаткам пальцев должен учитывать различные факторы качества, такие как искажение, которые могут повлиять на достоверность наблюдаемой информации при определении пригодности отпечатка для целей сравнения. Информация, присутствующая в отпечатке скрытого или худшего качества, всегда проверяется в первую очередь, а затем проверяется отпечаток известного или самого высокого качества.

Сравнение оттисков гребня трения - это параллельная оценка информации, проанализированной в обоих отпечатках [6]. Печать с латентным или худшим качеством сравнивается с печатью известного или лучшего качества, чтобы минимизировать когнитивные искажения. Экзаменатор сначала оценивает информацию уровня 1 из анализа скрытой распечатки и сравнивает ее с информацией, полученной из анализа известной распечатки. Если информация совпадает, то эксперт затем оценивает информацию уровня 2 из анализа скрытой распечатки и сравнивает ее с информацией, полученной в результате анализа известной распечатки. Сравнение - это не простое упражнение «подсчет очков»; на самом деле, нет научной основы для минимального точечного порога или определенного количества характеристик, которые должны совпадать в двух отпечатках для решения, которое должно быть принято экспертом [6]. Информация уровня 3 обычно отмечается при оценке детализации уровня 2 и, если она видна, также сравнивается. Эти сравнительные измерения начинаются с фокальной точки, выбранной экзаменатором, и последовательно проходят через гребни всего отпечатка. Чтобы информация соответствовала обоим отпечаткам, путь гребня должен иметь надлежащее качество, четкость, относительное положение и иметь одинаковые отношения единиц. Решение об идентификации может быть принято только в том случае, если скрытые и известные отпечатки совпадают [6]. Из-за податливости поверхности трения и других факторов окружающей среды, отпечатки гребня трения одного и того же пальца никогда не будут выглядеть одинаково.

Оценка включает в себя принятие решения на основе результатов анализа и сравнения этапов процесса идентификации. Есть три возможных вывода, которые могут быть сделаны экспертом в соответствии с определением Научной Рабочей Группы по Исследованиям и Технологиям Анализа Гребня Трения (SWGFAST) [7]. Первый вывод - это индивидуализация (идентификация) и определение того, что достаточное количество информации, представленной в двух показах совпадает, что означает, что они получены из одного источника. Вторым выводом является исключение и определение того, что информация, представленная в двух показах, не совпадает, а это означает, что они не из одного источника. Третий вывод

является неубедительным решением и состоит в том, что невозможно сделать окончательное сравнение из-за не качественности или отсутствия сопоставимой области в известном образце.

Проверка - последний шаг в методологии ACE-V. Хотя проверка технически не является частью процесса идентификации, она служит формой экспертной оценки, обеспечивающей надежные и точные результаты [7]. Все индивидуализации, сделанные экспертом по отпечаткам пальцев, проверяются путем независимой экспертизы идентифицированных отпечатков вторым квалифицированным экспертом по скрытой печати в качестве механизма обеспечения качества. Проверка исключения или неокончательных решений также может быть выполнена, но это не требуется SWGFAST. Для использования отпечатков пальцев в качестве средства индивидуализации существуют два основных положения, которые также составляют основу процесса идентификации гребня трения [7]. Эти предпосылки касаются индивидуальности и стойкости трения кожи, которые были научно подтверждены с течением времени благодаря научным исследованиям и работе экспертов в области отпечатков пальцев.

Индивидуальность относится к тому факту, что отпечатки пальцев являются уникальными; ни одна из двух областей шкурки трения одинакова, даже у идентичных близнецов. Основой для этого утверждения является эмбриология и генетика человека, начиная с развития плода. Физиология кожи гребня трения начинается с развития подушечек, представляющих собой выпуклости ткани, которые начинают формироваться на кончиках пальцев примерно на восьмой неделе беременности. Степень сложности плоских подушечек (их размер, форма и расположение на пальце) значительно влияет на течение гребня или уровень детализации 1 [7]. Эти соляные подушечки регрессируют или область впитывается обратно в палец примерно на десятой или одиннадцатой неделе беременности, когда начинают формироваться гребни трения. Первичные гребни развиваются первыми, затем следует развитие вторичных гребней или появление борозд между папиллярными гребнями. Случайное развитие гребней трения и деталей их уровня 2 и уровня 3 является результатом бесконечного числа факторов окружающей среды, хотя большинство этих видов деятельности имеют генетический компонент [7]. Конечным результатом этих генетических и экологических различий во время формирования гребня трения является полная биологическая уникальность вплоть до структуры одного гребня.

Стойкость относится к тому факту, что гребни трения являются постоянными и остаются постоянными в течение всей жизни человека, вплоть до разложения после смерти, если иное не затронуто случайной травмой или умышленным увечьем. Основой для этого утверждения является анатомия человека и гистология кожи. Когда тело избавляется от отмерших клеток кожи, они заменяются новыми клетками кожи, которые образуются из нижнего или базального слоя эпидермиса [7]. Клетки, соединенные через клеточные соединения, заменяются одинаковым образом в течение всей жизни, если не происходит рубцевание. Таким образом, базальный слой действует как неизменная корневая система, которая является основой для постоянства гребней трения и соответствующих им деталей уровня 1, 2 и 3.

Заключение

Идентификация по отпечатку пальца - самая старая судебная дисциплина, известная человеку. С течением времени отпечатки пальцев оказались наиболее быстрым, надежным и экономически эффективным средством выявления

неизвестных умерших людей, особенно в условиях массовых бедствий. Восстановленные отпечатки можно вручную сравнить с известными записями до смерти или выполнить поиск с помощью автоматизированной системы отпечатков пальцев (AFIS) для проверки или установления личности. Идентификация останков по отпечаткам пальцев выполняет самую важную и трудную задачу судебно-медицинской идентификации: своевременное и точное уведомление семей о судьбе их близкого человека [11].

Работа выполнена за счет средств грантового финансирования научных исследований на 2018-2020 годы по проекту AP05131027 «Разработка биометрических методов и средств защиты информации».

Литература

1. Galton F (1892) Fingerprints. London: Macmillan & Co.
2. Herschel WJ (1916) The origin of fingerprints. London: Oxford University Press.
3. Henry ER (1922) Classification and uses of fingerprints. (5th edn), London: HM Stationary Office.
4. Browne DG, Brock (1954) A Fingerprints: Fifty years of scientific crime detection. New York: EP Dutton & Co.
5. Ashbaugh DR (1999) Quantitative and qualitative friction ridge analysis: An introduction to basic and advanced ridgeology. Boca Raton, FL: CRC Press LLC.
6. Symposium (1995) Report: Israel National Police: International symposium on Fingerprint Detection and identification. J for Ident 45: 578-584.
7. Wertheim K, Maceo A (2002) The critical stage of friction ridge and pattern formation. J for Ident 52: 35-85.
8. Dror IE, Charlton P, Peron AE (2006) Contextual information renders experts vulnerable to making erroneous identifications. Forensic Science International 156: 74-78.
9. Snady LZ (2005) Fingerprint evidence. L Law & Policy 143: 13-20.
10. Vokey JR, Tangen JM, Cole SA (2009) On the preliminary psychophysics of fingerprint identification. Quart J Exp Psycho 62: 1023-1040.
11. Senn DR, Stimson PG (2010) Forensic Dentistry. New York: CRC Press.

О ПРИМЕНЕНИИ ОДНОГО СПОСОБА КЛАСТЕРИЗАЦИИ ДЛЯ ПОИСКА МИНИМАЛЬНОЙ ЭНЕРГИИ ПЛОСКОГО КРИСТАЛЛА

Амирханова Г.А., Дуйсенбаева А.Ж.

e-mail: gulshat.aa@gmail.com, aigerim.95.05@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** В работе рассматривается три метода кластеризации для найденных векторов геометрических характеристик плоской кристаллической решетки. Первые два, K-means и Mini Batch K-means, применяются для решения задач с большим количеством образцов. Третий метод, Affinity Propagation,*

самостоятельно выбирает количество кластеров, на основе предоставленных данных. Для улучшения качества кластеризации методом *AffinityPropagation*, предложена матрица сходства, учитывающая симметрии фрагмента кристаллической решетки. Проведено экспериментальное сравнение методов, которое показало преимущество третьего метода.

Ключевые слова: *k-means*, *mini batch k-means*, *AffinityPropagation*, кластеризация, двумерные кристаллы, графеноподобные структуры.

Введение

Исследование плоских кристаллов началось еще в 1912 году, когда немецкий физик М. Лауэ сформулировал условия возникновения дифракционных максимумов при прохождении рентгеновских лучей через кристаллы [1]. В работе рассматривается задача нахождения энергетического минимума фрагмента кристаллической решетки. Решение данной задачи позволит получать новые материалы.

Для вычисления серии локальных минимумов применяется метод глобальной оптимизации – метод мултистарта. Далее решается задача кластеризации полученных локальных минимумов. Для этого применяются методы из библиотеки *scikit-learn* [2], которые осуществляют группировку и поиск центров в совокупности атомов, заданных координатами в двумерном пространстве. В работе рассматривается постановка задачи с фрагментом фиксированной длины.

В качестве рекомендуемого метода кластеризации выбран метод *AffinityPropagation* [3]. Для ускорения вычисления матрицы сходства применяется библиотека *mpi4py* [4-6].

1. Постановка задачи

Для задачи исследования свойств многослойного кусочно-однородного материала, модель которого предложена в [7, 8], был сгенерирован фрагмент с кристаллической решеткой, в которой было задано количество слоев, количество атомов и их тип, радиус отсечения и т.д.

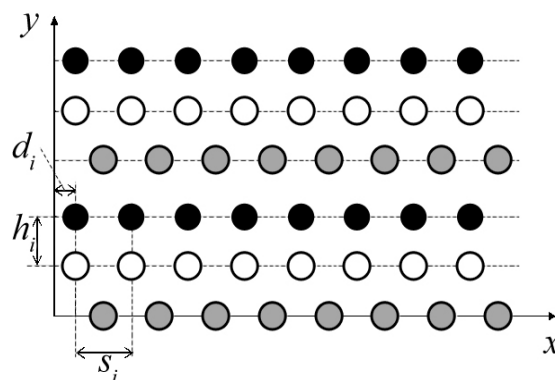


Рис.1 Двумерная модель материала

В рамках данной модели положение атомов определяется следующими параметрами:

$h_i, i = 1, \dots, K$ – расстояние между слоем с номером i и предыдущим слоем;

$d_i, i = 1, \dots, K$ – смещение первого атома в слое i с положительной абсциссой относительно нулевой отметки;

$s_i, i = 1, \dots, K$ – расстояние между атомами в слое i .

$d_i = 0, d_i < s_i, i = 2, \dots, K$

Совокупность значений перечисленных параметров будем называть конфигурацией, которая определяется вектором типов атомов длины k и вектором геометрических характеристик $z = (h_1, d_1, s_1, \dots, h_k, d_k, s_k)$ длины $3k$.

В работах [9, 10] была осуществлена минимизация $E(x)$ – энергии атомов, входящих в фрагмент. Для этого был использован метод Розенброка, и получена серия локальных минимумов и соответствующих им конфигураций.

В качестве методов кластеризации рассматриваются:

1. K-means
2. Mini Batch K-Means
3. AffinityPropagation

2. Метод K-Means

В алгоритме K-means данные кластеризуются, пытаясь разбить образцы в n групп с одинаковым отклонением, минимизируя критерий, известный как *инерция* или внутрикластерная сумма квадратов. Этот алгоритм требует указывать количество кластеров. Он хорошо масштабируется для большого количества образцов и использовался в широком диапазоне областей применения во множестве разных сфер.

Алгоритм K-means делит множество N образцов X в K непересекающихся кластеров C , каждый из которых описывается как среднее μ_j образцов в кластере. Эти средние обычно называются кластерными центрами; они, как правило, не являются X , хотя существуют в одном пространстве.

Алгоритм K-means предназначен для выбора центров, которые минимизируют инерцию или внутрикластерный критерий суммы квадратов:

$$\sum_{i=0}^n \min_{\mu_j \in C} (\|x_i - \mu_j\|^2)$$

Инерция может быть определена как мера того, насколько внутренне когерентны кластеры. Она обладает рядом недостатков:

– Инерция предполагает, что кластеры дугообразные и изотропные, что не всегда так. Она плохо реагирует на вытянутые скопления или множества неправильной формы.

– Инерция не является нормализованной метрикой: мы просто знаем, что более низкие значения лучше, а ноль оптимален. Но в пространствах с высоким количеством измерений евклидовы расстояния имеют тенденцию становиться раздутыми (т.н. «проклятие размерности»). Выполнение алгоритма уменьшения размерности, такого как анализ главных компонент (PCA) перед кластеризацией K-means, может решить эту проблему и ускорить вычисления.

3. Метод Mini Batch K-Means

MiniBatchKMeans — это вариант алгоритма KMeans, который использует мини-пакеты для сокращения времени вычислений, в то же время пытаясь оптимизировать ту же целевую функцию. Мини-пакеты — это подмножества входных данных, которые выбираются случайным образом в каждой обучающей итерации. Эти мини-пакеты значительно сокращают объем вычислений, необходимых для сходимости к локальному решению. В отличие от других

алгоритмов, которые уменьшают время сходимости k -means, мини-пакетные k -means дают результаты, которые, как правило, лишь немного хуже, чем стандартный алгоритм.

Данный алгоритм перебирает два основных шага, аналогично базовому k -means. На первом этапе образцы b выбираются случайным образом из набора данных, чтобы сформировать мини-пакет. Затем они назначаются ближайшему центру. На втором этапе обновляются центры. В отличие от k -means, это делается для каждой выборки. Для каждой выборки в мини-пакете назначенный центр обновляется путем взятия среднего значения потока выборки и всех предыдущих выборок, назначенных этому центру. Это приводит к уменьшению скорости изменения центра с течением времени. Эти шаги выполняются до тех пор, пока не будет достигнута сходимость или заданное количество итераций.

MiniBatchKMeans сходится быстрее, чем KMeans, но качество результатов снижено. На практике эта разница в качестве оказалась небольшой.

4. Метод Affinity Propagation

Affinity Propagation создает кластеры путем отправки сообщений между парами образцов до сходимости. Набор данных затем описывается с использованием небольшого числа образцов, которые определены как наиболее представительные из других образцов. Сообщения, отправленные между парами, представляют пригодность для одного образца быть образцом другого, который обновляется в ответ на значения из других пар. Это обновление происходит итеративно до сходимости, после чего выбираются окончательные образцы и, следовательно, происходит окончательная кластеризация.

Affinity Propagation может быть интересным, поскольку данный метод самостоятельно выбирает количество кластеров на основе предоставленных данных. Для этой цели служат два важных параметра — это *свойство*, который контролирует, сколько образцов используется, и *коэффициент демпфирования*, который заглушает сообщения об ответственности и доступности, чтобы избежать численных колебаний при обновлении этих сообщений.

Основным недостатком Affinity Propagation является его сложность. Алгоритм имеет временную сложность порядка $O(N^2T)$, где N — количество образцов и T — количество итераций до сходимости. Кроме того, сложность памяти порядка $O(N^2)$ если используется плотная матрица сходства, но сводимая, если используется разреженная матрица сходства. Это делает Affinity Propagation наиболее подходящим для небольших и средних наборов данных.

5. Индекс Дэвиса-Боулдина

В качестве метрики для оценки метода кластеризации использован индекс Дэвиса-Боулдина.

Если основные метки истинности неизвестны, индекс Дэвиса-Боулдина можно использовать для оценки модели, причем более низкий индекс Дэвиса-Боулдина относится к модели с лучшим разделением между кластерами.

Индекс определяется как среднее сходство между каждым кластером C_i для $i = 1, \dots, K$ и его самым похожим C_j . В контексте этого индекса сходство определяется как мера R_{ij} которая меняет

- s_i — среднее расстояние между каждой точкой кластера i и центра этого кластера, также известно как диаметр кластера.
- d_{ij} — расстояние между центрами кластеров i и j .

Простейшим является случай, когда R_{ij} является неотрицательной и симметричной:

$$R_{ij} = \frac{s_i + s_j}{d_{ij}}$$

Тогда индекс Дэвиса-Боулдина определяется как:

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} R_{ij}$$

При этом следует заметить, что $DB = 0$ — минимально возможная оценка. Значения ближе к нулю указывают на лучшее разделение.

При обычном использовании индекс Дэвиса-Болдина применяется к результатам кластерного анализа.

6. Результаты численных экспериментов

Мы провели численный эксперимент на компьютере с процессором Intel Core i7-4710HQ с частотой 2.50 GHz, ОЗУ 16 Гб, 64-разрядная система. По графику видно, что с увеличением количества выбранных кластеров качество кластеризации повышается незначительно, притом, что время, требуемое для вычислений, возрастает (рис. 2)

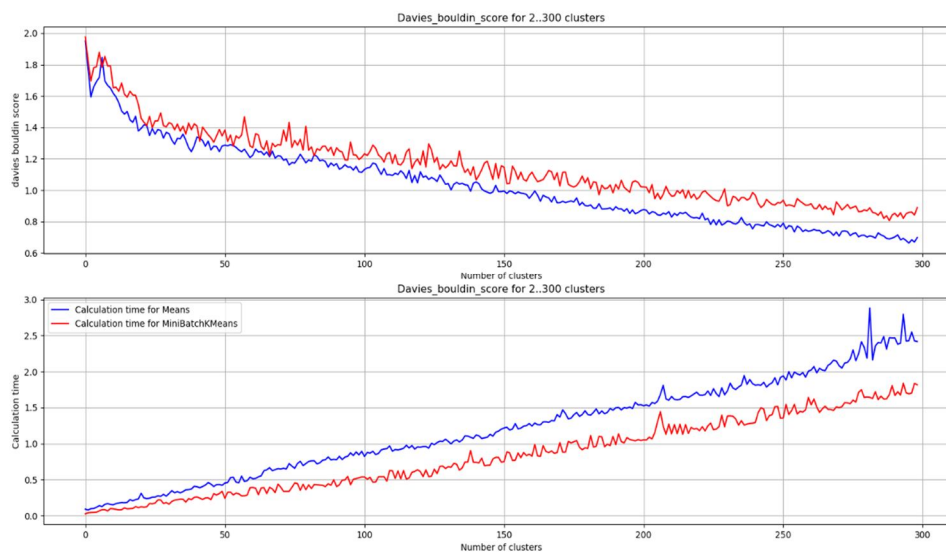


Рис.2 Зависимость индекса Дэвиса-Боулдина и времени вычисления кластеров от количества кластеров

Выбрав количество кластеров, равное 60, мы сравнили эффективность трех методов кластеризации по индексу Дэвиса-Боулдина и времени вычисления (табл. 1). Судя по результатам, все три способа эффективны, однако первые два работают значительно быстрее.

Таблица 1. Оценка методов кластеризации с помощью метрики индекс Дэвиса-Боулдина

Метод кластеризации	K-Means	Mini Batch K-Means	Affinity Propagation
индекс Дэвиса-Боулдина	1.242	1.276	1.214
Время, сек.	0.331	0.296	1.821

Далее в нашем эксперименте была применена предложенная функция расстояния между двумя конфигурациями (и вычислена матрица сходства), учитывающая симметрию поворотов и сдвигов на воображаемом торе, которая позволила снизить количество кластеров, получаемое методом AffinityPropagation.

Таблица 2. Качество кластеризации методом AffinityPropagation до и после оптимизации

	до оптимизации	после оптимизации
количество кластеров	75	60

Заключение

В качестве методов кластеризации были выбраны и исследованы методы K-means, Mini Batch K-Means и AffinityPropagation.

Методы K-means и Mini Batch K-Means подходят для большого набора данных и имеют хорошую скорость работы, но для решения нашей задачи не вполне пригодны из-за особенностей алгоритмов поиска кластеров

Метод AffinityPropagation в настоящее время исследуется, для него стало возможным сократить количество получаемых кластеров на 20%.

Работа выполнена по гранту КН МОН РК №AP05133366 по теме: Развитие методов оптимизации, ориентированных на широкий класс прикладных задач нанотехнологий (2018-2020 гг.).

Литература

1. Физика конденсированного состояния: уч.пос. / Н.И. Филимонова, Р.П. Дикарева. — Новосибирск: Изд-во НГТУ, 2016. — 136 с.
2. Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.
3. Brendan J. Frey and Delbert Dueck, “Clustering by Passing Messages Between Data Points”, Science Feb. 2007
4. L. Dalcin, P. Kler, R. Paz, and A. Cosimo, Parallel Distributed Computing using Python, Advances in Water Resources, 34(9):1124-1139, 2011. <http://dx.doi.org/10.1016/j.advwatres.2011.04.013>
5. L. Dalcin, R. Paz, M. Storti, and J. D’Elia, MPI for Python: performance improvements and MPI-2 extensions, Journal of Parallel and Distributed Computing, 68(5):655-662, 2008. <http://dx.doi.org/10.1016/j.jpdc.2007.09.005>
6. L. Dalcin, R. Paz, and M. Storti, MPI for Python, Journal of Parallel and Distributed Computing, 65(9):1108-1115, 2005. <http://dx.doi.org/10.1016/j.jpdc.2005.03.010>
7. Лурье С.А., Посыпкин М.А., Соляев Ю.О. Метод идентификации масштабных параметров градиентной теории упругости на основе численных

экспериментов для плоских композитных структур //International Journal of Open InformationTechnologies, 2015. Т. 3. No 6. С. 1-6.

8. Ю.Г. Евтушенко, С.А. Лурье, М.А. Посыпкин, Ю.О. Соляев. Применение методов оптимизации для поиска равновесных состояний двумерных кристаллов // Журнал вычислительной математики и математической физики. 2016. Т. 56. No 12. С.50-59.

9. Амирханова Г.А., Горчаков А.Ю., Дуйсенбаева А.Ж., Посыпкин М.А. Применение метода точных штрафных функций к задаче минимизации энергии плоского кристалла // Материалы XIV Международной Азиатской школы-семинара «Проблемы оптимизации сложных систем» 20-31 июля 2018 года. — Алматы, 2018. — Ч. 1 — С. 107–113.

10. Амирханова Г.А., Дуйсенбаева А.Ж. Сравнение двух вариантов метода мультистарта // Матер. науч. конф. «Инновационные IT и Smart-технологии», посвященной 70-летию юбилею профессора Утепбергенова И.Т. (20 марта 2019 г). — Алматы: ИИВТ МОН РК, 2019. — С. 304–309.

УДК 551.501: 629.195.1

ИЗМЕНЕНИЯ СОДЕРЖАНИЙ МАЛЫХ ГАЗОВ В АТМОСФЕРЕ КАЗАХСТАНА ПО ДАННЫМ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ

Ахмеджанов А.Х., Ахметов Н.Д., Караданов Т.К.

E-mail: model_lab@mail.ru

*АО «Национальный центр космических исследований и технологий»
АК МЦРОАП РК, Республика Казахстан, г. Алматы*

Аннотация. Приводятся результаты обработки спутниковых данных и анализа содержания малых газов в атмосфере Казахстана. Содержание в атмосферном столбе диоксида азота (NO_2) по спутниковым данным в теплое время года максимальное, в холодное время минимальное. Анализ спутниковых данных за период 2004.10.01 по 2019.04.19 показал, что значение содержания диоксида азота в атмосферном столбе увеличилось на 8,57 %. За период с 1980 по 2018 годы содержание диоксида серы (SO_2) уменьшилось на 38%. Наибольшее содержание SO_2 наблюдаются в зимние месяцы, а минимальные в теплое время года. Максимумы содержания оксида углерода (CO) наблюдаются в периоды март-апрель, минимум на июль по октябрь месяцы. В среднем по территории Казахстана общее содержания CO в атмосферном столбе за последние 38 лет уменьшилось на 4,7%. Средние и максимальные значения содержания CO в атмосфере Казахстана имеют тенденцию к понижению.

Развитие промышленного производства и транспортных систем на территории Казахстана сопровождается ростом загрязнения атмосферы в результате эмиссии различных газовых примесей, в том числе окислов азота и серы.

Оценка интегрального содержания NO_2 в пограничном слое атмосферы по наблюдениям солнечного излучения дана в работе [1]. Результаты анализа

спутниковых наблюдений тропосферного содержания NO_2 над Московским регионом представлены в [2]. Около 40% загрязнений оксидами азота производится из точечных источников на электростанциях [3]. Диоксид азота отличается относительно продолжительным периодом существования в атмосфере. Выбросы NO_2 считаются одной из основных причин образования фотохимического смога. Соединяясь с парами воды в атмосфере, они образуют азотную кислоту, и, вместе с оксидами серы, являются причиной образования кислотных дождей. Выполнена оценка накопления и перемещения диоксидов азота и серы в Армении по материалам дистанционного зондирования Земли [4].

Содержание CO в атмосфере характеризуется значительной пространственно-временной изменчивостью [5]. В целом общее содержание CO в Северном полушарии выше, чем в Южном полушарии, что объясняется большим вкладом антропогенных источников CO . Важной составляющей изменчивости CO в атмосферном столбе является временной ход [6]. В низких широтах в межгодовой изменчивости CO проявляется квазидвухлетняя цикличность [7]. Анализ спутниковых данных позволил впервые выявить особенности сезонного хода и трендов CO и прояснить степень влияния локальных источников на крупномасштабное загрязнение атмосферы оксидом углерода [8].

Целью данной работы является исследование особенностей содержания диоксидов азота и серы в атмосфере Казахстана на основе данных дистанционного зондирования Земли из космоса.

Для анализа выполнена выборка и обрабатывались данные зондирующего комплекса AIRS (Atmospheric InfraRed Sounder) на платформе Aqua, который из четырех приборов: 2378-канальный ИК-сканер AIRS, 4-канальный спектрометр AIRS/VisNIR, а также 13-канальный микроволновой радиометр AMSU-A1 и 2-канальный AMSU-A2. Для оценки пространственно-временной изменчивости содержания малых газов в атмосферном столбе были использованы многолетние данные измерений сенсора OMI, размещенного на спутнике Aura, а также данные спутникового прибора MOPITT (Measurements Of Pollution In The Troposphere). Последний прибор является 8-ми канальным газокорреляционным ИК-радиометром, измеряющим отраженную солнечное излучение и излучение земной поверхности и атмосферы, радиацию в диапазонах 2.3 мкм и 4.7 мкм.

Построено сглаженное распределение содержания диоксида азота в атмосферном столбе ($*10^{+15} \text{ 1/см}^2$) по территории Казахстана за 2018 год по данным дистанционного зондирования (рис. 1).

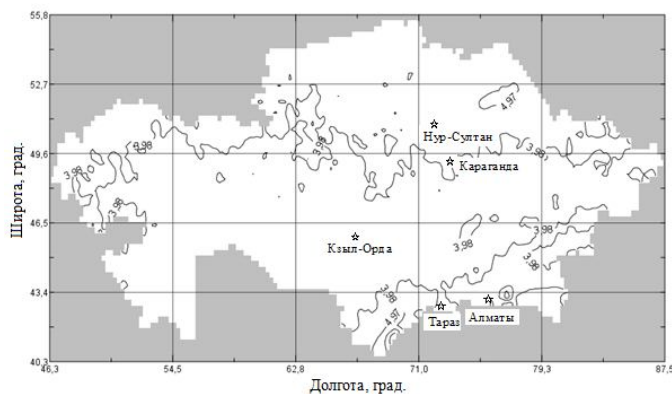


Рисунок 1 - Сглаженное распределение содержания диоксида азота в атмосферном столбе ($*10^{+15} \text{ 1/см}^2$) по территории Казахстана за 2018 год по данным дистанционного зондирования

Выбранные по наземным измерениям города Нур-Султан (Астана), Алматы и Тараз, в которых по данным наземных измерений [9] наблюдаются повышенные концентрации метана, попадают в зоны повышенного содержания диоксида азота в атмосферном столбе, отмеченные на рис.1 звездочками, что свидетельствует об удовлетворительном качественном соответствии данных дистанционного зондирования и наземных измерений.

Статистический анализ значений содержания NO_2 в атмосферном столбе по территории Казахстана показал, что среднее значение равно $3.67\text{e}+15$ $1/\text{cm}^2$, медиана $3.72\text{e}+15$ $1/\text{cm}^2$, среднеквадратичное отклонение $8.71 \text{e}+14$ $1/\text{cm}^2$. Изменчивость за период 2004-2019 годы содержания в столбе NO_2 , усредненного по территории, показана на рис. 2. За исследуемый период (13 лет) увеличение значения содержания диоксида азота в атмосферном столбе составило 8,57 %. Внутригодовые изменения содержание в столбе диоксида азота за 2016, 2017 и 2018 годы по спутниковым данным сенсора OMI показаны на рис. 3. Максимумы изменения содержания диоксида азота в атмосферном столбе по территории Казахстана наблюдаются в летние месяцы (июль) и минимумы зимой (январь). В то же время в холодный период амплитуды изменений имеют максимумы, а в теплый период интервал изменений минимальный.

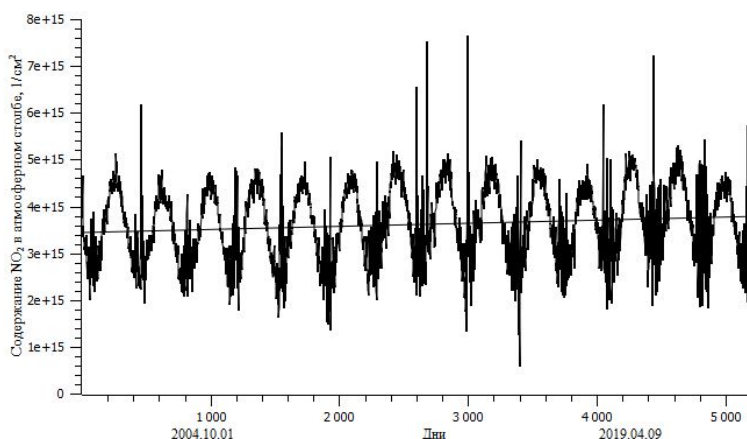


Рисунок 2 - Изменения содержания диоксида азота в атмосферном столбе по территории Казахстана за период с октября 2004 года по апрель 2019 года

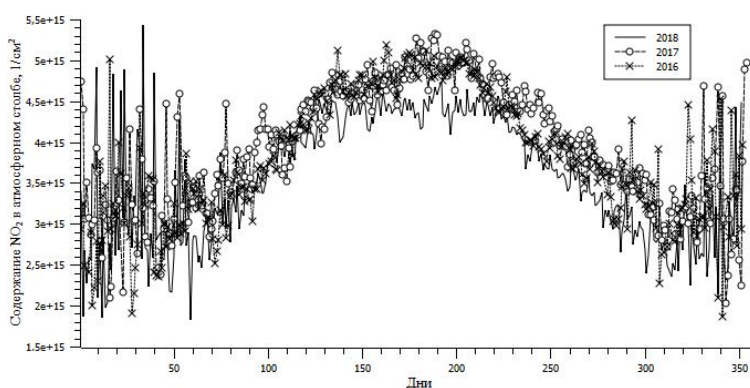


Рисунок 3 - Внутригодовые изменения содержание в столбе диоксида азота за 2016, 2017 и 2018 годы по спутниковым данным

Многолетние изменения и тренд содержания диоксида серы в атмосферном столбе за период с 1980 по 2018 годы по территории Казахстана показаны на рис. 4. Наблюдается тенденция к уменьшению содержания SO_2 в атмосферном столбе по Казахстану. За период с 1980 по 2018 годы содержание диоксида серы уменьшилось на 38%. Внутригодовые изменения содержания диоксида серы по территории Казахстана за последние 10 лет представлены на рис. 5. В основном наибольшие содержание SO_2 наблюдаются в зимние месяцы, а минимальные в теплое время года. Максимум в летнее время выпал на 2010 год, когда на европейской территории России наблюдалась температурная аномалия [10]. Результаты статистического анализа содержания SO_2 в атмосферном столбе по Казахстану представлены в таблице 2. Статистический анализ содержания SO_2 в атмосферном столбе по Казахстану показал, что среднее значение находится в диапазоне от $2,46 \text{ мг/м}^3$ в 2015 году до $3,16 \text{ мг/м}^3$ в 2010 году. Стандартная ошибка изменялась от $0,07 \text{ мг/м}^3$ в 2009 и 2015 годах до $0,23 \text{ мг/м}^3$ в 2012 году. Медиана находилась в диапазоне $2,38 \text{ мг/м}^3$ в 2016 году до $2,99 \text{ мг/м}^3$ в 2010 году.

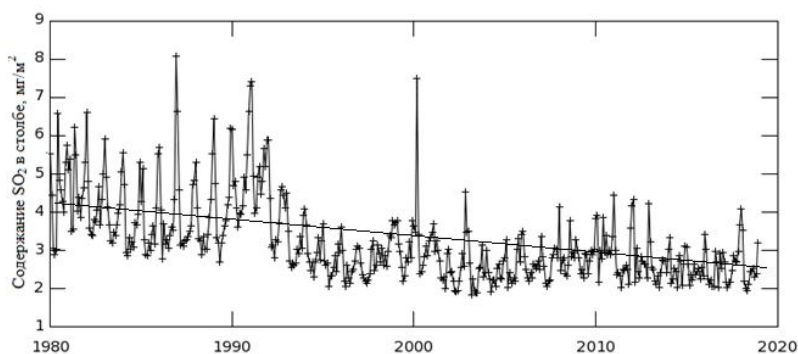


Рисунок 4 - Многолетние изменения содержания диоксида серы в атмосферном столбе по территории Казахстана

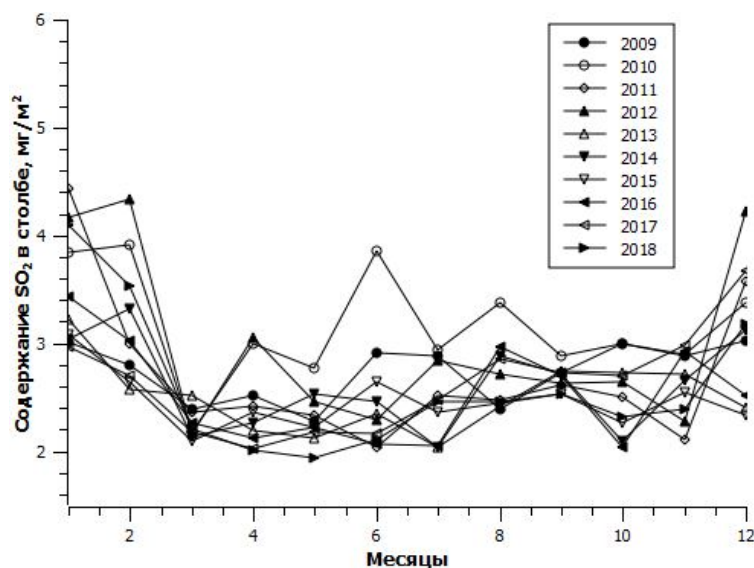


Рисунок 5 - Внутригодовые изменения содержания диоксида серы по территории Казахстана за 2009 - 2018 годы

Сезонная изменчивость общего содержания CO по территории Казахстана за период 2007-2017 годы по данным прибора MOPITT (спутник TERRA) представлена на рис. 6. Максимумы содержания CO наблюдаются в периоды март-апрель, минимум на июль по октябрь месяцы. Это может быть связано с отопительными сезонами года. Локальный максимум содержания CO 2010 года отражен на представленном графике. Объяснение этим обстоятельствам могут быть аномально жаркое летнее время и сопровождающие их степные пожары в эти годы. Многолетние изменения содержания оксида углерода в атмосфере по территории Казахстана за период 1980-2018 годы показаны на рис. 7. За рассматриваемый период средние и максимальные значения содержания CO имеют тенденцию к понижению. Общее содержание CO в атмосферном столбе за последние 38 лет уменьшилось на 4,7%.

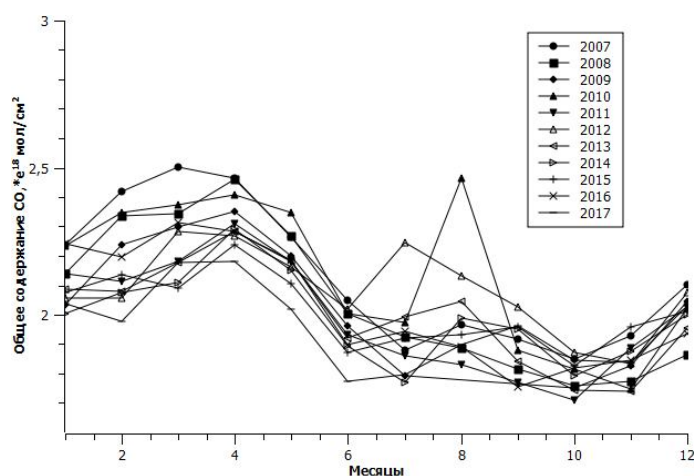


Рисунок 6 - Сезонная изменчивость общего содержания CO по территории Казахстана за период 2007-2017 годы

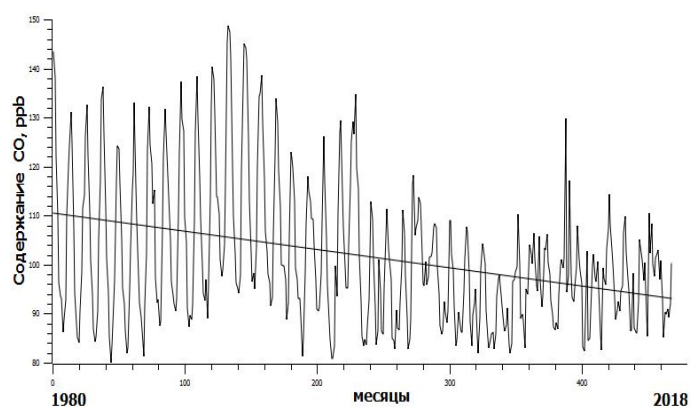


Рисунок 7 - Многолетние изменения общего содержания CO по территории Казахстана за период 1980-2018 годы

Исследования выполнены в рамках грантового проекта № AP05135191 МОН РК "Исследовать особенности распределения малых газов, загрязняющих атмосферу, по территории Казахстана на основе данных дистанционного зондирования Земли из космоса".

Литература

1. Иванов В.А., Постыляков О.В. Оценка интегрального содержания NO₂ в пограничном слое атмосферы по наблюдениям рассеянной в зените солнечной радиации. // Оптика атмосферы и океана. 2010. - Том 23. - № 06. - С. 471-474.
2. Ситнов С. Анализ спутниковых наблюдений тропосферного содержания NO₂ над Московским регионом // Известия РАН. Физика атмосферы и океана. 2011. -Т. 47. - № 2. - С. 184-203.
3. Ahmad Kamruzzaman Majumder, V. Krishna Murthy, 3Sanjay Nath Khanal, Dhiraj Giri. NO₂ concentration in Banepa valley, Nepal // Kathmandu university journal of science, engineering and technology. September 2008. - Vol. I. - no. V.- pp 1-11.
4. Станкевич С.А., Налбандян М.А., Андреасян Д.М., Титаренко О.В. Оценка накопления, пространственного распределения и перемещения диоксидов азота и серы в атмосфере бассейна реки Дебед по материалам дистанционного зондирования Земли и возможные риски для здоровья населения // Современные проблемы дистанционного зондирования Земли из космоса. 2017. - Т. 14. - № 2. - С. 240-249.
5. Ситнов С. А., Мохов И. И., Джола А. В.. Общее содержание оксида углерода в атмосфере над российскими регионами по спутниковым данным // Изв. РАН. Физика атмосферы и океана, 2017, том 53, № 1, с. 38–55.
6. Арефьев В.Н., Кашин Ф.В., Орозалиев М.Д. и др. Структура временных вариаций окиси углерода в толще атмосферы центральной части Евразии (станция мониторинга “Иссык-Куль”) // Изв. РАН. Физика атмосферы и океана. 2013. Т. 49. № 2. С. 165–170.
7. Ситнов С.А. Анализ квазидвухлетней изменчивости общего содержания окиси углерода и ее связи с квазидвухлетней изменчивостью общего содержания озона // Изв. РАН. Физика атмосферы и океана. 2008. том 44. № 4. С. 494–502.
8. Novelli P.C., Masarie K.A., Lang P.M. Distributions and recent changes in atmospheric carbon monoxide // J. Geophys. Res. 1998. V. 103. P. 19015–19033.
9. Информационные бюллетени о состоянии окружающей среды 2017-2018 г.г. РГП Казгидромет МЭ РК // <https://kazhydromet.kz/ru/bulleten/okrsreda> (3.06.2019 г.)
10. Сидоренков Н. С., Сумерова К. А. Биения колебаний температуры как причина аномально жаркого лета 2010 г. на европейской территории России // Метеорол. и гидрол. 2012. - № 6. - С. 81-94.

РЕГУЛЯРИЗАЦИЯ ДВИЖЕНИЯ РОБОТА ПО ЗАДАННОЙ ТРАЕКТОРИИ

Ахметжанов¹ М.А., Алпысбай¹ Н.Э., Даулетбек² Е.Т.

e-mail: maks714@mail.ru, alpysbainursultan@gmail.com, y.dauletbek@iitu.kz

¹*Институт информационных и вычислительных технологий КН МОН РК, Казахстан,*

²*Международный Университет Информационных Технологий, Казахстан, Алматы*

***Аннотация.** В статье представлен кинематический анализ движения планарного параллельного робота на основании построенной 3D модели. В ходе анализа учитывались такие элементы исследования движения как ограничения движения, свойства материала, масса, скорости и контакты компонентов. Приведены результаты моделирования движения робота. В результате анализа определены максимальные перемещения звеньев, а также установлена адекватность разработанных подвижных частей и отсутствие конфликтующих связей. Получены траектории движения выходных звеньев модулей роботизированного комплекса. Проведенный анализ позволяет сделать выводы о корректности траекторий перемещения модулей для выполнения операций обработки изделий различной сложности. Представлены результаты математического моделирования.*

Введение

При создании новых типов высокопроизводительного инновационного оборудования для различных применений (механической обработки, резки, сварки, сборки, окраски) возникают проблемы разработки новых схемно-технических решений их компоновок, обеспечивающих сложные виды движений выходного звена в пространстве и высокую точность позиционирования. Одним из вариантов решения этой проблемы является использование механизмов параллельной структуры [1–4].

Классические механизмы параллельной структуры с 6 степенями свободы (например, гексапод) весьма сложны в управлении. Однако существует возможность упростить задачу управления, построив механизм с 6 степенями свободы в виде соединения двух механизмов параллельной структуры с 3 степенями свободы каждый [5].

В настоящее время механизмы параллельной структуры [9] находят всё более широкое применение в промышленности. В сравнении с более распространенными последовательными аналогами параллельные роботы обладают существенным превосходством по точности позиционирования, мощности и допустимой нагрузке [6–8]. Одной из важнейших подзадач построения параллельных механизмов является определение их конструктивных характеристик (размеров неподвижного основания, подвижной платформы, диапазонов длин или углов поворота приводных звеньев) в соответствии с необходимыми размерами рабочего пространства, т.е. диапазоном изменений выходных координат. Решение этой подзадачи рассмотрено в настоящей работе на примере планарного 3-RPR механизма с тремя степенями свободы (перемещение вдоль осей x и y и поворот вокруг оси z). Механизм состоит из трех кинематических цепей, содержащих одну приводную поступательную и две пассивных вращательных кинематических пары. Плоский механизм параллельной структуры представленный на рис.1 имеет одну вращательную кинематическую пару, сопряженную с основанием, одну вращательную кинематическую пару, сопряженную с выходной платформой, и одну промежуточную вращательную кинематическую пару. Модуль имеет перемещения вдоль горизонтальных осей x и y , а также вращение вокруг вертикальной оси z .

Теоретическая часть

Рассмотрим задачу оптимизации закона движения платформы для планарного робота с тремя степенями свободы 3-RPR, кинематическая схема которого показана на рисунке 1.

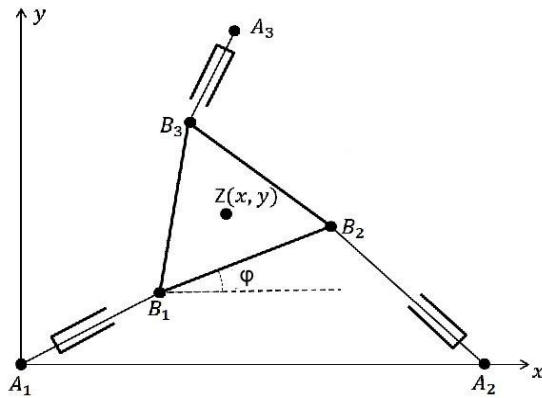


Рис.1. Кинематическая схема плоского робота с тремя степенями свободы

В названии число 3 означает количество степеней свободы, а буквы R,P,R соответствуют типам подвижных соединений. Этот робот представляет собой плоский манипулятор с шестью ротационными шарнирами для закрепления штанг и тремя призматическими двигателями. Двигатели изменяют длину штанг, соединяющих неподвижное основание $A_1A_2A_3$ с углами треугольной плоской платформы $B_1B_2B_3$, управляя ее движением. Будем считать основание и платформу равносторонними треугольниками.

Положение манипулятора характеризуется следующими параметрами:

(x,y) - координаты центра масс подвижной платформы

ϕ - угол поворота платформы относительно нейтрального положения

l_1, l_2, l_3 – длины штанг A_1B_1, A_2B_2, A_3B_3 соответственно.

Постановка задач оптимизации движений платформы зависит от практических приложений и требует, в первую очередь, определения критериев качества такого движения. Например, критериями качества могут быть потребляемая мощность, минимизация ускорений, вибраций, увеличений рабочей области и т.д. при заданных ограничениях.

В данном исследовании мы рассмотрим задачу оптимизации закона движения манипулятора по заданной траектории.

Пусть траектория центра тяжести движущейся платформы задана в плоскости координат (x,y) параметризацией

$$\begin{aligned} x_c &= x(p), \\ y_c &= y(p). \end{aligned} \tag{1}$$

Пусть функция $\varphi = \varphi(t)$ определяет угол поворота платформы в каждый момент времени.

Определим закон изменения параметра $p = p(t)$ в зависимости от времени так, чтобы движение платформы, определяемое соотношениями

$$\begin{aligned} x_c(t) &= x(p(t)), \\ y_c(t) &= y(p(t)), \\ \varphi &= \varphi(t), \end{aligned} \tag{2}$$

имело минимально возможную кинетическую энергию за рассматриваемый интервал движения и при этом обеспечивало минимальные силы инерции за весь период движения. Пренебрежем массами штанг, шарниров и двигателей и

действием других внешних сил. Такая постановка имеет смысл, если стоит задача уменьшения энергопотребления манипулятора и при этом надо получить закон движения по траектории, обеспечивающий движения платформы с минимальной инерцией.

Будем считать, что положения и скорости платформы в начальный и конечный моменты времени заданы, а масштаб времени выбран так, чтобы время менялось в интервале $[0,1]$. Допустим также, что в начальный и конечный момент времени скорости равны нулю, т.е. эти моменты являются точками остановки движения:

$$\begin{aligned}x(0) &= x_0, x(1) = x_1, \\y(0) &= y_0, y(1) = y_1, \\ \varphi(0) &= \varphi_0, \varphi(1) = \varphi_1, \\ \dot{x}(0) &= 0, \dot{x}(1) = 0, \\ \dot{y}(0) &= 0, \dot{y}(1) = 0, \\ \dot{\varphi}(0) &= 0, \dot{\varphi}(1) = 0.\end{aligned}\tag{3}$$

Здесь мы предположили, что в начальный и конечный момент времени скорости равны нулю, т.е. эти моменты являются точками остановки движения. Точки над функциями, как обычно, обозначают производные по времени. Штрихи над функциями $x(p)$, $y(p)$ будут обозначать дифференцирование по параметру p .

Положим, что в начальный и конечный момент времени функция $p(t)$ удовлетворяет условиям

$$\begin{aligned}p(0) &= p_0, \\ p(1) &= p_1.\end{aligned}\tag{4}$$

Запишем выражения для скоростей и ускорений точек платформы в инерциальной системе координат, связанной с неподвижным основанием:

$$\begin{aligned}\vec{v} &= \vec{v}_C + \vec{\omega} \times \vec{r}, \\ \vec{a} &= \vec{a}_C + 2\vec{\omega} \times \vec{v} + \vec{\omega} \times (\vec{\omega} \times \vec{v}).\end{aligned}$$

Здесь \vec{r} , \vec{v} , \vec{v}_C , \vec{a} , \vec{a}_C , $\vec{\omega}$ - радиус-вектор, скорость, скорость центра масс, ускорение, ускорение центра масс и угловая скорость платформы. Векторы $2\vec{\omega} \times \vec{v}$, $\vec{\omega} \times (\vec{\omega} \times \vec{v})$ отвечают за кориолисово и центробежное ускорения.

Суммарная сила инерции есть

$$F_i = - \iint_{B_1 B_2 B_3} \rho \vec{a} dS.$$

В предположении, что масса платформы в основном сосредоточена вблизи центра тяжести, мы можем приближенно записать, что сила инерции есть

$$F_i = - \iint_{B_1 B_2 B_3} \delta(\vec{r} - \vec{r}_C) \vec{a} dS = -m\vec{a}_C.$$

Здесь m – масса платформы.

В качестве меры величины сил инерции за все время движения будем рассматривать интеграл:

$$\|F_i\|^2 = \int_0^1 m^2 a_c^2(t) dt = m^2 \int_0^1 (\ddot{x}^2 + \ddot{y}^2) dt.$$

Запишем кинетическую энергию движения платформы, пренебрегая массой стержней и шарниров:

$$E_k = \frac{1}{2} \int_0^1 m (\dot{x}^2(p(t)) + \dot{y}^2(p(t))) dt + \frac{1}{2} \int_0^1 I \dot{\varphi}^2(t) dt.$$

Здесь I - момент инерции платформы. Согласно нашему предположению, вся масса сосредоточена в центре масс платформы, поэтому кинетической энергией вращательного движения можно пренебречь по сравнению с энергией поступательного движения, тогда

$$E_k = \frac{1}{2} \int_0^1 m (\dot{x}^2(p(t)) + \dot{y}^2(p(t))) dt.$$

Введем функционал качества движения в виде суммы кинетической энергии и некоторой меры сил инерции следующим образом:

$$J(p) = \frac{1}{2} \int_0^1 m (\dot{x}^2(t) + \dot{y}^2(t)) dt + \frac{\alpha m^2}{2} \int_0^1 (\ddot{x}^2(t) + \ddot{y}^2(t)) dt. \quad (5)$$

Здесь $\alpha > 0$ – некоторый положительный параметр, определяющий вес сил инерции в функционале качества, который мы будем называть регуляризатором.

Экспериментальная часть

Рассматривалась модель робота 3-RPR со следующими геометрическими параметрами: длины штанг в исходном положении равны 30 см, радиус окружности, описанной вокруг подвижной платформы и неподвижного основания равен соответственно 12 см и 42 см, угол поворота равен нулю. Ограничения на длины штанг не рассматривались.

Самым простым является случай, когда центр масс движется по прямолинейной траектории. В этом случае параметризация траектории может быть задана в виде:

$$\begin{aligned} x &= p, y = kp + b, \\ x_c(t) &= x(t), y_c(t) = kx(t) + b. \end{aligned}$$

На рисунке 2 показаны профили скорости и ускорения при разных значениях регуляризирующего параметра α . Очевидно, что увеличение параметра уменьшает ускорения и делает профиль скоростей более плавным. Расчет показал, что увеличение α от 0.01 до 0.5 приводит к возрастанию кинетической энергии платформы возрастает на 7.5%. Однако, за счет регуляризации профиля скорости удается существенно уменьшить силы инерции. Максимальные значения уменьшаются почти на 50% (см. рисунок 2).

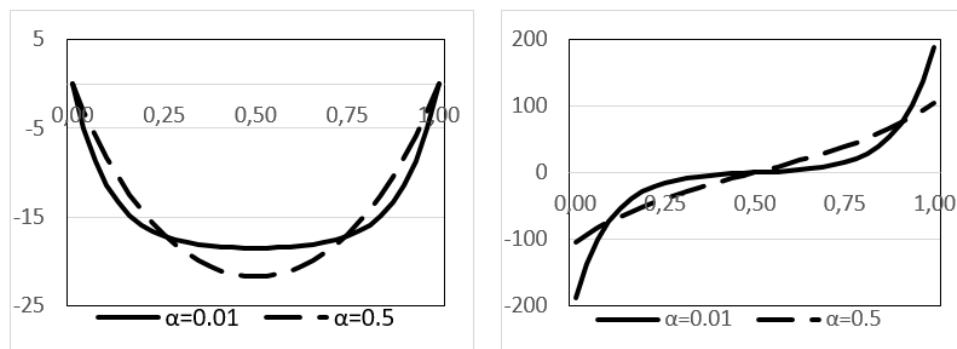


Рис.2. Профили скорости и ускорения при разных значения регуляризатора α

На основе полученных данных при расчете, мы смогли достичь последовательным положением платформы с течением времени. Платформа совершает поступательное движение.

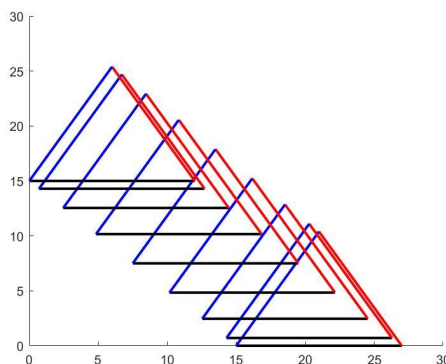


Рис.3. Последовательные положения платформы при $\alpha = 0.5$, платформа движется с правого нижнего положения до левого верхнего

Заключение

В конечном итоге мы смогли получить, обширную формулу по которым определены профили скорости и ускорения при разных значения регуляризатора α , и прийти к формулировке, закон движения по прямолинейной траектории, на которой должен достигаться минимум функционала качества и должен удовлетворять необходимому условию минимума. Представленные возможности применения кинематического анализа позволяют получить результаты, которые могут быть применены при проектировании подобных конструкций и упростить дальнейшие исследования в данной области. Применение предложенной методики переноса имеющихся ограничений в пространство входных координат робота-станка параллельной структуры и использование в качестве целевой функции модифицированной метрики позволяют проводить оптимизацию времени позиционирования рабочего инструмента и, в конечном итоге, снизить общую длительность обработки изделия.

Литература

1. Glazunov V.A., Koliskor A.S., Krainev A.F., 1991. The Spatial Mechanisms with Parallel Structure. M.: Nauka, pp. 95.

2. Rybak L.A., Chichvarin A.V., Mamaev Y.A., Gaponenko E.V. Synthesis of Control System of Single- and a Two-Section Manipulators with Parallel, Вестник БГТУ им. В.Г. Шухова 2018, №9 Kinematics // Fundamental Problems of Engineering and Technology. 2012, (2-6). Pp. 60-68.

3. Rybak L.A., Gaponenko E.V., Mamaev Y.A., New Technologies of High-Speed Mechanical Processing on Machine Tools with Parallel Kinematics // Fundamental Problems of Engineering and Technology, 2012, (6): www.science-education.ru/106-7430.

4. Evtushenko Y.G., Posypkin M.A. Nonuniform covering method as applied to multicriteria optimization problems with guaranteed accuracy // Computational Mathematics and Mathematical Physics. – 2013. – Т. 53. – №. 2. – С. 144.

5. Rybak L.A., Gosselin C., Chichvarin A.V., Grinek A.V. The method of structural synthesis of parallel mechanisms with the set of the trajectory of output link on the plane // Advances in Environmental Biology. 2014. №8 (4). Pp. 855-861.

6. Посыпкин М.А., Ахметжанов М.А., Намазбаев Б.Д. Методы аппроксимации решения систем уравнений и неравенств. Труды науч. конф. ИИВТ МОН РК «Современные проблемы информатики и вычислительных технологий». Алматы, 2018. - С.228-235.

7. Посыпкин М.А., Ахметжанов М.А., Намазбаев Б.Д. Определение рабочей области робота с помощью метода неравномерных покрытий. Материалы XIV Международной Азиатской школы-семинара «Проблемы оптимизации сложных систем». Алматы. 2018. Часть 2. –С.368.

8. Rybak L.A., Gaponenko E.V., Chichvarin A.V. Determination of Static Error of the Machine with Parallel Kinematic // World Applied Sciences Journal. 2014. №30 (9). Pp. 1193–1198.

9. Rybak L.A., Gaponenko E.V., Mamaev Y.A. Algorithm for correcting the trajectory of the working body of the parallel robot based on special provisions // World Applied Sciences Journal. 2015. №25 (9). Pp. 824-834.

ДВИЖЕНИЕ РАБОЧЕГО ИНСТРУМЕНТА ДЛЯ РОБОТА 3RPR ПО ЗАДАННОЙ ТРАЕКТОРИИ С МИНИМАЛЬНОЙ СУММАРНОЙ ДЛИНОЙ ШТАНГ

Ахметжанов М.

maks714@mail.ru

*Институт информационных и вычислительных технологий, Алматы,
Казахстан*

Рассмотрим плоское движение подвижной платформы робота типа 3RPR с тремя степенями свободы. В качестве координат, определяющих движение платформы, примем координаты центра масс платформы, которые будем обозначать через (x, y) либо (x_C, y_C) в зависимости от контекста и угол поворота платформы φ .

Пусть закон траектория движения центра масс платформы известна. Обычно эта траектория зависит от технологических условий и определяется выполняемой

роботом задачей. Допустим, что центральная точка платформы является стартовой позицией робота.

Зададимся вопросом: как следует менять угол поворота платформы, чтобы во все время движения робота по заданной траектории суммарная длина штанг, приводящих робот в движение, была минимальна?

Длина штанги в каждый момент движения задается длинами векторов A_1B_1 , A_2B_2 , A_3B_3

Итак, нам следует во все время движения минимизировать суммарную длину штанг.

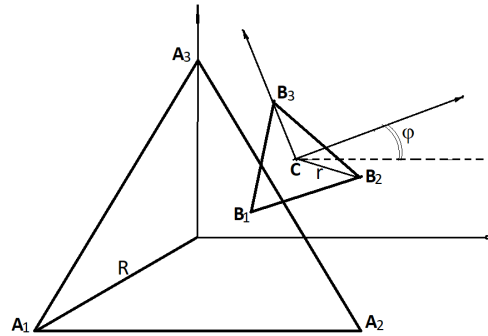


Рис. 1. Подвижная и неподвижная платформы и связанные системы отсчета

Построим функционал качества для данной задачи. При заданном законе движения центра и угла поворота платформы координаты точек B_i определяются соотношениями:

$$\begin{aligned} x_i^B &= x(t) + \xi_i \cos \varphi(t) - \eta_i \sin \varphi(t), \\ y_i^B &= y(t) + \xi_i \sin \varphi(t) + \eta_i \cos \varphi(t). \end{aligned} \quad (3.1)$$

$$i = \overline{1, 3}.$$

Здесь

$$\begin{aligned} \xi_1 &= -r\sqrt{3}/2, \quad \eta_1 = -r/2, \\ \xi_2 &= r\sqrt{3}/2, \quad \eta_2 = -r/2, \\ \xi_3 &= 0, \quad \eta_3 = r. \end{aligned} \quad (3.2)$$

есть координаты вершин $B_1B_2B_3$ в системе с центром в точке C в подвижной системе координат, жестко связанной с платформой $B_1B_2B_3$, r – радиус окружности, описанной вокруг $B_1B_2B_3$.

С учетом (3.2) координаты вершин (3.1) расписываются в виде:

$$\begin{aligned} x_1^B &= x_C - 0.5\sqrt{3}r \cos \varphi + 0.5r \sin \varphi \\ y_1^B &= y_C - 0.5\sqrt{3}r \sin \varphi - 0.5r \cos \varphi, \\ x_2^B &= x_C + 0.5\sqrt{3}r \cos \varphi + 0.5r \sin \varphi, \\ y_2^B &= y_C + 0.5\sqrt{3}r \sin \varphi - 0.5r \cos \varphi, \\ x_3^B &= x_C - r \sin \varphi, \\ y_3^B &= y_C + r \cos \varphi, \end{aligned} \quad (3.3)$$

В неподвижной системе координат с началом в центре тяжести основания координаты неподвижного основания задаются как:

$$\begin{aligned}x_1^A &= -\sqrt{3}R/2, y_1^A = -R/2, \\x_2^A &= \sqrt{3}R/2, y_2^A = -R/2, \\x_3^A &= 0, y_3^A = R.\end{aligned}\quad (3.4)$$

Для каждого момента времени требуется решить задачу минимизации:

$$J[\varphi] = \sum_{i=1}^3 (x_i^B - x_i^A)^2 + (y_i^B - y_i^A)^2 \xrightarrow{\varphi} \min. \quad (3.5)$$

Подставляя в (3.5) значения координат (3.3), (3.4), получаем:

$$\begin{aligned}J[\varphi] &= (x_C - 0.5\sqrt{3}r \cos \varphi + 0.5r \sin \varphi + \sqrt{3}R/2)^2 + (y_C - 0.5\sqrt{3}r \sin \varphi - 0.5r \cos \varphi + R/2)^2 \\&+ (x_C + 0.5\sqrt{3}r \cos \varphi + 0.5r \sin \varphi - \sqrt{3}R/2)^2 + (y_C + 0.5\sqrt{3}r \sin \varphi - 0.5r \cos \varphi + R/2)^2 \\&+ (x_C - r \sin \varphi)^2 + (y_C + r \cos \varphi - R)^2 = J_1 + J_2 + J_3 \xrightarrow{\varphi} \min\end{aligned}\quad (3.6)$$

Проведем ряд промежуточных преобразований, слагаемых выражения (3.6).

$$\begin{aligned}J_1 &= (x_C + \sqrt{3}R/2)^2 - (rx_C + r\sqrt{3}R/2)(\sqrt{3} \cos \varphi - \sin \varphi) + 0.25r^2 (\sqrt{3} \cos \varphi - \sin \varphi)^2 / 4 + \\&+ (y_C + R/2)^2 - (ry_C + rR/2)(\sqrt{3} \sin \varphi + \cos \varphi) + r^2 (\sqrt{3} \sin \varphi + \cos \varphi)^2 / 4 = \\&= (x_C + \sqrt{3}R/2)^2 + (y_C + R/2)^2 + r^2 - r(\sqrt{3}x_C + 3R/2 + y_C + R/2) \cos \varphi \\&+ r(x_C + \sqrt{3}R/2 - \sqrt{3}y_C - \sqrt{3}R/2) \sin \varphi = \\&(x_C + \sqrt{3}R/2)^2 + (y_C + R/2)^2 + r^2 - r \cos \varphi (\sqrt{3}x_C + y_C + 2R) \\&+ r \sin \varphi (x_C - \sqrt{3}y_C).\end{aligned}\quad (3.7)$$

$$\begin{aligned}J_2 &= (x_C + 0.5\sqrt{3}r \cos \varphi + 0.5r \sin \varphi - \sqrt{3}R/2)^2 + (y_C + 0.5\sqrt{3}r \sin \varphi - 0.5r \cos \varphi + R/2)^2 \\&= (x_C - \sqrt{3}R/2)^2 + (x_C - \sqrt{3}R/2)r(\sqrt{3} \cos \varphi + \sin \varphi) + 0.25r^2 (\sqrt{3} \cos \varphi + \sin \varphi)^2 \\&+ (y_C + R/2)^2 + (y_C + R/2)r(\sqrt{3} \sin \varphi - \cos \varphi) + 0.25r^2 (\sqrt{3} \sin \varphi - \cos \varphi)^2 = \\&r \cos \varphi (\sqrt{3}x_C - 2R - y_C) + r \sin \varphi (x_C + y_C \sqrt{3}) + \\&(y_C + R/2)^2 + (x_C - \sqrt{3}R/2)^2 + r^2.\end{aligned}\quad (3.8)$$

$$\begin{aligned}J_3 &= (x_C - r \sin \varphi)^2 + (y_C + r \cos \varphi - R)^2 = \\&x_C^2 - 2x_C r \sin \varphi + r^2 \sin^2 \varphi + \\&+ (y_C - R)^2 + 2(y_C - R)r \cos \varphi + r^2 \cos^2 \varphi = \\&+ 2r((y_C - R) \cos \varphi - x_C \sin \varphi) + r^2 + x_C^2 + (y_C - R)^2.\end{aligned}\quad (3.9)$$

Обозначим через r_i расстояния от центра до i -й вершины основания:

$$\begin{aligned} r_1^2 &= (y_c + R/2)^2 + (x_c + R\sqrt{3})^2 \\ r_2^2 &= (y_c + R/2)^2 + (x_c - R\sqrt{3})^2, \\ r_3^2 &= x_c^2 + (y_c - R)^2. \end{aligned}$$

Отсюда и из (3.7) - (3.9) следует, что минимизация функционала (3.5) эквивалентна минимизации следующего функционала:

$$\begin{aligned} J[\varphi] &= r_1^2 + r_2^2 + r_3^2 + 3r^2 - r \cos \varphi (\sqrt{3}x_c + y_c + 2R) + r \sin \varphi (x_c - \sqrt{3}y_c) \\ &+ r \cos \varphi (\sqrt{3}x_c - y_c - 2R) + r \sin \varphi (x_c + \sqrt{3}y_c) + \\ &- 2r(x_c \sin \varphi - (y_c - R) \cos \varphi) = \\ &= r \sin \varphi (x_c - \sqrt{3}y_c + x_c + \sqrt{3}y_c - 2x_c) + \\ &r \cos \varphi (-\sqrt{3}x_c - y_c - 2R + \sqrt{3}x_c - y_c - 2R + 2y_c - 2R) = \\ &r_1^2 + r_2^2 + r_3^2 + 3r^2 - 6Rr \cos \varphi \xrightarrow{\varphi} \min. \end{aligned} \quad (3.11)$$

Не составляет большого труда исследовать экстремумы функции (3.11) относительно переменной φ .

Если не рассматривать ограничения на углы поворота, то, очевидно, минимум достигается при $\cos \varphi = +1$, т.е. платформа должна находиться в положении $\varphi = 0$ во все время движения. Вторая производная функции

$$J(\varphi) = r_1^2 + r_2^2 + r_3^2 + 3r^2 - 6Rr \cos \varphi$$

есть

$$\frac{d^2}{d\varphi^2} J(\varphi) = 6Rr \cos^2 \varphi$$

При $\varphi=0$ эта функция строго положительна, что гарантирует наличие строгого минимума в этой точке.

Для проверки результата рассмотрим разные положения платформы для разных углов поворота и рассчитаем суммарную длину штанг для трех разных положений центра подвижной платформы

1. Платформа находится в центре основания, $x_c = 0.5L$, $y_c = \sqrt{3}L/6$

	$-\pi/2$	$-\pi/3$	$-\pi/6$	0	$\pi/6$	$\pi/3$	$\pi/2$
11	1.826	1.528	1.265	1.155	1.265	1.528	1.826
12	1.826	1.528	1.265	1.155	1.265	1.528	1.826
13	1.826	1.528	1.265	1.155	1.265	1.528	1.826
11+12+13	5.477	4.583	3.796	3.464	3.796	4.583	5.477

2. Платформа находится в крайнем нижнем положении, $x_C = 0.5L$, $y_C = 0$

	$-\pi/2$	$-\pi/3$	$-\pi/6$	0	$\pi/6$	$\pi/3$	$\pi/2$
11	1.31	1.041	0.923	1.041	1.857	1.607	1.857
12	1.857	1.607	1.31	1.041	1.31	1.041	1.31
13	0.577	0.577	0.577	0.577	0.577	0.577	0.577
11+12+13	3.745	3.225	2.81	2.659	3.745	3.225	3.745

3. Платформа находится на ребре A_1A_3 , $x_C = 0.25L$, $y_C = \sqrt{3}L/8$

	$-\pi/2$	$-\pi/3$	$-\pi/6$	0	$\pi/6$	$\pi/3$	$\pi/2$
11	1.239	0.971	0.672	0.439	0.485	0.753	1.049
12	2.543	2.251	1.967	1.787	1.794	1.986	2.273
13	1.31	1.041	0.923	1.041	1.31	1.607	1.857
11+12+13	5.092	4.263	3.562	3.267	3.59	4.346	5.18

Как видим на численных примерах, минимальная суммарная длина штанг достигается при поступательном движении платформы, т.е. при постоянном нулевом угле поворота.

Условия реализации движения с минимальной суммарной длиной штанг

Полученное выше требование постоянства угла поворота платформы указывает на то, что оптимальным в этом смысле является поступательное движение с нулевым углом поворота. С другой стороны, формула (3.11) указывает на то, что при желании уменьшить суммарную длину штанг при ненулевых углах поворота, следует стремиться к значениям угла φ , близким к нулю. Заметим, однако, что, фиксируя нулевое значение угла поворота, мы тем самым уменьшаем степень свободы платформы на единицу. При этом движение платформы становится поступательным.

Но, как показано в работе [2], такое движение является сингулярным, т.к. якобиан преобразования скоростей вырождается. Более того, в этой работе построение траектории проводится так, чтобы во все время движения избегать нулевого угла поворота.

Однако, как оказалось, в этом случае, даже если якобиан преобразования равен нулю, удастся однозначно решить и прямую и обратную задачи кинематики.

Рассмотрим, как в этом случае решаются обратная и прямая задачи кинематики.

Сингулярное движение робота 3RPR при поступательном движении и решении задач кинематики

Для рассматриваемого робота имеем три связи

$$\left(x + \xi_i \cos \varphi - \eta_i \sin \varphi - x_i^A\right)^2 + \left(y + \xi_i \sin \varphi + \eta_i \cos \varphi - y_i^A\right)^2 = l_i^2, \quad (4.1)$$

$$i = \overline{1,3}.$$

Координаты рабочего органа (x, y, φ)

В качестве координат актуаторов возьмем длины стержней либо другие три параметра. Кинематическая сингулярность проявляется в том, что якобиан

перехода от li к координатам подвижной платформы равен нулю. В данном случае это так и есть.

Якобиан используется при переходе от скоростей актуаторов к обобщенным скоростям рабочего инструмента. А именно, в возможности выразить скорости рабочего инструмента через скорости актуаторов. Продифференцируем (4.1) по времени.

Получим следующие соотношения:

$$\begin{aligned} & (x + \xi_i \cos \varphi - \eta_i \sin \varphi - x_i^A)(\dot{x} - (\xi_i \sin \varphi + \eta_i \cos \varphi)\dot{\varphi}) + \\ & (y + \xi_i \sin \varphi + \eta_i \cos \varphi - y_i^A)(\dot{y} + (\xi_i \cos \varphi - \eta_i \sin \varphi)\dot{\varphi}) = l_i \dot{l}_i, \\ & i = \overline{1,3}. \end{aligned}$$

Отсюда следует, что связь между скоростями актуаторов и скоростями платформы задается в виде:

$$\begin{pmatrix} l_1 & 0 & 0 \\ 0 & l_2 & 0 \\ 0 & 0 & l_3 \end{pmatrix} \begin{pmatrix} \dot{l}_1 \\ \dot{l}_2 \\ \dot{l}_3 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\varphi} \end{pmatrix} = 0$$

где

$$\begin{aligned} a_{i1} &= -(x + \xi_i \cos \varphi - \eta_i \sin \varphi - x_i^A) \\ a_{i2} &= -(y + \xi_i \sin \varphi + \eta_i \cos \varphi - y_i^A) \\ a_{i3} &= (x + \xi_i \cos \varphi - \eta_i \sin \varphi - x_i^A)(\xi_i \sin \varphi + \eta_i \cos \varphi) - \\ & (y + \xi_i \sin \varphi + \eta_i \cos \varphi - y_i^A)(\xi_i \cos \varphi - \eta_i \sin \varphi) \\ & i = \overline{1,3} \end{aligned}$$

В монографии Мерле [2] показано, что якобиан равен нулю, если три штанги пересекаются в одной точке. Именно это происходит в случае поступательного движения, в котором в угол поворота платформы равен нулю.

То есть, в случае поступательного движения матрица $\|a\|$ перехода от координат актуаторов к координатам рабочего инструмента вырождена.

Из общей теории следует, что в этом случае возможны ненулевые движения платформы, при которых актуаторы неподвижны.

Проверим, имеется ли такая возможность для рассматриваемого поступательного движения с нулевым углом поворота. Итак, пусть во все время движения угол поворота платформы остается постоянным и равным нулю.

Для нашего робота имеем из кинематических соотношений:

$$\begin{aligned} a_{11} &= -(2x - r\sqrt{3} \cos \varphi + r \sin \varphi + R\sqrt{3})/2 \\ a_{21} &= -(2x + r\sqrt{3} \cos \varphi + r \sin \varphi - R\sqrt{3})/2 \\ a_{31} &= -(x - r \sin \varphi) \\ a_{12} &= -(2y - r\sqrt{3} \sin \varphi - r \cos \varphi + R)/2 \\ a_{22} &= -(y + r\sqrt{3} \sin \varphi - r \cos \varphi + R)/2 \\ a_{32} &= -(y + r \cos \varphi - R) \end{aligned}$$

$$\begin{aligned}
a_{13} &= (2x - r\sqrt{3} \cos \varphi + r \sin \varphi + R\sqrt{3})(-\sqrt{3} \sin \varphi - \cos \varphi)r/2 - \\
& (2y - r\sqrt{3} \sin \varphi - r \cos \varphi + R)(-\sqrt{3} \cos \varphi + \sin \varphi)r/2 \\
a_{23} &= (2x + r\sqrt{3} \cos \varphi + r \sin \varphi - R\sqrt{3})(\sqrt{3} \sin \varphi - \cos \varphi)r/2 - \\
& (y + r\sqrt{3} \sin \varphi - r \cos \varphi + R)(\sqrt{3} \cos \varphi + \sin \varphi)r/2 \\
a_{33} &= (x - r \sin \varphi)r \cos \varphi - (y + r \cos \varphi - R)(-r \sin \varphi)
\end{aligned}$$

В случае поступательного движения с нулевым углом поворота платформы элементы матрицы задаются выражениями:

$$\begin{aligned}
a_{i1} &= -(x + \xi_i - x_i^A) \\
a_{i2} &= -(y + \eta_i - y_i^A) \\
a_{i3} &= (x + \xi_i - x_i^A)\eta_i - (y + \eta_i - y_i^A)\xi_i \\
i &= \overline{1,3}
\end{aligned}$$

Координаты вершин основания и подвижной платформ задаются в терминах радиусов описанных окружностей в виде:

$$\begin{aligned}
x_1^A &= -\sqrt{3}R/2, \quad y_1^A = -R/2, \\
x_2^A &= \sqrt{3}R/2, \quad y_2^A = -R/2, \\
x_3^A &= 0, \quad y_3^A = R. \\
\xi_1 &= -r\sqrt{3}/2, \quad \eta_1 = -r/2, \\
\xi_2 &= r\sqrt{3}/2, \quad \eta_2 = -r/2, \\
\xi_3 &= 0, \quad \eta_3 = r.
\end{aligned} \tag{4.3}$$

Введем обозначение

$$c \triangleq R - r \tag{4.5}$$

Дифференцируя ограничения, получим:

$$\begin{aligned}
l_i \dot{l}_i + (x_i^A - \xi_i - x)\dot{x} + (y_i^A - \eta_i - y)\dot{y} + ((x + \xi_i - x_i^A)\eta_i - (y + \eta_i - y_i^A)\xi_i)\dot{\varphi} = 0, \\
i = \overline{1,3}.
\end{aligned}$$

Соответственно уравнение, связывающее скорости актуаторов и платформы имеет вид:

$$\begin{pmatrix} l_1 \dot{l}_1 \\ l_2 \dot{l}_2 \\ l_3 \dot{l}_3 \end{pmatrix} + \begin{pmatrix} -\sqrt{3}R/2 - x + r\sqrt{3}/2 & -R/2 - y + r/2 & a_{13} \\ \sqrt{3}R/2 - x - r\sqrt{3}/2 & -R/2 - y + r/2 & a_{23} \\ -x & R - y - r & a_{33} \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\varphi} \end{pmatrix} = 0$$

Здесь

$$a_{13} = -(x - r\sqrt{3}/2 + \sqrt{3}R/2)r/2 + (y - r/2 + R/2)\sqrt{3}r/2 = (-x + y\sqrt{3})r/2$$

$$a_{23} = -(x + r\sqrt{3}/2 - \sqrt{3}R/2)r/2 - (y - r/2 + R/2)\sqrt{3}r/2 = (-x - y\sqrt{3})r/2$$

$$a_{33} = xr$$

Сингулярность соответствует тому, что возможны ненулевые движения платформы, если актуаторы зафиксированы, т.е. возможны *ненулевые* решения системы уравнений

$$\begin{pmatrix} -\sqrt{3}R/2 - x + r\sqrt{3}/2 & -R/2 - y + r/2 & (-x + \sqrt{3}y)r/2 \\ \sqrt{3}R/2 - x - r\sqrt{3}/2 & -R/2 - y + r/2 & (-x - \sqrt{3}y)r/2 \\ -x & R - y - r & xr \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\phi} \end{pmatrix} = 0$$

Вычислим определитель системы, чтобы удостовериться, что она действительно вырождена.

Сложим и вычтем первые два уравнения и преобразуем систему к виду:

$$\begin{pmatrix} -2x & -c - 2y & -xr \\ \sqrt{3}c & 0 & -\sqrt{3}yr \\ -x & c - y & xr \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\phi} \end{pmatrix} = 0$$

Вычислим определитель системы, чтобы удостовериться, что положение с $\phi=0$ является сингулярным.

$$J = \det \begin{pmatrix} -2x & -c - 2y & -xr \\ \sqrt{3}c & 0 & -\sqrt{3}yr \\ -x & c - y & -xr \end{pmatrix} = \sqrt{3}r \times$$

$$\det \begin{pmatrix} -2x & -c - 2y & -x \\ c & 0 & -y \\ -x & c - y & x \end{pmatrix} = 3\sqrt{3}r \det \begin{pmatrix} -x & -y & 0 \\ c & 0 & -y \\ -x & c - y & x \end{pmatrix} =$$

$$= 3\sqrt{3}r [-xy^2 + xyc - xy(c - y)] = 0.$$

Отсюда следует, что положение $\phi=0$ является сингулярным. Проверим, возможны ли ненулевые решения системы при нулевых скоростях штанг и нулевой угловой скорости для данного угла. С учетом обозначений (4.5) имеем:

$$\begin{pmatrix} 2x & 2y + c & xr \\ -\sqrt{3}c & 0 & \sqrt{3}yr \\ x & y - c & -xr \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\phi} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Легко видеть, что ранг системы равен двум, причем любые пары уравнений могут быть решены независимо и имеют единственное решение с нулевыми скоростями. Например, из первых двух строк имеем:

$$\begin{aligned} 2x\dot{x} + (c + 2y)\dot{y} &= 0 \\ -\sqrt{3}c\dot{x} &= 0 \end{aligned}$$

Отсюда следует, что

$$\begin{aligned} \dot{x} &= 0 \\ (c + 2y)\dot{y} &= 0 \end{aligned}$$

Отсюда возможны два варианта:

$$\dot{x} = \dot{y} = 0$$

и

$$y = -c/2 = \text{const}$$

Оба этих варианта соответствуют неподвижному положению платформы $x = x_0 = \text{const}, y = y_0 = \text{const}$.

Третье уравнение является условием совместности этих решений:

$$x\dot{x} + (y - c)\dot{y} = 0,$$

которое, очевидно, выполнено. Таким образом, даже при нулевом значении якобиана в данном конкретном случае при нулевых скоростях актуаторов платформа остается неподвижной.

Определение положений актуаторов для заданного движения платформы – это **обратная кинематическая задача**. В данном случае она решается просто, т.к. матрица при удлинении штанг диагональная. Для расчета длин штанг достаточно подставить в уравнения закон движения штанги.

Сложнее оказывается решить **прямую задачу кинематики**: определить положение и скорости рабочего инструмента для заданного закона движения актуаторов.

Таким образом, однородная система уравнений имеет только решения с нулевыми скоростями платформы при нулевых скоростях актуаторов. Это означает, что возможно движение с нулевым углом поворота, при этом удлинения штанг должны быть согласованы. Независимо могут задаваться только две длины.

Рассмотрим теперь, как определить закон движения платформы, соответствующий поступательному движению платформы при нулевом угле платформы. Запишем уравнения движения.

$$\begin{pmatrix} 2x & 2y + c & xr \\ -\sqrt{3}c & 0 & \sqrt{3}yr \\ x & y - c & -xr \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \\ 0 \end{pmatrix} = \begin{pmatrix} l_1\dot{l}_1 + l_2\dot{l}_2 \\ l_2\dot{l}_2 - l_1\dot{l}_1 \\ l_3\dot{l}_3 \end{pmatrix} \quad (4.6)$$

Введем обозначения:

$$p(t) = 0.5(l_1^2(t) + l_2^2(t)), \quad q = 0.5(l_2^2(t) - l_1^2(t)), \quad r = 0.5l_3^2 \quad (4.7)$$

В данных обозначениях система уравнений (4.6) запишется в виде:

$$\begin{aligned} 2x\dot{x} + (2y + c)\dot{y} &= \dot{p} \\ -\sqrt{3}c\dot{x} &= \dot{q} \\ \dot{r} &= x\dot{x} + (y - c)\dot{y}. \end{aligned}$$

Решения $x(t)$, $y(t)$ определяются из первых двух уравнений:

$$\begin{aligned} x &= \frac{-q(t)}{\sqrt{3}c} + C \\ \dot{y} &= \frac{\dot{p} - 2x\dot{x}}{(2y + c)} \end{aligned} \quad (4.8)$$

Движение третьей штанги не является независимым и определяется из условия, обеспечивающего поступательный характер движения, а именно интегрированием соотношения:

$$\dot{r} = x\dot{x} + (y - c)\dot{y}. \quad (4.9)$$

Если положить, что в начальный момент времени платформа находилась в центре, то

$$\begin{aligned} x(0) &= 0, \quad y(0) = 0, \\ p(0) &= 0.5((R - r)^2 + (R + r)^2) = c^2, \\ q(0) &= 0, \quad r(0) = 0.5c^2. \end{aligned} \quad (4.10)$$

Тогда, с учетом начального условия (4.10), имеем:

$$x(t) = \frac{-q(t)}{\sqrt{3}c} = \frac{l_1^2(t) - l_2^2(t)}{2\sqrt{3}c}.$$

Отсюда определяется скорость перемещения по абсциссе:

$$\dot{x}(t) = \frac{l_1\dot{l}_1(t) - l_2\dot{l}_2(t)}{\sqrt{3}c}.$$

Для определения неизвестной функции $y(t)$ нам нужно решить нелинейное обыкновенное дифференциальное уравнение

$$\dot{y} = \frac{\dot{p} - 2x\dot{x}}{(2y + c)} = \frac{\sqrt{3}c\dot{p} - (2q/(\sqrt{3}c))\dot{q}}{\sqrt{3}c(2y + c)}. \quad (4.11)$$

Перепишем уравнение (4.11) в виде:

$$\begin{aligned}\sqrt{3c}(2y+c)\dot{y} - \sqrt{3c}\dot{p} + 2q\dot{q}/(\sqrt{3c}) &= 0 \\ 2\sqrt{3c}y\dot{y} + \sqrt{3c}^2\dot{y} - \sqrt{3c}\dot{p} + 2q\dot{q}/(\sqrt{3c}) &= 0 \\ \frac{d}{dt}\left(3cy^2 + 3c^2y - 3cp + \frac{q^2}{c}\right) &= 0\end{aligned}\quad (4.12)$$

Имея в виду начальные условия (4.10) и интегрируя последнее соотношение в (4.12), получаем:

$$\left(3cy^2 + 3c^2y - 3cp + \frac{q^2}{c}\right) = -3c^3.$$

Отсюда получаем квадратное уравнение для функции $y(t)$:

$$y^2(t) + cy(t) - p(t) + \frac{q^2(t)}{3c^2} + c^2 = 0.$$

Вводя обозначение

$$Q(t) = -p(t) + \frac{q^2(t)}{3c^2} + c^2,$$

запишем уравнение для функции $y(t)$ в виде:

$$y^2(t) + cy(t) + Q(t) = 0$$

Решая это квадратное уравнение, получаем, что

$$y(t) = -0.5c \pm \sqrt{D}, \quad (4.13)$$

где

$$\begin{aligned}D &= 0.25c^2 - Q(t) = p(t) - \frac{q^2(t)}{3c^2} - 0.75c^2, \\ D(0) &= c^2 - 0.75c^2 = 0.25c^2\end{aligned}\quad (4.14)$$

Чтобы определить, какой знак следует выбрать перед корнем, проверим начальные условия. Так как $l_2 = l_1 = d$ в начальный момент времени, то из формулы (4.13) следует, что

$$\begin{aligned}y(0) &= -0.5c \pm \sqrt{D(0)} = -0.5c \pm \sqrt{D(0)} = -0.5c \pm 0.5c \\ y_1 &= 0 \\ y_2 &= -c\end{aligned}$$

Очевидно, решение, удовлетворяющее начальному условию должно иметь знак «+».

Т.е. решение системы уравнений (4.8) задается формулами

$$\begin{aligned} x(t) &= \frac{l_1^2(t) - l_2^2(t)}{2\sqrt{3}c} \\ y(t) &= -0.5c + \sqrt{D} = -0.5c + \sqrt{p(t) - \frac{q^2(t)}{3c^2} - 0.75c^2} \end{aligned} \quad (4.15)$$

С учетом обозначений (4.7), получаем, что

$$D = p(t) - \frac{q^2(t)}{3c^2} - 0.75c^2 = \frac{l_1^2(t) + l_2^2(t)}{2} - \frac{(l_2^2(t) - l_1^2(t))^2}{12c^2} - \frac{3}{4}c^2. \quad (4.16)$$

Т.е. ордината центра тяжести платформы определяется формулой

$$y(t) = -0.5c + \sqrt{\frac{l_1^2(t) + l_2^2(t)}{2} - \frac{(l_2^2(t) - l_1^2(t))^2}{12c^2} - \frac{3}{4}c^2}.$$

При этом поступательное движение возможно, только если движение третьей штанги согласовано с удлинениями первых двух согласно уравнению (4.9).

Удлинение третьей штанги находим геометрически. Длина третьей штанги равна расстоянию между точками A_3 и B_3 . Координаты этих точек в неподвижной системе координат есть $(0, R)$ и $(x(t), y(t)+r)$. Соответственно, расстояние между ними есть

$$l_3(t) = \sqrt{x^2(t) + (y(t) - c)^2} \quad (4.17)$$

Это же соотношение можно получить интегрированием кинематического соотношения (4.9):

$$\begin{aligned} \dot{r} &= \frac{1}{2} \frac{d}{dt} (x^2 + (y - c)^2) \\ r(t) &= \frac{1}{2} (x^2 + (y - c)^2) + C = \frac{1}{2} (x^2 + (y - c)^2). \\ l_3 &= \sqrt{2r(t)} = \sqrt{x^2 + (y - c)^2} \end{aligned}$$

Для проверки формулы (4.17) мы визуализировали закон движения платформы при заданном законе удлинения двух стержней A_1B_1 и A_2B_2 .

В прилагаемом далее примере геометрические параметры платформ были взяты в виде: $R=17.32$, $r=5.77$, что соответствует $c=11.55$ и сторонам платформ 30 и 10 соответственно; ограничения на длины штанг были заданы $l_{\max}=21.93$, $l_{\min}=1.15$.

Закон удлинения штанг был задан как

$$\begin{aligned} l_1(t) = f_1(t) &= (R - r) + 0.4(l_{\max} - l_{\min}) \sin(2\pi t) \\ l_3(t) = f_2(t) &= (R - r) + 0.4(l_{\max} - l_{\min}) \sin(\pi t). \end{aligned} \quad (4.18)$$

Чтобы исключить выход за пределы допустимых значений закон корректировался по формулам

$$l_1(t) = \max[\min(f_1(t), l_{\max}), l_{\min}]$$

$$l_2(t) = \max[\min(f_2(t), l_{\max}), l_{\min}].$$

Была решена прямая кинематическая задача, т.е. рассчитан закон движения платформы $x(t)$ и $y(t)$ по заданным удлинениям штанг. При этом длина третьей штанги рассчитывалась по формуле (4.17), а дискриминант D рассчитывался по формуле (4.16).

Последовательные положения платформы показаны на рисунке 2, показана также траектория центра, построенная по функциям $x(t)$ и $y(t)$.

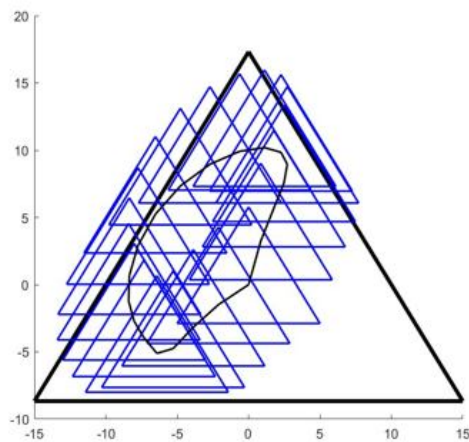


Рис. 2. Положения платформы в разные моменты времени при поступательном перемещении

Графики для функций координат центра подвижной платформы и законы удлинения штанг в данном модельном примере показаны на рисунках 3 и 4 соответственно.

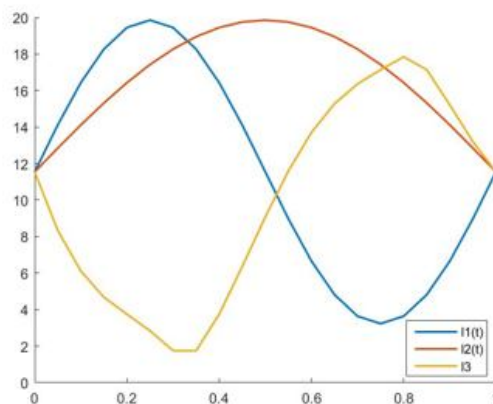


Рис. 3. Закон движения центра платформы по координатно

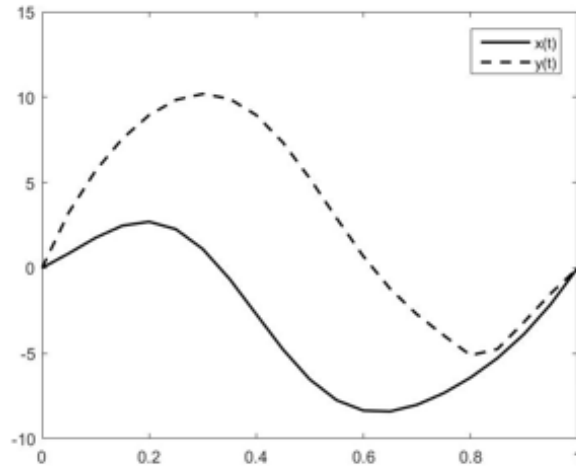


Рис. 4. Заданные удлинения первых двух штанг и рассчитанное удлинение третьей штанги с учетом ограничений

Литература

1. L.A. Rybak¹, S.Y. Khalapyan², E.V. Gaponenko, Issues of planning trajectory of parallel robots taking into account zones of singularity, - IOP Conf. Series: Materials Science and Engineering 327 (2018) 042092 doi:10.1088/1757-899X/327/4/042092
2. Merlet J. P. Parallel robots. – Springer Science & Business Media, 2006. – Т. 128.

О СПЕЦИАЛЬНЫХ КЛАССАХ ПОЗИТИВНЫХ ПРЕДПОРЯДКОВ

Бадаев С.А., Калмурзаев Б.С.

e-mail: sbadaev@gmail.com, birzhan.kalmurzayev@gmail.com

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Аннотация. Данная работа посвящена изучению структуры позитивных предпорядков на множестве натуральных чисел. Нас интересует сводимость, определенная следующим образом: позитивный предпорядок P вычислимо сводится к позитивному предпорядку Q (символический $P \leq_c Q$), если существует вычислимая функция f такая, что $xPy \Leftrightarrow f(x)Qf(y)$ для всех $x, y \in \omega$. В работе определяются такие понятия, как предполнота, слабая предполнота, ϵ -полнота, равномерно конечная предполнота и эффективная неотделимость на предпорядках. В работе показано существование универсальных позитивных предпорядков специальных - выше указанных типов. А также показано существование бесконечно много не сравнимых позитивных предпорядков таких типов. Наиболее важным случаем предпорядков являются предпорядки, которые представляют из себя отношение эквивалентности. Показано, что для любого не универсального в.п. отношения эквивалентности R существуют бесконечное

число попарно не сравнимых слабо предполных минимальных в.п. отношений эквивалентности, не сравнимых с R . Краткое изложение предмета доклада и основных содержащихся в нем выводов.

Ключевые слова: позитивный предпорядок, позитивное отношение эквивалентности, предполный предпорядок, слабо предполный предпорядок, е-полный предпорядок, равномерно конечно предполный предпорядок, эффективно неотделимые предпорядки, вычислимая сводимость.

Работа посвящена изучению позитивных предпорядков, заданных на множестве натуральных чисел ω . Предпорядком называют рефлексивное и транзитивное бинарное отношение. А под позитивным предпорядком понимаем вычислимо перечислимый предпорядок. Если для каких-то чисел x, y пара $(x, y) \in S$ для некоторого предпорядка S , то мы будем писать xSy . Изучение позитивных предпорядков только набирает обороты, поэтому научных работы по позитивным предпорядкам очень мало. А вот частный случай позитивных предпорядков – позитивные эквивалентности изучены достаточно хорошо. Классическими работами в этой области можно считать работы Ю.Л. Ершова [1], Бернарди и Сорби [2], Лахлана [3], Бадаева и т.д. Поэтому в данной работе мы попытаемся расширить некоторые понятия в отношениях эквивалентности до предпорядков. Если R и S – это предпорядки на ω , то говорят, что R вычислимо сводится (или сводится) к S (обозначается через $R \leq_c S$), если существует вычислимая функция f , такая что для любых $x, y \in \omega$ выполнено $(xRy \Leftrightarrow f(x)Sf(y))$. Предпорядки R и S называют эквивалентными (обозначается через $R \equiv S$), если имеют место соотношения $R \leq_c S$ и $S \leq_c R$. Предпорядки R и S являются вычислимо изоморфными, если существует вычислимая перестановка множества ω , сводящая R к S . Напомним некоторые понятия в отношениях эквивалентности. Позитивная эквивалентность S называется предполным, если она не тривиальная и для любой частично рекурсивной функции ϕ найдется общерекурсивная функция F такая, что $\phi(n)SF(n)$ для всех $n \in \text{dom}(\phi)$.

Определение [3]. Пусть S – позитивная эквивалентность. S называется е-полным (кратко о textension complete), если для всякой конечной функции F , и всякой $i \in \omega \setminus \text{dom}(F)$ и отношения эквивалентности R над $\{i\} \cup \text{dom}(F)$, мы можем эффективно найти j такое что, если F является вложением $\langle \text{dom}(F), R \upharpoonright \text{dom}(F) \rangle$ в $\langle \omega, S \rangle$, тогда $F \cup \{(i, j)\}$ является вложением $\langle \text{fld } R, R \rangle$ в $\langle \omega, S \rangle$.

Из той же работы Лахлана [3] известно, что все е-полные позитивные эквивалентности между собой изоморфны, и все предполные позитивные эквивалентности тоже между собой изоморфны.

Определение. Пусть P – предпорядок. Отношение эквивалентности $E \Leftrightarrow \{(x, y) : xPy \& yPx\}$ назовем носителем предпорядка P и обозначим через $\text{supp}(P)$.

Если предпорядок P является позитивным, тогда $\text{supp}(P)$ является позитивной эквивалентностью. Однако, обратное не верно. То есть существуют не позитивные предпорядки с позитивными носителями. К примеру, пусть X – не вычислимо перечислимое множество, тогда предпорядок $Q = \{(2x, 2x + 1) : x \in X\}$ не является позитивным, но его носитель $\text{supp}(Q) = Id$ является вычислимым этношением эквивалентности.

Замечание. Пусть P_1, P_2 предпорядки. Если $P_1 \leq_c P_2$, то $\text{supp}(P_1) \leq_c \text{supp}(P_2)$. Но обратное неверно. К примеру, $\text{supp}(Q) \leq_c \text{supp}(Id)$ но Q вычислимо не сводится к Id .

Определение. Предпорядок P назовем предполным (слабо предполным, e -полным, ufr , или эффективно неотделимым), если носитель предпорядка $supp(P)$ является предполным (слабо предполным, e -полным, ufr , или эффективно неотделимым).

Теорема. Для любого положительного предпорядка P существует предполный положительный предпорядок Q такой, что $P \leq_c Q$.

Доказательство. Пусть P – произвольный положительный предпорядок и E – универсальное предполное отношение эквивалентности. Так как E – универсальное отношение эквивалентности $supp(P) \leq_c E$. Пусть f – сводящая вычислимая функция $supp(P) \leq_c E$. Определим положительный предпорядок Q следующим образом: для любых $x, y \in \omega$

$$xQy \Leftrightarrow xEy \vee (\exists u)(\exists v)[xEf(u) \& yEf(v) \& uPv].$$

Не сложно увидеть, что $P \leq_c Q$ и $supp(Q) = E$. Так как $supp(Q) = E$ – предполное положительное отношение эквивалентности, Q является предполным положительным предпорядком. Так как E – предполная положительная эквивалентность, Q так же является предполным положительным предпорядком.

Следствие. (а) Существует универсальный предполный положительный предпорядок. б) Существует бесконечно много неэквивалентных предполных положительных предпорядков.

Доказательство.

(а) В качестве P , в доказательстве теоремы, достаточно взять универсальный положительный предпорядок.

(б) Возьмем вычислимую последовательность положительных предпорядков $\{P_k\}_{k \in \omega}$ со следующими свойствами:

- $supp(P_k) = Id_{k+1}$
- $(\forall i \leq j \leq k)[[i]Id_{n+1}P_k[j]Id_{n+1}]$.

То есть каждый P_k – это положительный предпорядок, фактор которого является линейным порядком, изоморфным линейному порядку $\{0 < 1 < 2 < \dots < k\}$. Очевидно, что для любых $i < j$ P_j вычислимо не сводится к P_i . В качестве P будем брать положительный предпорядок P_k и пусть Q_k – это положительный предпорядок, построенный для P_k в доказательстве теоремы. Если $Q_k \equiv_c Q_n$ для некоторых $k < n$, то из $Q_n \leq_c Q_k$ вытекает, что $P_n \leq_c P_k$, что невозможно.

Следствие. В утверждении теоремы вместо предполного предпорядка можно взять такие классы, как

- слабо предполная
- e -полная
- ufr
- $e.i.$

Список литературы

1. Ershov Yu.L., Positive equivalence. Algebra and logic, Vol. 10, (1973), Pp. 378–394.

2. Bernardi C., and Sorbi A., Classifying positive equivalence relations. Journal of Symbolic Logic, Volume 48, (1983), Pp. 529–537.

3. Lachlan A.H., A note on positive equivalence relations, Zeitschr. f. math. Logik und Grundlagen d. Math., Volume 33, (1987), Pp. 43–46.

ОБ АППРОКСИМАЦИИ ЗАДАЧИ УПРАВЛЕНИЯ ДЛЯ ИНТЕГРО-ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ ПАРАБОЛИЧЕСКОГО ТИПА

Бакирова Э.А.^{1,2}, Асанова А.Т.^{1,2}

e-mail: bakirova1974@mail.ru; assanova@math.kz

¹Институт математики и математического моделирования МОН РК,
Казахстан

²Институт информационных и вычислительных технологий МОН РК,
Казахстан

Аннотация. Исследуется линейная краевая задача с параметром для интегро-дифференциальных уравнений параболического типа. С помощью дискретизации пространственной переменной рассматриваемая задача аппроксимируется линейной двухточечной краевой задачей с параметром для системы интегро-дифференциальных уравнений. Для решения полученной задачи применяется метод параметризации. Аппроксимирующая задача сведена к эквивалентной задаче, состоящей из специальной задачи Коши для системы интегро-дифференциальных уравнений Фредгольма, краевых условий и условий непрерывности решения в точках разбиения. Получены условия однозначной разрешимости аппроксимирующей задачи в терминах разрешимости системы алгебраических уравнений относительно параметров. Установлены оценки аппроксимации решения исходной задачи решениями аппроксимирующей линейной двухточечной задачи с параметром для системы интегро-дифференциальных уравнений

Задачи управления, которые также называются краевыми задачами с параметрами и задачей идентификации параметров для системы обыкновенных дифференциальных и интегро-дифференциальных уравнений с параметрами активно исследуются в последние десятилетия. Модели, описывающие процессы реакция-диффузия, приводят к задачам управления для интегро-дифференциальных уравнений параболического типа [1-17]. Вопросы существования, единственности и устойчивости решения задач с параметрами весьма существенны для разработки численных методов идентификации параметров математических моделей, описываемых интегро-дифференциальными уравнениями параболического типа [1-17].

В предлагаемой работе исследуется линейная краевая задача с параметром для интегро-дифференциального уравнения параболического типа. Путем дискретизации пространственной переменной рассматриваемая задача аппроксимируется двухточечной краевой задачей с параметрами для системы интегро-дифференциальных уравнений Фредгольма с вырожденным ядром. С

помощью введения дополнительных параметров [18-23] как значений искомого решения в некоторых точках отрезка $[0, T]$, где рассматривается задача, полученная задача сводится к эквивалентной задаче, состоящей из специальной задачи Коши для системы интегро-дифференциальных уравнений Фредгольма, краевых условий и условий непрерывности решения в точках разбиения. Используя интегральное уравнение, эквивалентное специальной задаче Коши для системы интегро-дифференциальных уравнений Фредгольма, и свойство вырожденности ядра интегрального слагаемого, получено представление решения специальной задачи Коши через введенные параметры при предположении обратимости некоторой матрицы. С учетом этого представления из краевых условий и условий непрерывности решения построена система алгебраических уравнений относительно параметров. Установлены условия однозначной разрешимости рассматриваемой задачи в терминах разрешимости системы алгебраических уравнений относительно параметров. Получены условия разрешимости исходной задачи для интегро-дифференциального уравнения параболического типа в терминах разрешимости аппроксимирующей двухточечной краевой задачи с параметрами.

Рассматривается линейная двухточечная краевая задача с параметром для интегро-дифференциального уравнения параболического типа

$$\frac{\partial u}{\partial t} = a(x,t) \frac{\partial^2 u}{\partial x^2} + c(x,t)u + \int_0^T k(x,s)u(x,s)ds + b(x,t)\mu(x) + f(x,t), \quad (x,t) \in \Omega = (0,\omega) \times (0,T), \quad (1)$$

$$b_0(x)\mu(x) + b_1(x)u(x,0) + b_2(x)u(x,T) = \varphi_1(x), \quad x \in [0,\omega], \quad (2)$$

$$c_0(x)\mu(x) + c_1(x)u(x,0) + c_2(x)u(x,T) = \varphi_2(x), \quad x \in [0,\omega], \quad (3)$$

$$u(0,t) = \psi_1(t), \quad u(\omega,t) = \psi_2(t), \quad t \in [0,T], \quad (4)$$

где $u(x,t)$ - искомая функция, $\mu(x)$ - неизвестный функциональный параметр, функции $a(x,t) \geq a_0 > 0$, $c(x,t) \leq 0$, $k(x,t)$, $b(x,t)$, $f(x,t)$ - непрерывны по t и гельдерово непрерывны по x на Ω ; функции $b_k(x)$, $c_k(x)$, $k = \overline{0,2}$, $\varphi_1(x)$, $\varphi_2(x)$ - непрерывны на $[0,\omega]$; функции $\psi_1(t)$, $\psi_2(t)$ непрерывны на $[0,T]$. Предполагается, что граничные функции достаточно гладкие и удовлетворяют условиям согласования.

Решением краевой задачи (1)-(4) является пара функций $(u^*(x,t), \mu^*(x))$, где непрерывная на Ω , имеющая непрерывные частные производные по t первого порядка, по x второго порядка, функция $u^*(x,t)$, удовлетворяет интегро-дифференциальному уравнению (1) при $\mu(x) = \mu^*(x)$, $x \in [0,\omega]$, и краевым условиям (2)-(4).

Возьмем $\forall h > 0$ и произведем дискретизацию по x : $x_i = ih$, $i = \overline{0,P}$, $Ph = \omega$.

Введем обозначения

$$u_i(t) = u(ih,t), \quad \mu_i = \mu(ih), \quad a_i(t) = a(ih,t), \quad c_i(t) = c(ih,t), \quad b_i(t) = b(ih,t), \\ k_i(t) = k(ih,t), \quad f_i(t) = f(ih,t), \quad b_{k,i} = b_k(ih), \quad c_{k,i} = c_k(ih), \quad k = \overline{0,2}, \quad \varphi_{1,i} = \varphi_1(ih), \\ \varphi_{2,i} = \varphi_2(ih), \quad i = \overline{0,P}.$$

Задачу (1)-(3) заменяем следующей двухточечной краевой задачей для системы обыкновенных дифференциальных уравнений

$$\frac{du_i}{dt} = a_i(t) \frac{u_{i+1} - 2u_i + u_{i-1}}{2h} + c_i(t)u_i + \int_0^T k_i(s)u_i(s)ds + b_i(t)\mu_i + f_i(t), \quad i = \overline{1, P-1}, \quad (5)$$

$$b_{0,i}\mu_i + b_{1,i}u_i(0) + b_{2,i}u_i(T) = \varphi_{1,i}, \quad i = \overline{0, P}, \quad (6)$$

$$c_{0,i}\mu_i + c_{1,i}u_i(0) + c_{2,i}u_i(T) = \varphi_{2,i}, \quad i = \overline{0, P}, \quad (7)$$

$$u_0(t) = \psi_1(t), \quad u_P(t) = \psi_2(t), \quad t \in [0, T], \quad (8)$$

С учетом условия (8) из (6) и (7) получим две группы уравнений для определения μ_0 и μ_P :

$$\begin{aligned} b_{0,0}\mu_0 &= \varphi_{1,0} - b_{1,0}\psi_1(0) - b_{2,0}\psi_1(T) \\ b_{0,P}\mu_P &= \varphi_{1,P} - b_{1,P}\psi_2(0) - b_{2,P}\psi_2(T) \end{aligned} \quad , \quad (9)$$

$$\begin{aligned} c_{0,0}\mu_0 &= \varphi_{2,0} - c_{1,0}\psi_1(0) - c_{2,0}\psi_1(T) \\ c_{0,P}\mu_P &= \varphi_{2,P} - c_{1,P}\psi_2(0) - c_{2,P}\psi_2(T) \end{aligned} \quad . \quad (10)$$

Отметим, что уравнения (9) и (10) являются также условиями согласования данных. Находя из уравнений μ_0 и μ_P и подставляя их в соотношения (10) получим условия согласования исходных данных задачи (1)-(4). И наоборот, определяя из уравнений (10) μ_0 и μ_P и подставляя их в соотношения (9) снова получим условия согласования исходных данных задачи (1)-(4).

Таким образом, функции $u_0(t)$, $u_P(t)$ и параметры μ_0 , μ_P являются известными.

С учетом этого, задачу (5)-(8) перепишем в векторно-матричной форме

$$\frac{du}{dt} = A(t)u + \int_0^T K(s)u(s)ds + B(t)\mu + f(t), \quad u, \mu \in R^{P-1}, \quad t \in (0, T), \quad (11)$$

$$B_0\mu + B_1u(0) + B_2u(T) = \Phi_1, \quad \Phi_1 \in R^{P-1}, \quad (12)$$

$$C_0\mu + C_1u(0) + C_2u(T) = \Phi_2, \quad \Phi_2 \in R^{P-1}, \quad (13)$$

где $u(t) = (u_1(t), u_2(t), \dots, u_{P-1}(t))'$, $\mu = (\mu_1, \mu_2, \dots, \mu_{P-1})'$ - неизвестные функция и параметр,

$$A(t) = \begin{pmatrix} -\frac{2a_1(t)}{h^2} + c_1(t) & \frac{a_1(t)}{h^2} & 0 & \dots & 0 & 0 & 0 \\ \frac{a_2(t)}{h^2} & -\frac{2a_2(t)}{h^2} + c_2(t) & \frac{a_2(t)}{h^2} & \dots & 0 & 0 & 0 \\ 0 & \frac{a_3(t)}{h^2} & -\frac{2a_3(t)}{h^2} + c_3(t) & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & \frac{a_{P-1}(t)}{h^2} & -\frac{2a_{P-1}(t)}{h^2} + c_{P-1}(t) \end{pmatrix},$$

$$K(s) = \text{diag}\{k_1(s), k_2(s), \dots, k_{P-1}(s)\}, \quad B(t) = \text{diag}\{b_1(t), b_2(t), \dots, b_{P-1}(t)\},$$

$$f(t) = \left(\frac{a_1(t)}{h^2} \psi_1(t) + f_1(t), f_2(t), \dots, \frac{a_{P-1}(t)}{h^2} \psi_2(t) + f_{P-1}(t) \right)',$$

$$B_i = \text{diag}\{b_{i,1}, b_{i,2}, \dots, b_{i,P-1}\}, \quad C_i = \text{diag}\{c_{i,1}, c_{i,2}, \dots, c_{i,P-1}\}, \quad i = 0, 1, 2,$$

$$\Phi_1 = (\varphi_{1,1}(t), \varphi_{1,2}(t), \dots, \varphi_{1,P-1}(t))', \quad \Phi_2 = (\varphi_{2,1}(t), \varphi_{2,2}(t), \dots, \varphi_{2,P-1}(t))'.$$

Здесь $(P-1) \times (P-1)$ -матрицы $A(t)$, $B(t)$, $K(s)$ и $(P-1)$ -вектор $f(t)$ непрерывны на $[0, T]$, B_i , C_i - постоянные $(N-1) \times (N-1)$ -матрицы, $i = 0, 1, 2$, $\|u\| = \max_{i=1, P-1} |u_i|$.

Решением задачи (11)-(13) является пара $(u^*(t), \mu^*)$, где непрерывная на $[0, T]$ и непрерывно дифференцируемая на $(0, T)$ функция $u^*(t)$, удовлетворяет интегро-дифференциальному уравнению (11) при $\mu = \mu^*$ и краевым условиям (12), (13).

Для решения задачи с параметром (11)-(13) используется подход, разработанный в [24], на основе алгоритмов метода параметризации и численных методов решения задач Коши. Схема метода. Берутся точки $t_0 = 0 < t_1 < \dots < t_{N-1} < t_N = T$ и осуществляется разбиение интервала $[0, T]$ на N подинтервалов: $[0, T] = \bigcup_{r=1}^N [t_{r-1}, t_r)$, которое обозначим через Δ_N [20]. Сужение функции $u(t)$ на r -ый интервал $[t_{r-1}, t_r)$ обозначим через $x_r(t)$, т.е. $u_r(t) = u(t)$ для $t \in [t_{r-1}, t_r)$, $r = \overline{1, N}$.

Пусть $C([0, T], R^{P-1})$ - пространство непрерывных на $[0, T]$ функций $u: [0, T] \rightarrow R^{P-1}$ с нормой $\|u\|_1 = \max_{t \in [0, T]} \|u(t)\|$; $C([0, T], \Delta_N, R^{(P-1)N})$ - пространство систем функций $u[t] = (u_1(t), u_2(t), \dots, u_N(t))$, где $u_r: [t_{r-1}, t_r) \rightarrow R^n$ непрерывны на $[t_{r-1}, t_r)$ и имеют конечные левосторонние пределы $\lim_{t \rightarrow t_r-0} u_r(t)$ для всех $r = \overline{1, N}$, с нормой

$$\|u[\cdot]\|_2 = \max_{r=1, N} \sup_{t \in [t_{r-1}, t_r)} \|u_r(t)\|.$$

Вводятся дополнительные параметры $\lambda_r = u_r(t_{r-1})$, $r = \overline{1, N}$, обозначим через $\lambda_{N+1} = \mu$, затем на каждом r -ом интервале $[t_{r-1}, t_r)$ производится следующая замена функции $u_r(t)$:

$$u_r(t) = z_r(t) + \lambda_r, \quad t \in [t_{r-1}, t_r), \quad r = \overline{1, N}, \quad \text{и} \quad \mu = \lambda_{N+1}.$$

Задача (11)--(13) переходит к линейной краевой задаче с параметрами

$$\frac{dz_r}{dt} = A(t)(z_r + \lambda_r) + \sum_{j=1}^N \int_{t_{j-1}}^{t_j} K(s)(z_j(s) + \lambda_j) ds + B(t)\lambda_{N+1} + f(t), \quad t \in [t_{r-1}, t_r), \quad (14)$$

$$z_r(t_{r-1}) = 0, \quad r = \overline{1, N}, \quad (15)$$

$$B_0 \lambda_{N+1} + B_1 \lambda_1 + B_2 \lim_{t \rightarrow T-0} z_N(t) + B_2 \lambda_N = \Phi_1, \quad (16)$$

$$C_0 \lambda_{N+1} + C_1 \lambda_1 + C_2 \lim_{t \rightarrow T-0} z_N(t) + C_2 \lambda_N = \Phi_2, \quad (17)$$

$$\lambda_s + \lim_{t \rightarrow t_s-0} z_s(t) = \lambda_{s+1}, \quad s = \overline{1, N-1}, \quad (18)$$

где соотношения (18) является условием склеивания решения в точках разбиения отрезка $[0, T]$. Заметим, что условия (18) и интегро-дифференциальные уравнения (14) также обеспечивают непрерывность производных решения в этих точках.

Решением задачи с параметрами (14)-(18) является пара $(z^*[t], \lambda^*)$, где функция $z^*[t] = (z_1^*(t), z_2^*(t), \dots, z_N^*(t)) \in C([0, T], \Delta_N, R^{(P-1)N})$ с непрерывно дифференцируемыми на $[t_{r-1}, t_r)$ компонентами $z_r^*(t)$, и $\lambda^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*, \lambda_{N+1}^*) \in R^{(P-1)(N+1)}$, удовлетворяет интегро-дифференциальному уравнению с параметрами (14), начальным условиям (15), соотношениям (16)-(18) при $\lambda_j = \lambda_j^*$, $j = \overline{1, N+1}$.

Если пара $(u^*(t), \mu^*)$ является решением задачи (11)--(13), то пара $(z^*[t], \lambda^*)$ с элементами

$$z^*[t] = (z_1^*(t), z_2^*(t), \dots, z_N^*(t)) \in C([0, T], \Delta_N, R^{(P-1)N}), \quad \lambda^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*, \lambda_{N+1}^*) \in R^{(P-1)(N+1)},$$

где $\lambda_r^* = u_r^*(t_{r-1})$, $z_r^*(t) = u_r^*(t) + u_r^*(t_{r-1})$, $t \in [t_{r-1}, t_r)$, $r = \overline{1, N}$, $\lambda_{N+1}^* = \mu^* \in R^{P-1}$, будет решением задачи (14)--(18). И обратно, если пара $(\tilde{z}[t], \tilde{\lambda})$ с элементами $\tilde{z}[t] = (\tilde{z}_1(t), \tilde{z}_2(t), \dots, \tilde{z}_N(t)) \in C([0, T], \Delta_N, R^{(P-1)N})$, $\tilde{\lambda} = (\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_N, \tilde{\lambda}_{N+1}) \in R^{(P-1)(N+1)}$, является решением задачи (14)--(18), то пара $(\tilde{u}(t), \tilde{\mu})$, определяемая равенствами

$$\tilde{u}(t) = \tilde{z}_r(t) + \tilde{\lambda}_r, \quad t \in [t_{r-1}, t_r), \quad r = \overline{1, N}, \quad \tilde{u}(T) = \lim_{t \rightarrow T-0} \tilde{z}_N(t) + \tilde{\lambda}_N \quad \text{и} \quad \tilde{\mu} = \tilde{\lambda}_{N+1},$$

будет решением исходной краевой задачи с параметром (11)-(13).

Используя фундаментальную матрицу $X_r(t)$ дифференциального уравнения $\frac{dx}{dt} = A(t)x$, $t \in [t_{r-1}, t_r)$, $r = \overline{1, N}$, сведем решение специальной задачи Коши для интегро-дифференциального уравнения с параметрами (14), (15) к эквивалентной системе интегральных уравнений

$$z_r(t) = X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) A(\tau) d\tau \lambda_r + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) \sum_{j=1}^N \int_{t_{j-1}}^{t_j} K(s) [u_j(s) + \lambda_j] ds d\tau + \\ + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) B(\tau) d\tau \lambda_{N+1} + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) f(\tau) d\tau, \quad t \in [t_{r-1}, t_r), \quad r = \overline{1, N}, \quad (19)$$

Пусть $\xi = \sum_{j=1}^N \int_{t_{j-1}}^{t_j} K(s) u_j(s) ds$ и перепишем систему интегральных уравнений (19) в виде

$$z_r(t) = X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) d\tau \xi + \\ + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) \left[A(\tau) \lambda_r + \sum_{j=1}^N \int_{t_{j-1}}^{t_j} K(s) \lambda_j ds + B(\tau) \lambda_{N+1} + f(\tau) \right] d\tau, \quad t \in [t_{r-1}, t_r), \\ r = \overline{1, N}. \quad (20)$$

Умножив обе части (20) на $K(t)$, интегрируя на $[t_{r-1}, t_r]$ и суммируя по r , получим систему линейных алгебраических уравнений относительно $\xi \in R^{P-1}$

$$\xi = G(\Delta_N) \xi + \sum_{r=1}^{N+1} V_r(\Delta_N) \lambda_r + g(f, \Delta_N), \quad (21)$$

с $(P-1) \times (P-1)$ - матрицами

$$G(\Delta_N) = \sum_{r=1}^N \int_{t_{r-1}}^{t_r} K(\tau) X_r(\tau) \int_{t_{r-1}}^{\tau} X_r^{-1}(s) ds d\tau, \\ V_r(\Delta_N) = \int_{t_{r-1}}^{t_r} K(\tau) X_r(\tau) \int_{t_{r-1}}^{\tau} X_r^{-1}(s) A(s) ds d\tau + \sum_{j=1}^N \int_{t_{j-1}}^{t_j} K(\tau) X_r(\tau) \int_{t_{j-1}}^{\tau} X_r^{-1}(\tau_1) d\tau_1 d\tau \int_{t_{r-1}}^{t_r} K(s) ds, \\ r = \overline{1, N},$$

$$V_{N+1}(\Delta_N) = \sum_{r=1}^N \int_{t_{r-1}}^{t_r} K(\tau) X_r(\tau) \int_{t_{r-1}}^{\tau} X_r^{-1}(s) B(s) ds d\tau$$

и $(P-1)$ - векторами $g(f, \Delta_N) = \sum_{r=1}^N \int_{t_{r-1}}^{t_r} K(\tau) X_r(\tau) \int_{t_{r-1}}^{\tau} X_r^{-1}(s) f(s) ds d\tau$.

Систему (21) запишем в виде

$$[I - G(\Delta_N)] \xi = \sum_{r=1}^{N+1} V_r(\Delta_N) \lambda_r + g(f, \Delta_N), \quad (22)$$

где I – единичная матрица размерности $P-1$.

Специальная задача Коши (14), (15) эквивалентна системе интегральных уравнений (19). Эта система в силу вырожденности ядра будет эквивалентна системе алгебраических уравнений (21) относительно $\xi \in R^{P-1}$. Вопросы однозначной разрешимости специальной задачи Коши исследованы в [19, 20]. Установлено, что при достаточно малом шаге $h > 0 : Nh = T$ разбиения отрезка $[0, T]$ специальная задача Коши будет однозначно разрешима.

Пусть матрица $I - G(\Delta_N)$ обратима, т.е. существует $[I - G(\Delta_N)]^{-1}$. Тогда согласно (22) вектор $\xi \in R^{P-1}$ определяется равенством

$$\xi = [I - G(\Delta_N)]^{-1} \sum_{r=1}^{N+1} V_r(\Delta_N) \lambda_r + [I - G(\Delta_N)]^{-1} g(f, \Delta_N). \quad (23)$$

В (20) вместо ξ подставляя правую часть (23) получим представление функции $z_r(t)$ через λ_r , $r = \overline{1, N+1}$:

$$\begin{aligned} z_r(t) = & \sum_{j=1}^N \left\{ X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) d\tau \left[[I - G(\Delta_N)]^{-1} V_j(\Delta_N) + \int_{t_{j-1}}^{t_j} K(s) ds \right] \right\} \lambda_j + \\ & + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) A(\tau) d\tau \lambda_r + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) \left[[I - G(\Delta_N)]^{-1} V_{N+1}(\Delta_N) + B(\tau) \right] d\tau \lambda_{N+1} + \\ & + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) \left[[I - G(\Delta_N)]^{-1} g(f, \Delta_N) + f(\tau) \right] d\tau, \quad t \in [t_{r-1}, t_r], \quad r = \overline{1, N}. \quad (24) \end{aligned}$$

Введем обозначения

$$\begin{aligned} D_{r,j}(\Delta_N) &= X_r(t_r) \int_{t_{r-1}}^{t_r} X_r^{-1}(\tau) d\tau \left[[I - G(\Delta_N)]^{-1} V_j(\Delta_N) + \int_{t_{j-1}}^{t_j} K(s) ds \right], \quad r \neq j, \\ & r, j = \overline{1, N}, \\ D_{r,r}(\Delta_N) &= X_r(t_r) \int_{t_{r-1}}^{t_r} X_r^{-1}(\tau) d\tau \left[[I - G(\Delta_N)]^{-1} V_r(\Delta_N) + \int_{t_{j-1}}^{t_j} K(s) ds \right] + X_r(t_r) \int_{t_{r-1}}^{t_r} X_r^{-1}(\tau) A(\tau) d\tau, \\ D_{r,N+1}(\Delta_N) &= X_r(t_r) \int_{t_{r-1}}^{t_r} X_r^{-1}(\tau) d\tau [I - G(\Delta_N)]^{-1} V_{N+1}(\Delta_N) + X_r(t_r) \int_{t_{r-1}}^{t_r} X_r^{-1}(\tau) B(\tau) d\tau, \\ F_r(\Delta_N) &= X_r(t_r) \int_{t_{r-1}}^{t_r} X_r^{-1}(\tau) d\tau [I - G(\Delta_N)]^{-1} g(f, \Delta_N) + X_r(t_r) \int_{t_{r-1}}^{t_r} X_r^{-1}(\tau) f(\tau) d\tau, \\ & r = \overline{1, N}. \end{aligned}$$

Тогда из (24) получим

$$\lim_{t \rightarrow t_r, 0} z_r(t) = \sum_{j=1}^{N+1} D_{r,j}(\Delta_N) \lambda_j + F_r(\Delta_N), \quad r = \overline{1, N}. \quad (25)$$

Подставив соответствующие правые части (25) в краевые условия (16), (17) и условия склеивания (18), получим систему линейных алгебраических уравнений относительно параметров λ_r , $r = \overline{1, N+1}$:

$$\begin{aligned} [B_1 + D_{N,1}(\Delta_N)]\lambda_1 + B_2 \sum_{j=2}^{N-1} D_{N,j}(\Delta_N)\lambda_j + B_2[I + D_{N,N}(\Delta_N)]\lambda_N + [B_0 + D_{N,N+1}(\Delta_N)]\lambda_{N+1} = \\ = \Phi_1 - B_2 F_N(\Delta_N), \end{aligned} \quad (26)$$

$$\begin{aligned} [C_1 + D_{N,1}(\Delta_N)]\lambda_1 + C_2 \sum_{j=2}^{N-1} D_{N,j}(\Delta_N)\lambda_j + C_2[I + D_{N,N}(\Delta_N)]\lambda_N + [C_0 + D_{N,N+1}(\Delta_N)]\lambda_{N+1} = \\ = \Phi_2 - C_2 F_N(\Delta_N), \end{aligned} \quad (27)$$

$$[I + D_{s,s}(\Delta_N)]\lambda_s - [I - D_{s,s+1}(\Delta_N)]\lambda_{s+1} + \sum_{\substack{j=1 \\ j \neq s, j \neq s+1}}^{N+1} D_{s,j}(\Delta_N)\lambda_j = -F_s(\Delta_N), \quad s = \overline{1, N-1}. \quad (28)$$

Матрицу, соответствующую левой части системы уравнений (26)-(28) обозначим через $Q_*(\Delta_N)$ и систему запишем в виде

$$Q_*(\Delta_N)\lambda = -F_*(\Delta_N), \quad \lambda \in R^{(P-1)(N+1)}, \quad (29)$$

где $F_*(\Delta_N) = (-\Phi_1 + B_2 F_N(\Delta_N), -\Phi_2 + C_2 F_N(\Delta_N), F_1(\Delta_N), \dots, F_{N-1}(\Delta_N)) \in R^{(P-1)(N+1)}$.

Пусть $\alpha = \max_{t \in [0, T]} \|A(t)\|$, $\beta = \max_{t \in [0, T]} \|B(t)\|$, $L_N = \max_{r=1, N} (t_r - t_{r-1})$, $\chi = \int_0^T \|K(t)\| dt$.

Справедлива

Теорема 1. Пусть при некотором разбиении Δ_N матрицы $I - G(\Delta_N): R^{P-1} \rightarrow R^{P-1}$ и $Q_*(\Delta_N): R^{(P-1)(N+1)} \rightarrow R^{(P-1)(N+1)}$ обратимы.

Тогда задача с параметром (11)–(13) имеет единственное решение – пару $(u^*(t), \mu^*)$ для произвольных $f(t) \in C([0, T], R^n)$, $\Phi_1, \Phi_2 \in R^{P-1}$ и справедлива оценка

$$\max(\|u^*\|_1, \|\mu^*\|) \leq Y(\Delta_N) \max(\|f\|_1, \|\Phi_1\|, \|\Phi_2\|), \quad (30)$$

где

$$\begin{aligned} Y(\Delta_N) = & \left\{ e^{\alpha L_N} \left\| [I - G(\Delta_N)]^{-1} \right\| \left[\chi (e^{\alpha L_N} - 1 + e^{\alpha L_N} \chi) + \chi \right] + e^{\alpha L_N} - 1 + \right. \\ & \left. + e^{\alpha L_N} L_N \beta \left\| [I - G(\Delta_N)]^{-1} \right\| \left[\chi e^{\alpha L_N} + 1 \right] \left\| [Q_*(\Delta_N)]^{-1} \right\| \cdot [1 + \max(\|B_2\|, \|C_2\|)] \right\} \times \\ & \times \max \left\{ 1, e^{\alpha L_N} L_N \left[e^{\alpha L_N} \left\| [I - G(\Delta_N)]^{-1} \right\| \chi \right] + e^{\alpha L_N} L_N \left\| [I - G(\Delta_N)]^{-1} \right\| \left[\chi e^{\alpha L_N} + 1 \right] \right\}. \end{aligned}$$

Таким образом, Теорема 1 устанавливает условия однозначной разрешимости задачи с параметром (11)–(13) в терминах матриц $I - G(\Delta_N)$ и $Q_*(\Delta_N)$, соответствующих систем линейных алгебраических уравнений (22) и (29). Оценка (30) дает нам непрерывную зависимость решения задачи с параметром (11)–(13) от правой части системы (11) – функции $f(t)$, данных граничных условий Φ_1, Φ_2 .

Теорема 2. Пусть

- (i) выполнены все предположения относительно данных задачи (1)-(4);
- (ii) при некотором разбиении Δ_N матрицы $I - G(\Delta_N): R^{P-1} \rightarrow R^{P-1}$ и $Q_*(\Delta_N): R^{(P-1)(N+1)} \rightarrow R^{(P-1)(N+1)}$ обратимы.

Тогда существует единственное решение задачи (1)-(4) - пара $(u^*(x, t), \mu^*(x))$ для которой справедлива оценка

$$\max \left(\max_{i=1, P} \max_{x \in [(i-1)h, ih]} |u^*(x, t) - u_i^*(t)|, \max_{i=1, P} \max_{x \in [(i-1)h, ih]} |\mu^*(x) - \mu_i^*| \right) \leq \Psi(\Delta_N)h, \quad (31)$$

где $\Psi(\Delta_N)$ вычисляется через коэффициенты интегро-дифференциального уравнения (1), граничные функции и константу $Y(\Delta_N)$.

Работа выполнена в рамках проекта № AP05132455 по грантовому финансированию МОН РК на 2018-2020 гг.

Литература

1. Sloan I. and Thomée V. Time discretization of an integro-differential equation of parabolic type // SIAM Journal on Numerical Analysis, 23(1986), 1052–1061.
2. Cannon J. and Lin Y. Non-classical H1 projection and Galerkin methods for non-linear parabolic integro-differential equations // Calcolo, 25(1988), 187–201.
3. Thomée V. and Zhang N.-Y. Error estimates for semidiscrete finite element methods for parabolic integro-differential equations // Mathematics of Computation, 53(1989), 121–139.
4. Lin Y., Thomée V., and Wahlbin L. Ritz-Volterra projections to finite element spaces and applications to integro-differential and related equations // SIAM Journal on Numerical Analysis, 28(1991), 1047–1070.
5. Pani A. and Peterson E. Finite element methods with numerical quadrature for parabolic integrodifferential equations // SIAM Journal on Numerical Analysis, 33(1996), 1084–1105.
6. Lin Y. Semi-discrete finite element approximations for linear parabolic integrodifferential equations with integrable kernels // Journal of Integral Equations and Applications, 10(1998), 51–83.
7. Pani A. and Sinha R. Finite element approximation with quadrature to a time dependent parabolic integro-differential equation with nonsmooth initial data // Journal of Integral Equations and Applications, 13(2001), 35–72.
8. Pani A. and Fairweather G. H1-Galerkin mixed finite element methods for parabolic partial integro-differential equations // IMA Journal of Numerical Analysis, 22(2002), 231–252.
9. McLean W., Sloan I. and Thomée V. Time discretization via Laplace transformation of an integro-differential equation of parabolic type // IMA Journal of Numerical Analysis, 24(2004), 439–463.
10. Matache A.-M., Schwab C. and Wihler T. Fast numerical solution of parabolic integrodifferential equations with applications in finance // SIAM Journal on Scientific Computing, 27(2005), 369–393.
11. Sinha R., Ewing R. and Lazarov R. Some new error estimates of a semidiscrete finite volume element method for a parabolic integro-differential equation with nonsmooth initial data // SIAM Journal on Numerical Analysis, 43(2006), 2320–2344.
12. Zhang N. On fully discrete Galerkin approximations for partial integrodifferential equations of parabolic type // Mathematics of Computation, 60(1993), 133–166.

13. Dolejší V. and Vlasák M. Analysis of a BDF-DGFE scheme for nonlinear convection-diffusion problems // *Numerische Mathematik*, 110(2008), 405–447.
14. Volpert V. Elliptic partial differential equations Vol. 2: Reaction-Diffusing Equations, Birkhauser Springer, Basel etc., 2014.
15. Douglas J. and Jones B. Numerical methods for integro-differential equations of parabolic and hyperbolic types // *Numerische Mathematik*, 4(1962), 96–102.
16. Yanik E. and Fairweather G. Finite element methods for parabolic and hyperbolic partial integro-differential equations // *Nonlinear Analysis*, 12(1988), 785–809.
17. Pani A., Thomée V., and Wahlbin L. Numerical methods for hyperbolic and parabolic integro-differential equations // *Journal of Integral Equations and Applications*, 4(1992), 533–584.
18. Dzhumabayev, D.S. Criteria for the unique solvability of a linear boundary-value problem for an ordinary differential equation // *U.S.S.R. Computational Mathematics and Mathematical Physics*. 1989. – Vol. 29, No. 1. – P. 34-46.
19. Dzhumabaev, D.S. A method for solving the linear boundary value problem for an integro-differential equation // *Computational Mathematics and Mathematical Physics*. 2010. – Vol. 50, No. 7. – P. 1150-1161.
20. Dzhumabaev, D.S. An algorithm for solving a linear two-point boundary value problem for an integrodifferential equation // *Computational Mathematics and Mathematical Physics*. 2013. – Vol. 53, No. 6. – P. 736-758.
21. Dzhumabaev, D.S., Bakirova, E.A. Criteria for the unique solvability of a linear two-point boundary value problem for systems of integro-differential equations // *Differential Equations*. 2013. – Vol. 49, No. 9. – P. 1087-1102.
22. Dzhumabaev, D.S. Necessary and sufficient conditions for the solvability of linear boundary-value problems for the Fredholm integrodifferential equations // *Ukrainian Mathematical Journal*. 2015. – Vol. 66, No. 8. – P. 1200-1219.
23. Dzhumabaev, D.S. Solvability of a linear boundary value problem for a Fredholm integro-differential equation with impulsive inputs // *Differential Equations*. 2015. – Vol. 51, No. 9. – P. 1180-1196.
24. Dzhumabaev, D.S. On one approach to solve the linear boundary value problems for Fredholm integro-differential equations // *Journal of Computational and Applied Mathematics*. 2016. – Vol. 294, - P. 342-357.

ИЗМЕНЕНИЕ ПУБЛИКАЦИОННОЙ АКТИВНОСТИ В ОБЛАСТИ ОБРАБОТКИ ЕСТЕСТВЕННОГО ЯЗЫКА

**Барахнин В.^{1,2}, Дуйсенбаева А.⁴, Кожемякина О.¹, Кучин Я.³,
Якунин К.^{3,4}, Мухамедиев Р.^{3,4}**
e-mail: ravil.muhamedyev@gmail.com

¹*Институт вычислительных технологий СО РАН, Новосибирск, Россия*

²*Новосибирский государственный университет, Новосибирск, Россия*

³*Институт информационных и вычислительных технологий, Алматы,
Казахстан*

⁴*Казахский национальный исследовательский технический университет,
Алматы, Казахстан*

Аннотация. *Обработка естественного языка (NLP) является быстро развивающейся областью исследований. Для оценки динамики развития разделов NLP публикационная активность (ежегодное количество публикаций и цитат) рассматривается в соответствии с библиометрическими базами данных Science direct и E-library. Дифференциальные показатели были использованы для определения скорости и ускорения увеличения количества публикаций. NLP рассматривается с точки зрения как решенных проблем, так и используемых методов. Расчеты проводились для следующих задач NLP: проверка грамматики, извлечение информации, категоризация текста и т.д. Области с высокой скоростью роста (проверка грамматики, извлечение информации), машинный перевод, вопросно-ответные системы) и области, которые потеряли ранее существующую динамику роста публикационной активности (поиск информации, анализ мнений, классификация текста). В качестве методов NLP мы рассматриваем машинное обучение, нейронные сети, глубокое обучение и т.д. Наиболее динамично развивающимися методами являются те, которые связаны с машинным обучением, глубоким обучением и нейронными сетями. Публикация деятельности по решению проблем NLP с использованием машинного обучения рассматривается отдельно. Выявлено, что машинное обучение наиболее интенсивно используется в системах распознавания речи, автоматического перевода и вопросно-ответных систем.*

Ключевые слова: *Обработка естественного языка, Машинное обучение, Библиометрические индикаторы, Scientometrics, Глубокое обучение*

1 Введение

Область обработки естественного языка (NLP) характеризуется широким спектром методов и задач, некоторые из которых уже имеют приемлемые решения, реализованные в виде программного обеспечения, в то время как другие требуют интенсивных исследований.

Развитие каждого научного направления, в том числе обработка текстов на естественном языке, сопровождается повышением или уменьшением интереса исследователей, что отражается в изменении библиометрических показателей. Последний включает количество публикаций, индекс цитирования, количество соавторов, индекс Хирша и другие. Выявление «горячих» областей, в которых эти показатели более важны, позволяет нам лучше понять ситуацию в науке и, если возможно, сосредоточить усилия на прорывных областях.

Для выявления логики изменения публикационной активности в [1] реализованы дифференциальные показатели. Их применение позволяет оценить скорость и ускорение изменений в библиометрических показателях. Таким образом, внедренные индикаторы могут более наглядно показать рост или снижение интереса исследователей к определенным разделам NLP, характеризующимся определенными ключевыми словами.

В данной статье количество публикаций и количество ссылок на статьи с выбранными ключевыми терминами рассматриваются как анализируемые показатели. Дифференциальные метрики позволяют оценить динамику изменения использования выбранных ключевых терминов авторами научных публикаций, что косвенно свидетельствует о росте или снижении интереса исследователей к научной области, обозначаемой этим термином.

2 Цели исследования

В рамках этой работы мы определили следующие области анализа.

Во-первых, мы рассматриваем область NLP с точки зрения задач, которые должны быть решены, среди которых, основываясь на вышеупомянутом обзоре и таксономии NLP в [2], мы включаем следующее (группа «Задачи NLP» или «задачи»): проверка грамматики, извлечение информации, категоризация текста, диалоговые системы, распознавание речи, машинный перевод, поиск информации, вопросно-ответные системы, анализ мнений и анализ тональности, интеллектуальные советники, автоматическое реферирование, поиск информации.

Во-вторых, область NLP характеризуется быстрым ростом технологий и методов, способствующих решению вышеуказанных задач. Мы включаем следующие методы (группа «научные методы NLP» или «техника»): машинное обучение, нейронные сети, глубокое обучение, нечеткая логика, логика первого порядка, представление знаний, эволюционные вычисления и генетическое программирование, система, основанная на правилах, неконтролируемое обучение. Кластеризация, контролируемое обучение, статистические методы, байесовские сети, семантические сети, определение ключевых слов, лексическая близость, онтология, слияние информации, таксономия.

3 Методология исследования

Для оценки динамики изменений публикационной активности мы, как и в [1], будем использовать показатели совокупного годового темпа роста (Compound Annual Growth Rate - CAGR), D1 и D2, которые рассчитываются следующим образом:

$$CAGR = \left(\frac{\text{Конечное значение}}{\text{Начальное значение}} \right)^{\frac{1}{T-1}} - 1, \quad (1)$$

где T количество периодов.

$$D1_i^j(t_k) = \beta \times \frac{dn_i^j(t_k)}{dt} + \gamma \times \frac{dc_i^j(t_k)}{dt}, \quad (2)$$

$$D2_i^j(t_k) = \beta' \times \frac{d(dn_i^j(t_k)/dt)}{dt} + \gamma' \times \frac{d(dc_i^j(t_k)/dt)}{dt} \quad (3)$$

где n_i и c_i количество публикаций и количество ссылок, которые определяются номером поискового запроса i , и $\beta, \gamma, \beta', \gamma'$ являются некоторыми эмпирическими коэффициентами, которые регулируют «вес» вклада числа публикаций, скорость и ускорение изменения количества публикаций n_i и скорость и ускорение изменения количества цитат c_i , соответственно.

Одна из крупнейших библиометрических баз данных - Science Direct, которая содержит около 2500 научных журналов и 26000 электронных книг [3], была использована для анализа публикаций на английском языке.

Ведущая российская научная электронная библиотека eLIBRARY.ru использовалась для изучения динамики публикаций на эту тему на русском языке. По состоянию на середину 2018 года в базе данных eLIBRARY.ru насчитывается более 30,7 миллиона статей, в том числе публикаций последних десятилетий XX века. По нашим запросам были найдены некоторые публикации указанного периода, но их количество было очень небольшим, поэтому статистика в таблицах

приводится с 2005 года. Подчеркнем, что, хотя в базе данных eLIBRARY.ru есть много англоязычных публикаций из зарубежных. В изданиях мы ограничивались запросами только на русском языке, поскольку изучение англоязычных публикаций проводилось отдельно для базы данных ScienceDirect.

Для каждой библиометрической базы данных были сформированы соответствующие поисковые запросы, в том числе вышеуказанные термины в сочетании с NLP, NLP и ML и т.д. Результатами запросов были годовые объемы публикаций и цитат с 2005 года. Данные за 2018 год не используются из-за их неполноты.

На основании результатов, полученных для каждого поискового запроса, были рассчитаны показатели CAGR, D1, D2. Константы $\beta, \gamma, \beta', \gamma'$ предполагаются равными 0,95.

4 Анализ данных и интерпретация

Анализ статистических данных, описанных выше, выявил следующие тенденции.

В настоящее время количество публикаций и цитат практически во всех (за некоторыми исключениями) рассматриваемых областях демонстрирует рост. Использование дифференциальных индикаторов D1 и D2 позволяет более четко представить динамику этого роста.

В русскоязычном сегменте обработка естественного языка, по-видимому, достигла уровня насыщения, когда за двумя периодами взрывного роста числа публикаций следует замедление роста практически во всех рассматриваемых областях исследований.

На рисунке 1 указаны разделы группы «Задачи NLP» с ростом индекса D2 (ускорение). Аналогичным образом, на рисунке 2 раскрываются разделы группы «NLP и ML», которые демонстрируют положительное значение показателя D2 (ускорение).

Наконец, на рисунке 3 показаны разделы группы «Научные методы NLP», которые можно отнести к наиболее популярным, с большим и возрастающим положительным значением D2 (ускорение).

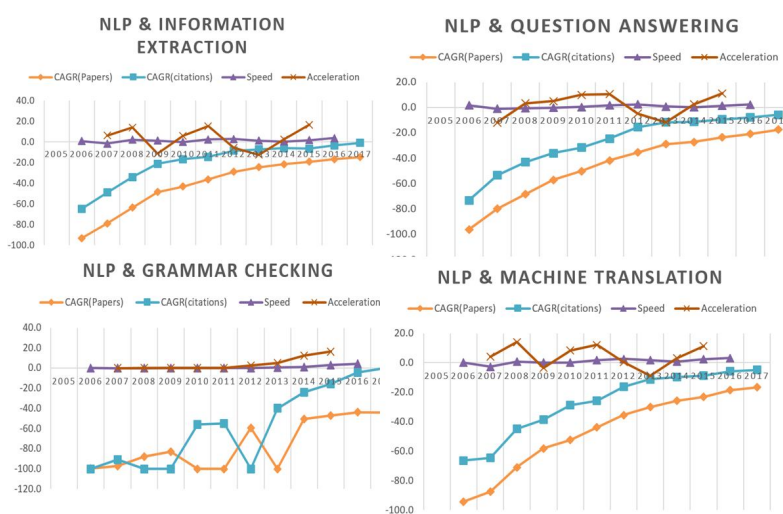


Рис. 1. Разделы группы «NLP задачи» с ростом показателя D2 (ускорение)

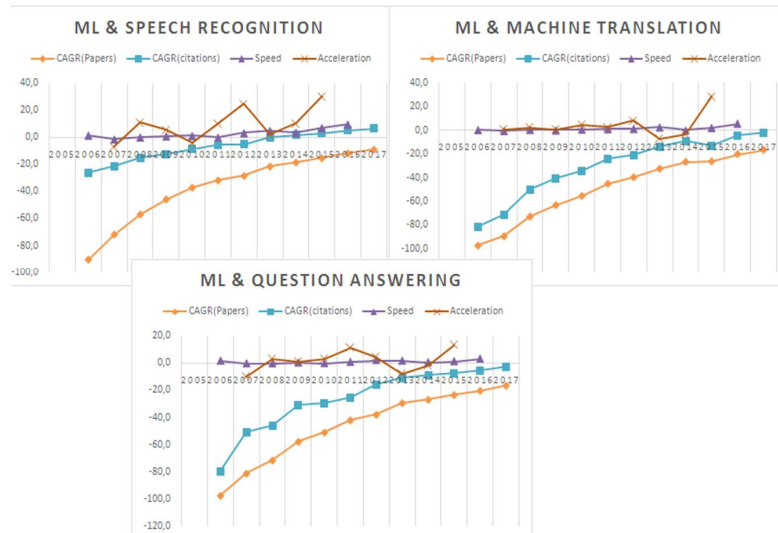


Рис. 2. Разделы группы «NLP & ML», демонстрирующие положительное значение D2 (ускорение)

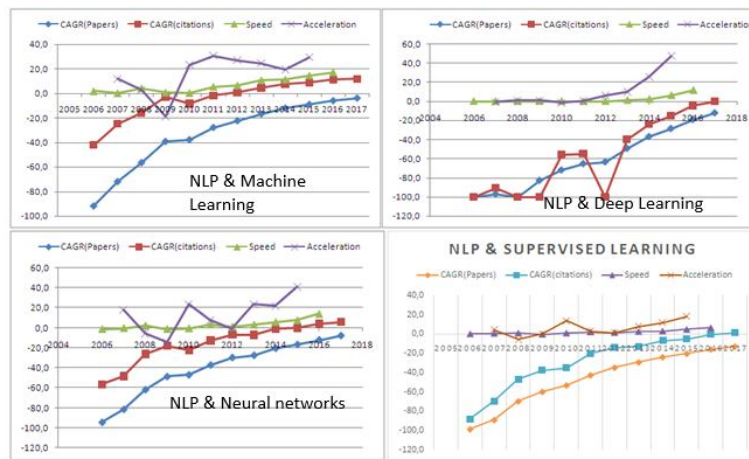


Рис. 3. Группа «Научные методы NLP», которую можно отнести к наиболее популярным, с большим и возрастающим положительным значением D2 (ускорение)

5 Заключение

Данные показывают интересную картину изменений для большинства разделов, где за периодом начального роста следует уменьшение, а затем происходит повторное ускорение (ответы на вопросы, распознавание речи). Причины этого явления следует исследовать, но можно предположить, что эта динамика характеризует, с одной стороны, интенсивность развития в области исследований, а с другой - понимание новой концепции исследователями и ее применение в исследованиях. Только некоторые из представленных областей исследований (проверка грамматики, NLP и глубокое обучение) характеризуются постоянным ростом цитирований и публикаций (показатель D2 имеет только положительные значения за весь рассматриваемый период ниже) (рисунки 1, 3).

В целом, анализ области исследований NLP показывает значительную динамику роста публикационной активности в разделах «Проверка грамматики»,

«Извлечение информации», «Машинный перевод» и «Ответы на вопросы» (положительное значение D2) (рисунки 1,2).

В области научных методов высокая положительная ценность D2 принадлежит исследованиям, связанным с машинным обучением, глубоким обучением, нейронными сетями и, как следствие, так называемым обучением с учителем. В этом случае машинное обучение особенно интенсивно используется при решении задач распознавания речи, автоматического перевода и вопросно-ответных систем (Рис.3). Умеренно положительное значение D2 отмечено для разделов «Эволюционные вычисления и генетическое программирование», «Обучение без учителя», «Байесовские сети», «Семантические сети». Нечеткая логика, представление знаний, система, основанная на правилах, статистические методы, лексическая близость, онтология имеют отрицательное значение D2.

Работа финансировалась за счет гранта BR05236839 Министерства образования и науки Республики Казахстан «Разработка информационных технологий и систем для стимулирования устойчивого развития личности как одна из основ развития цифрового Казахстана».

Литература

1. Muhamedyev R. et al. New bibliometric indicators for prospectivity estimation of research fields //Annals of Library and Information Studies (ALIS). – 2018. – Т. 65. – №. 1. – С. 62-69.

2. Natural language Processing.

https://en.wikipedia.org/wiki/Natural_language_processing

3. ElsevierScienceDirect.

https://en.wikipedia.org/wiki/Wikipedia:Elsevier_ScienceDirect

АВТОМАТИЗАЦИЯ ИССЛЕДОВАНИЯ ТЕРМО- НАПРЯЖЕННОГО СОСТОЯНИЯ СТЕРЖНЯ ОГРАНИЧЕННОЙ ДЛИНЫ ПРИ ВОЗДЕЙСТВИИ ЛОКАЛЬНОГО БОКОВОГО ТЕПЛООВОГО ПОТОКА

**Бегалиева К.Б.^{1,2}, Аршидинова М.Т.^{1,2}, Кудайкулов А.¹,
Ташев А.А.¹**

kalamkas_b@mail.ru, mukaddas_arshidi@mail.ru, kudaykulov2006@mail.ru,
azattash@mail.ru

¹*Institute of Information and Computational Technologies CS MES RK*

²*Al-Farabi Kazakh National University, Almaty, Kazakhstan*

Аннотация: Целью данной работы является автоматизировать процесс исследования термо-напряженно-деформированного состояния стержня постоянного сечения и ограниченной длины. При этом на замкнутой локальной поверхности подведен тепловой поток постоянной интенсивности. Остальные части боковой поверхности теплоизолированы. Через площади поперечных

сечений двух концов стержня происходит конвективный теплообмен с окружающими этих площадей сред. При этом коэффициенты теплообмена и температуры окружающих сред могут быть разными. Процесс автоматизации исследования опирается на законы сохранения энергии. На основе последнего строится разрешающие системы линейных алгебраических уравнений с учетом существующих естественных граничных условий. Решая построенные системы строится закон распределения по длине исследуемого стержня температуры, трех составляющих деформаций и напряжения, а также перемещения. Определяются величины удлинения стержня и возникающего осевого усилия.

Ключевые слова: локальный тепловой поток, теплоизоляция, удлинение, осевая сжимающая сила, температура, перемещение, деформация, напряжение.

1. Введение

Данная работа посвящена к автоматизации исследования термо-напряженно-деформированного состояния стержня постоянного сечения и ограниченной длины. При этом на замкнутой локальной боковой поверхности стержня подводится тепловой поток постоянной интенсивности. Остальная часть боковой поверхности стержня полностью теплоизолирована. Через площади поперечных сечений двух концов стержня происходит конвективный теплообмен с окружающей средой. При этом коэффициенты теплообменов и температуры окружающих сред двух концов стержня разные. Для определения поля температуры используется закон сохранения энергии в форме функционала полной тепловой энергии с учетом существующих разнородных видов источников тепла, физико-механических свойств материала стержня и его геометрических размеров. Пользуясь этим строится закон распределения температуры по длине исследуемого стержня. По ним определяется величина удлинения стержня, в случае заземления одного конца. Если оба конца стержня заземлены, то определяется величина возникающего осевого усилия. Также определяются законы распределения термоупругих, температурных и упругих составляющих деформаций, и напряжений в зависимости от наличия локального теплового потока, теплоизоляций теплообменов, геометрии стержня и физико-механических свойств материала стержня. Для определения поля перемещения используется закон сохранения энергии в форме функционала потенциальной энергии упругой деформации с учетом наличия поля температуры. Далее определяется поле перемещения с учетом реальных условий эксплуатации. Разработанная программа позволяет варьировать значениями исходных данных.

В данной работе для решения конкретной прикладной задачи используется фундаментальные законы сохранения энергии в сочетании построенных квадратичных сплайн функции.

2. Постановка задачи

Рассмотрим горизонтальный стержень ограниченной длины $L[cm]$, и постоянного поперечного сечения $F[cm^2]$. Горизонтальную ось ox направим слева на право. Она совпадает с осью стержня. Боковые поверхности участков $(0 \leq x \leq l_1)$ и $(l_2 \leq x \leq L)$ стержня полностью теплоизолированы. На боковой поверхности участка $(l_1 \leq x \leq l_2)$ стержня подводится тепловой поток постоянной интенсивности $q \left[\frac{W}{cm^2} \right]$. Через площадь поперечных сечений левого и правого

конца стержня происходит теплообмен с окружающей средой. Коэффициенты теплообменов h_1 (при $x=0$) и h_2 (при $x=L$ $\left[\frac{W}{cm^2 K} \right]$). Температура окружающих сред этих площадей T_{oc1} и T_{oc2} [$^{\circ}K$] - соответственно. Расчетная схема задачи приводится на рисунке 1.

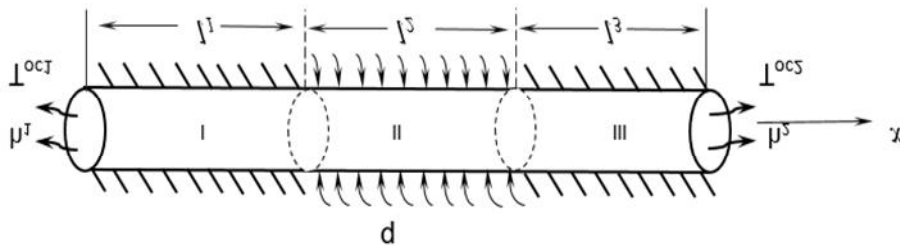


Рис. 1. Расчетная схема задачи

3. Решение проблемы с использованием закона сохранения энергии

Рассматриваемый стержень дискретизируем элементами длиной $L[cm]$. В пределах длины одного дискретного элемента поле температуры аппроксимируем полным полиномом второго порядка

$$T(x) = ax^2 + bx + c, \quad 0 \leq x \leq l \quad (1)$$

где a,b,c- постоянные, значение которых пока неизвестны. Закон распределения температуры в пределах длины одного дискретного элемента приводится на рисунке 2.

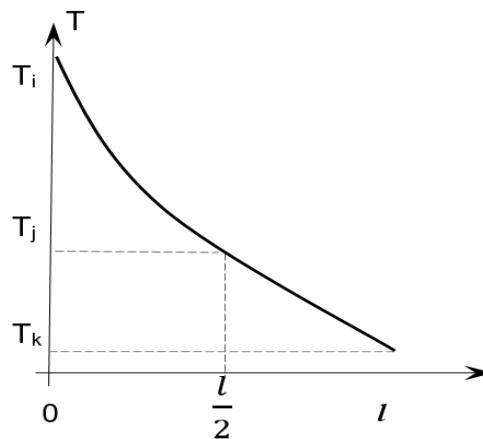


Рис. 2. Закон распределения температуры по длине одного дискретного элемента

В локальной системе координат $0 \leq x \leq l$ фиксируем три узла с координатами $x_i = 0$; $x_j = \frac{l}{2}$; $x_k = l$.

Значение температуры в этих сечениях обозначим соответственно

$$T(x = x_i = 0) = T_i; \quad T(x = x_j = \frac{l}{2}) = T_j; \quad T(x = x_k = l) = T_k \quad (2)$$

Тогда подставляя (2) в (1) получим следующую систему:

$$\left. \begin{aligned} a \cdot 0 + b \cdot 0 + c &= T_i \\ a \cdot \left(\frac{l}{2}\right)^2 + b \cdot \left(\frac{l}{2}\right) + c &= T_j \\ a \cdot l^2 + b \cdot l + c &= T_k \end{aligned} \right\} \quad (3)$$

Решая систему, определим, что

$$c = T_i; \quad b = \frac{4T_j - 3T_i - T_k}{l}; \quad a = \frac{2T_k + 2T_i - 4T_j}{l^2} \quad (4)$$

Подставляя (4) в (1) имеем

$$\begin{aligned} T(x) &= \frac{2T_k + 2T_i - 4T_j}{l^2} \cdot x^2 + \frac{4T_j - 3T_i - T_k}{l} \cdot x + T_i = \\ &= \left(\frac{2x^2 - 3lx + l^2}{l^2}\right) T_i + \left(\frac{4lx - 4x^2}{l^2}\right) T_j + \left(\frac{2x^2 - lx}{l^2}\right) T_k, \quad 0 \leq x \leq l \end{aligned} \quad (5)$$

Здесь введем обозначения

$$\varphi_i(x) = \frac{2x^2 - 3lx + l^2}{l^2}; \quad \varphi_j(x) = \frac{4lx - 4x^2}{l^2}; \quad \varphi_k(x) = \frac{2x^2 - lx}{l^2}, \quad 0 \leq x \leq l \quad (6)$$

Эти функции назовем квадратичными сплайн функциями в местной системе координат. Они имеют следующие свойства

$$\begin{aligned} \varphi_i(x) &= \begin{cases} 1, & \text{при } x = 0 \\ 0, & \text{при } x = \frac{l}{2}; \\ 0, & \text{при } x = l \end{cases} & \varphi_j(x) &= \begin{cases} 0, & \text{при } x = 0 \\ 1, & \text{при } x = \frac{l}{2}; \\ 0, & \text{при } x = l \end{cases} \\ \varphi_k(x) &= \begin{cases} 0, & \text{при } x = 0 \\ 0, & \text{при } x = \frac{l}{2}; \\ 1, & \text{при } x = l \end{cases} \end{aligned} \quad (7)$$

$$\varphi_i(x) + \varphi_j(x) + \varphi_k(x) = 1; \quad 0 \leq x \leq l \quad (8)$$

Градиент температуры в пределах длины одного дискретного элемента в местной системе координат имеет следующий вид:

$$\begin{aligned} \frac{dT}{dx} &= \frac{d\varphi_i(x)}{dx} T_i + \frac{d\varphi_j(x)}{dx} T_j + \frac{d\varphi_k(x)}{dx} T_k = \frac{4x - 3l}{l^2} T_i + \frac{4l - 8x}{l^2} T_j + \frac{4x - l}{l^2} T_k, \\ &0 \leq x \leq l \end{aligned} \quad (9)$$

Здесь следует отметить, что

$$\frac{d\varphi_i(x)}{dx} T_i + \frac{d\varphi_j(x)}{dx} T_j + \frac{d\varphi_k(x)}{dx} T_k = 0 \quad (10)$$

Для построения разрешающих систем уравнений с учетом естественных граничных условий дискретизируем исследуемый стержень тремя элементными разной длины. Для первого дискретного элемента функционал полной энергии с учетом теплоизоляции боковой поверхности имеет следующий вид:

$$J_1 = \int_{S(x=0)} \frac{h_1}{2} (T - T_{oc1})^2 ds + \int_{V_1} \frac{K_{xx}}{2} \left(\frac{\partial T}{\partial x} \right)^2 dv, \quad (0 \leq x \leq l_1) \quad (11)$$

Здесь первый интеграл по площади поперечного сечения левого конца имеет место только для точек этой поверхности. В локальной системе координат с учетом глобальной нумерации узлов можно переписать (11) в следующем виде:

$$\begin{aligned} J_1 &= \frac{F(x=0)h_1}{2} (T - T_{oc1})^2 + \frac{FK_{xx}}{2} \int_0^{l_1} \left[\left(\frac{4x-3l_1}{l_1^2} T_1 + \frac{4l_1-8x}{l_1^2} T_2 + \frac{4x-l_1}{l_1^2} T_3 \right) \right]^2 dx = \\ &= \frac{Fh_1}{2} (T_1 - T_{oc1})^2 + \frac{FK_{xx}}{6l_1} [7T_1^2 - 16T_1T_2 + 2T_1T_3 - 16T_2T_3 + 16T_2^2 + 7T_3^2] \end{aligned} \quad (12)$$

где h_l - коэффициент теплообмена. T_{oc1} – температура окружающей среды левого конца стержня.

Здесь надо сказать, что боковая поверхность первого дискретного элемента полностью теплоизолирована. В выражении (12), следует отметить, что сумма коэффициентов перед узловыми значениями температур будут равны нулю. Например, в первой скобке $(1-1) = 0$, а во второй скобке $[7-16+2-16+16+7] = 0$. Теперь на рисунке-1 рассмотрим второй дискретный элемент. Этот элемент является внутренней. Но на боковой поверхности этого элемента подведен тепловой постоянной интенсивности $q \left[\frac{W}{cm^2} \right]$. Поэтому для этого элемента функционал полной тепловой энергии будет иметь следующий вид:

$$\begin{aligned} J_2 &= \int_{V_2} \frac{K_{xx}}{2} \left(\frac{\partial T}{\partial x} \right)^2 dv + \int_{S_{\text{бок}}} q T ds = \frac{FK_{xx}}{2} \int_0^{l_2} \left[\left(\frac{4x-3l_2}{l_2^2} T_3 + \frac{4l_2-8x}{l_2^2} T_4 + \frac{4x-l_2}{l_2^2} T_5 \right) \right]^2 dx + \\ &+ qP \int_0^{l_2} \left[\left(\frac{2x^2-3l_2x+l_2^2}{l_2^2} \right) T_3 + \left(\frac{4l_2x-4x^2}{l_2^2} \right) T_4 + \left(\frac{2x^2-l_2x}{l_2^2} \right) T_5 \right] dx = \frac{FK_{xx}}{6l_2} [7T_3^2 - 16T_3T_4 + \\ &+ 2T_3T_5 - 16T_4T_5 + 16T_4^2 + 7T_5^2] + \frac{q\pi r l_2}{3} (T_3 + 4T_4 + T_5), \quad (l_1 \leq x \leq l_2) \quad (13) \end{aligned}$$

Здесь V_2 - объем второго дискретного элемента; $P=2\pi r$ – периметр поперечного сечения; Наконец переходим к последнему третьему дискретному элементу. Боковая поверхность этого элемента полностью теплоизолирована. Но через площади поперечного сечения правого конца происходит теплообмен с

окружающей ее средой. При этом коэффициент теплообмена h_2 , а температура окружающей среды T_{oc2} . Длина этого элемента – l_3 . Функционал полной тепловой энергии для третьего дискретного элемента будет следующей

$$\begin{aligned}
 J_3 &= \int_{V_3} \frac{K_{xx}}{2} \left(\frac{\partial T}{\partial x} \right)^2 dv + \int_{S(x=l_1+l_2+l_3)} \frac{h_2}{2} (T - T_{oc2})^2 ds = \\
 &= \frac{FK_{xx}}{2} \int_0^{l_3} \left[\left(\frac{4x-3l_3}{l_3^2} \right) T_5 + \left(\frac{4l_3-8x}{l_3^2} \right) T_6 + \left(\frac{4x-l_3}{l_3^2} \right) T_7 \right]^2 dx + \frac{Fh_2}{2} (T_7 - T_{oc2})^2 = \\
 &= \frac{FK_{xx}}{6l_3} [7T_5^2 - 16T_5T_6 + 2T_5T_7 - 16T_6T_7 + 16T_6^2 + 7T_7^2] + \frac{Fh_2}{2} (T_7 - T_{oc2})^2, \\
 & \quad (l_2 \leq x \leq L) \tag{14}
 \end{aligned}$$

где V_3 - объем третьего дискретного элемента; F – площадь поперечного сечения стержня; Тогда функционал полной тепловой энергии для исследуемого стержня имеет следующий вид:

$$\begin{aligned}
 J = J_1 + J_2 + J_3 &= \frac{Fh_1}{2} (T_1 - T_{oc1})^2 + \frac{FK_{xx}}{6l_1} [7T_1^2 - 16T_1T_2 + 2T_1T_3 - 16T_2T_3 + 16T_2^2 + 7T_3^2] + \\
 &+ \frac{FK_{xx}}{6l_2} [7T_3^2 - 16T_3T_4 + 2T_3T_5 - 16T_4T_5 + 16T_4^2 + 7T_5^2] + \frac{q\pi r l_2}{3} (T_3 + 4T_4 + T_5) + \\
 &+ \frac{FK_{xx}}{6l_3} [7T_5^2 - 16T_5T_6 + 2T_5T_7 - 16T_6T_7 + 16T_6^2 + 7T_7^2] + \frac{Fh_2}{2} (T_7 - T_{oc2})^2 \tag{15}
 \end{aligned}$$

Для построения разрешающих системы линейных алгебраических уравнений относительно узловых значений температур, минимизируется функционал J по T_1, T_2, \dots, T_7 .

$$\left. \begin{aligned}
 1) \frac{\partial J}{\partial T_1} = 0; &\Rightarrow Fh_1(T_1 - T_{oc1}) + \frac{FK_{xx}}{6l_1} [14T_1 - 16T_2 + 2T_3] = 0; \\
 2) \frac{\partial J}{\partial T_2} = 0; &\Rightarrow \frac{FK_{xx}}{6l_1} [-16T_1 + 32T_2 - 16T_3] = 0; \\
 3) \frac{\partial J}{\partial T_3} = 0; &\Rightarrow FK_{xx} \left[\left(\frac{2T_1 + 16T_2 + 14T_3}{6l_1} \right) + \left(\frac{14T_3 - 16T_4 + 2T_5}{6l_2} \right) \right] + \frac{q\pi r l_2}{3} = 0; \\
 4) \frac{\partial J}{\partial T_4} = 0; &\Rightarrow \frac{FK_{xx}}{6l_2} [-16T_3 + 32T_4 - 16T_5] + \frac{4q\pi r l_2}{3} = 0; \\
 5) \frac{\partial J}{\partial T_5} = 0; &\Rightarrow FK_{xx} \left[\left(\frac{2T_3 - 16T_4 + 14T_5}{6l_2} \right) + \left(\frac{14T_5 - 16T_6 + 2T_7}{6l_3} \right) \right] + \frac{q\pi r l_2}{3} = 0; \\
 6) \frac{\partial J}{\partial T_6} = 0; &\Rightarrow \frac{FK_{xx}}{6l_3} [-16T_5 + 32T_6 - 16T_7] = 0; \\
 7) \frac{\partial J}{\partial T_7} = 0; &\Rightarrow \frac{FK_{xx}}{6l_3} [2T_5 - 16T_6 + 14T_7] + Fh_2(T_7 - T_{oc2}) = 0.
 \end{aligned} \right\} \tag{16}$$

Решая эту систему вычисляются узловые значения температур T_1, T_2, \dots, T_7 . По ним строится закон распределения температуры по длине трех дискретных участков стержня:

$$\left. \begin{aligned} T^{(I)}(x) &= \varphi_i(x) \cdot T_1 + \varphi_j(x) \cdot T_2 + \varphi_k(x) \cdot T_3 \\ T^{(II)}(x) &= \varphi_i(x) \cdot T_3 + \varphi_j(x) \cdot T_4 + \varphi_k(x) \cdot T_5 \\ T^{(III)}(x) &= \varphi_i(x) \cdot T_5 + \varphi_j(x) \cdot T_6 + \varphi_k(x) \cdot T_7 \end{aligned} \right\} \quad (17)$$

или
$$T(x) = T^I(x) + T^{II}(x) + T^{III}(x).$$

Если коэффициент теплового расширения материала стержня $\alpha \left[\frac{1}{K} \right]$, является постоянной величиной, то величина термического удлинения стержня $\Delta l_T [cm]$, в случае заземления одним концом стержня определяется в соответствии теории теплофизики [13]:

$$\Delta l_T = \int_0^L \alpha \cdot T(x) dx. \quad (18)$$

где
$$T(x) = T^I(x) + T^{II}(x) + T^{III}(x) \quad (19)$$

Тогда подставляя (19) в (18) получим, что

$$\begin{aligned} \Delta l_T &= \alpha \left[\frac{l_1}{6} T_1 + \frac{2l_1}{3} T_2 + \frac{l_1}{6} T_3 + \frac{l_2}{6} T_3 + \frac{2l_2}{3} T_4 + \frac{l_2}{6} T_5 + \frac{l_3}{6} T_5 + \frac{2l_3}{3} T_6 + \frac{l_3}{6} T_7 \right] = \\ &= \frac{\alpha}{6} [l_1 T_1 + 4l_1 T_2 + (l_1 + l_2) T_3 + 4l_2 T_4 + (l_2 + l_3) T_5 + 4l_3 T_6 + l_3 T_7] \quad (20) \end{aligned}$$

Если оба конца стержня заземлены, то она не может удлиниться. Но возникает осевое сжимающее усилие $R [kG]$, которое определяется условия совместности деформации. Суть этого подхода заключается в следующем. Сначала рассмотрим горизонтальный стержень заземленный левым концом (рисунок-3).

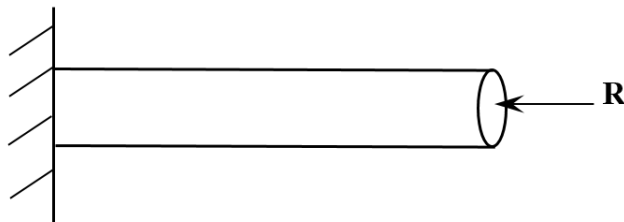


Рис. 3. Заземленный левым концом стержень под воздействием сжимающего усилия $R [kG]$

Этот стержень находится под воздействием сжимающего усилия R которая приложена на правом свободном конце. Тогда по закону Гука она сокращается на величину Δl_R

$$\Delta l_R = \frac{R \cdot L}{E \cdot F} \quad (21)$$

где $L = l_1 + l_2 + l_3$ – общая длина исследуемого стержня; $E \left[\frac{kG}{cm} \right]$ – модуль упругости материала стержня; $F [cm^2]$ – площадь стержня.

Если оба конца исследуемого стержня жестко защемлены, то естественно она не удлиняется и не укорачивается, т.е.

$$\Delta l_R + \Delta l_T = 0 \quad (22)$$

Тогда с учетом (21) определим величину возникающего в защемленном двумя концами исследуемого стержне осевую сжимающую силу R .

$$R = - \frac{E \cdot F \cdot \Delta l_T}{L} =$$

$$= - \frac{E \cdot F \cdot \alpha}{6L} [l_1 T_1 + 4l_1 T_2 + (l_1 + l_2) T_3 + 4l_2 T_4 + (l_2 + l_3) T_5 + 4l_3 T_6 + l_3 T_7] \quad (23)$$

В этом случае также возникает поле распределения термо-упругой составляющей напряжения $\sigma \left[\frac{kG}{cm^2} \right]$, которое определяется в соответствии закона Гука [1].

$$\sigma(x) = \frac{R}{F} = const \quad (24)$$

Тогда по закону Гука, поле распределения термо-упругой составляющей деформации имеет следующий вид:

$$\varepsilon(x) = \frac{\sigma(x)}{E} = const \quad (25)$$

Поле распределения температурной составляющей деформаций и напряжения определяются на основе общих теорий термоупругости:

$$\varepsilon_T(x) = -\alpha T(x), \quad (26)$$

$$\sigma_T(x) = E \cdot \varepsilon_T(x) = -\alpha E T(x), \quad (27)$$

Из этих соотношений в соответствии теории термоупругости определяются поле распределение упругих составляющих деформаций и напряжения:

$$\varepsilon_x(x) = \varepsilon - \varepsilon_T = \frac{\sigma(x)}{E} + \alpha \cdot T(x), \quad (28)$$

$$\sigma_x(x) = E \cdot \varepsilon_x(x) = \sigma(x) - \sigma_T(x), \quad (29)$$

Поле перемещения по длине одного дискретного элемента длиной l [см] аппроксимируем полным полиномом второго порядка

$$U(x) = \varphi_i(x)U_i + \varphi_j(x)U_j + \varphi_k(x)U_k, \quad 0 \leq x \leq l \quad (30)$$

Отсюда определяется градиент перемещения

$$\frac{\partial u}{\partial x} = \frac{4x-3l}{l^2}U_i + \frac{4l-8x}{l^2}U_j + \frac{4x-l}{l^2}U_k, \quad 0 \leq x \leq l \quad (31)$$

где $U_i = U(x=0)$; $U_j = U(x = \frac{l}{2})$; $U_k = U(x=l)$. Функционал потенциальной энергии упругих деформации при наличии поля температуры имеет следующий вид [3].

$$\Pi = \int_V \frac{\sigma_x(x)}{2} \varepsilon_x(x) dv - \int_V \alpha E \cdot T(x) \cdot \varepsilon_x(x) dv \quad (32)$$

где $\varepsilon_x(x) = \frac{\partial U}{\partial x} = \frac{4x-3l}{l^2}U_i + \frac{4l-8x}{l^2}U_j + \frac{4x-l}{l^2}U_k$ – упругое составляющее деформации;

$$\sigma_x(x) = E\varepsilon_x(x) = E \frac{\partial U}{\partial x} = E \left[\frac{4x-3l}{l^2}U_i + \frac{4l-8x}{l^2}U_j + \frac{4x-l}{l^2}U_k \right] \text{ – упругое составляющее}$$

напряжения.

С учетом этих соотношений выражения потенциальной энергии других деформаций для исследуемого стержня можно переписать в следующем виде:

$$\begin{aligned} \Pi &= \int_{V_1} \frac{\sigma_x(x)}{2} \varepsilon_x(x) dv - \int_{V_1} \alpha \cdot E \cdot T(x) \cdot \varepsilon_x(x) dv + \int_{V_2} \frac{\sigma_x(x)}{2} \varepsilon_x(x) dv - \\ &- \int_{V_2} \alpha \cdot E \cdot T(x) \cdot \varepsilon_x(x) dv + \int_{V_3} \frac{\sigma_x(x)}{2} \varepsilon_x(x) dv - \int_{V_3} \alpha \cdot E \cdot T(x) \cdot \varepsilon_x(x) dv = \\ &= \frac{EF}{2} \int_0^l \left[\frac{4x-3l}{l^2}U_1 + \frac{4l-8x}{l^2}U_2 + \frac{4x-l}{l^2}U_3 \right]^2 dx - \alpha EF \int_0^l \left[\frac{2x^2-3lx+l}{l^2}T_1 + \frac{4l-4x^2}{l^2}T_2 + \frac{2x^2-lx}{l^2}T_3 \right] \times \\ &\times \left[\frac{4x-3l}{l^2}U_1 + \frac{4l-8x}{l^2}U_2 + \frac{4x-l}{l^2}U_3 \right] dx + \frac{EF}{2} \int_0^l \left[\frac{4x-3l}{l^2}U_3 + \frac{4l-8x}{l^2}U_4 + \frac{4x-l}{l^2}U_5 \right]^2 dx - \\ &- \alpha EF \int_0^l \left[\frac{2x^2-3lx+l}{l^2}T_3 + \frac{4l-4x^2}{l^2}T_4 + \frac{2x^2-lx}{l^2}T_5 \right] \times \left[\frac{4x-3l}{l^2}U_3 + \frac{4l-8x}{l^2}U_4 + \frac{4x-l}{l^2}U_5 \right] dx + \\ &+ \frac{EF}{2} \int_0^l \left[\frac{4x-3l}{l^2}U_5 + \frac{4l-8x}{l^2}U_6 + \frac{4x-l}{l^2}U_7 \right]^2 dx - \alpha EF \int_0^l \left[\frac{2x^2-3lx+l}{l^2}T_5 + \frac{4l-4x^2}{l^2}T_6 + \frac{2x^2-lx}{l^2}T_7 \right] \times \\ &\times \left[\frac{4x-3l}{l^2}U_5 + \frac{4l-8x}{l^2}U_6 + \frac{4x-l}{l^2}U_7 \right] dx = \end{aligned}$$

$$\begin{aligned}
&= \frac{EF}{6l_1}(7U_1^2 + 16U_2^2 + 7U_3^2 - 16U_1U_2 + 2U_1U_3 - 16U_2U_3) - \alpha EF \left(-\frac{1}{2}T_1U_1 + \frac{2}{3}T_1U_2 - \right. \\
&- \frac{1}{6}T_1U_3 - \frac{2}{3}T_2U_1 + \frac{2}{3}T_2U_3 + \frac{1}{6}T_3U_1 - \frac{2}{3}T_3U_2) + \frac{EF}{6l_2}(7U_3^2 + 16U_4^2 + 7U_5^2 - 16U_3U_4 + \\
&+ 2U_3U_5 - 16U_4U_5) - \alpha EF \left(\frac{2}{3}T_3U_4 - \frac{1}{6}T_3U_5 - \frac{2}{3}T_4U_3 + \frac{2}{3}T_4U_5 + \frac{1}{6}T_5U_3 - \frac{2}{3}T_5U_4) + \\
&+ \frac{EF}{6l_3}(7U_5^2 + 16U_6^2 + 7U_7^2 - 16U_5U_6 + 2U_5U_7 - 16U_6U_7) - \alpha EF \left(\frac{2}{3}T_5U_6 - \frac{1}{6}T_5U_7 - \right. \\
&- \frac{2}{3}T_6U_5 + \frac{2}{3}T_6U_7 + \frac{1}{6}T_7U_5 - \frac{2}{3}T_7U_6 + \frac{1}{2}T_7U_7)
\end{aligned}$$

Из-за заземленности обоих концов стержня $U_1=U_7=0$.

Минимизируя этот функционал по узловым перемещениям U_2, U_3, \dots, U_6 строится разрешающая система линейных алгебраических уравнений относительно искомых величин с учетом одновременного наличия разнородных источников тепла и теплоизоляций

$$\begin{aligned}
1) \frac{\partial \Pi}{\partial U_2} = 0; &\Rightarrow \frac{EF}{6l_1}(32U_2 - 16U_3) - \alpha EF \left(\frac{2}{3}T_1 - \frac{2}{3}T_3\right) = 0; \\
2) \frac{\partial \Pi}{\partial U_3} = 0; &\Rightarrow EF \left[\left(\frac{14U_3 - 16U_2}{16l_1}\right) + \left(\frac{14U_3 - 16U_4 + 2U_5}{16l_2}\right) \right] - \alpha EF \left(-\frac{1}{6}T_1 + \frac{2}{3}T_2 - \frac{2}{3}T_4 + \frac{1}{6}T_5\right) = 0; \\
3) \frac{\partial \Pi}{\partial U_4} = 0; &\Rightarrow \frac{EF}{6l_2}(32U_4 - 16U_3 - 16U_5) - \alpha EF \left(\frac{2}{3}T_3 - \frac{2}{3}T_5\right) = 0; \\
4) \frac{\partial \Pi}{\partial U_5} = 0; &\Rightarrow EF \left[\left(\frac{14U_5 + 2U_3 - 16U_4}{16l_2}\right) + \left(\frac{14U_5 - 16U_6 + 2U_7}{16l_7}\right) \right] - \alpha EF \left(-\frac{1}{6}T_3 + \frac{2}{3}T_4 - \frac{2}{3}T_6 + \frac{1}{6}T_7\right) = 0; \\
5) \frac{\partial \Pi}{\partial U_6} = 0; &\Rightarrow \frac{EF}{6l_3}(32U_6 - 16U_5 - 16U_7) - \alpha EF \left(\frac{2}{3}T_5 - \frac{2}{3}T_7\right) = 0.
\end{aligned}$$

Решая последнюю систему определяются значения U_2, U_3, \dots, U_6 . По ним строится поле перемещения в пределах трех дискретных элементов

$$\left. \begin{aligned}
U^{(I)}(x) &= \frac{4lx - 4x^2}{l^2} \cdot U_2 + \frac{2x^2 - lx}{l^2} \cdot U_3 \\
U^{(II)}(x) &= \frac{2x^2 - 3lx + l^2}{l^2} \cdot U_3 + \frac{4lx - 4x^2}{l^2} \cdot U_4 + \frac{2x^2 - lx}{l^2} \cdot U_5 \\
U^{(III)}(x) &= \frac{2x^2 - 3lx + l^2}{l^2} \cdot U_5 + \frac{4lx - 4x^2}{l^2} \cdot U_6
\end{aligned} \right\} \quad (33)$$

4. Анализ полученных результатов

Для практического применения вышеизложенного метода и алгоритма, примем следующие исходные данные $L=30\text{см}; r=1\text{см}; \alpha=0,0000125 \frac{1}{\text{K}}$;

$$E = 2 \cdot 10^6 \frac{\text{кГ}}{\text{см}^2}; K_{xx} = 100 \frac{\text{watt}}{\text{см} \cdot ^\circ\text{K}}; h_1 = h_2 = 10 \frac{\text{watt}}{\text{см}^2 \cdot ^\circ\text{K}}; T_{oc1} = T_{oc2} = 40^\circ\text{K};$$

$$q = -500 \frac{\text{watt}}{\text{см}^2};$$

При этих исходных данных полученные решения приведены на рисунках 4-7.

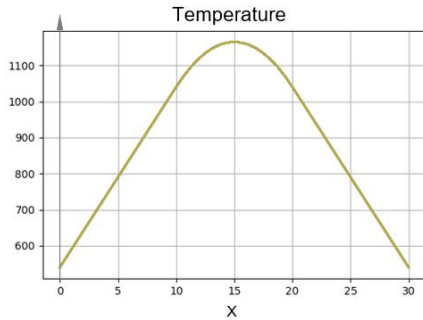


Рис.4. Зависимости температуры T по длине стержня по длине стержня

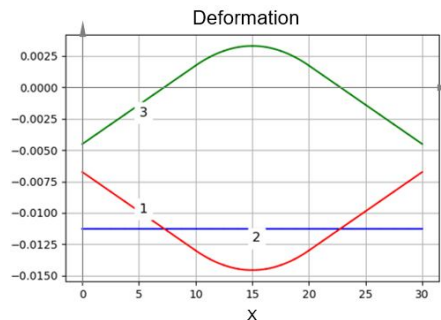


Рис.5. Зависимости деформации по длине стержня по длине стержня

$$1 - \varepsilon(x); 3 - \varepsilon_x(x); 2 - \varepsilon_T(x)$$

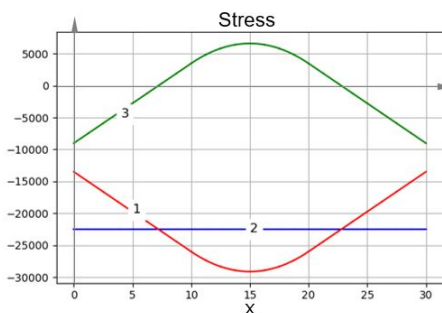


Рис. 6. Зависимости напряжений по длине стержня

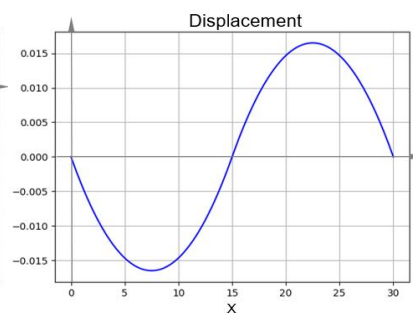


Рис.7. Зависимости перемещений по длине стержня

$$1 - \sigma(x); 3 - \sigma_x(x); 2 - \sigma_T(x)$$

Для численного решения рассматриваемой задачи, исследуемый стержень дискретизировался $n=24$ – дискретными элементами одинаковой длины $l = \frac{L}{n} = \frac{30}{24} = 1,25 \text{ см}$. При этом боковая поверхность первых и последних 8 элементов полностью теплоизолированы. На боковой поверхности серединного 8-элементов подведен тепловой поток $q = -500 \frac{\text{watt}}{\text{см}^2}$.

Из рисунка 4 видно, что из-за симметрии рассматриваемой задачи на концах стержня значения температуры будут $T(x=0)=T(x=L)=540^\circ\text{K}$. В то время наибольшая температура возникает в середине стержня, $T(x=\frac{L}{2})=1165^\circ\text{K}$. Из рисунка 4 также видно, что поле распределения температуры по длине исследуемого стержня имеет параболический характер.

Законы распределения трех составляющих деформаций приводится на рисунке-5. Здесь поле распределения температурной составляющей деформации $\varepsilon_T(x)$, имеет сжимающий характер по всей длине исследуемого стержня. Она имеет параболический вид выпуклостью вниз. Его значение на концах исследуемого стержня будет $\varepsilon_T(x=0) = \varepsilon_T(x=L) = -0.00675$. В то время в середине стержня она имеет $\varepsilon_T(x = \frac{L}{2}) = -0.014562$. Отсюда видно, что в середине стержня значения температурной составляющей деформации будет 2,157 раза больше чем в конца стержня. В то время значение термоупругой составляющей деформаций $\varepsilon(x)$ будет по всей длине стержня постоянной и $\varepsilon(x) = -0,0112$. Поведение упругой составляющей деформации будет переменного знака. Например, серединной части стержня, куда подведена тепловой поток упругий составляющий деформации $\varepsilon_x(x)$ имеет растягивающий характер. Амплитуда $\varepsilon_x(x = \frac{L}{2}) = -0.033$. В то время первая и последняя 1/3 части стержня будет испытывать сжимающий $\varepsilon_x(x)$. Следует отметить, что на концах стержня $\varepsilon_x(x=0) = \varepsilon_x(x=L) = -0.0045$. В общем поле распределения $\varepsilon_x(x)$ имеет параболический характер, выпуклостью в верх. В соответствующих обобщенного закона Гука, поведения составляющих напряжения будут подобающими соответствующих составляющих деформаций (рисунок 6).

Значения термоупругой составляющей напряжения $\sigma(x)$ по всей длине стержня будет постоянной $\sigma(x) = -22528 \frac{\kappa\Gamma}{\text{см}^2}$. Но закон распределения температурной составляющей напряжения $\sigma_T(x)$ будет иметь параболический характер, выпуклостью вниз, и она имеет сжимающий характер по всей длине стержня. При этом значения $\sigma_T(x)$ на концах стержня будет $\sigma_T(x=0) = \sigma_T(x=L) = -13500 \frac{\kappa\Gamma}{\text{см}^2}$. В то время в середине стержня $\sigma_T(x = \frac{L}{2}) = -29125 \frac{\kappa\Gamma}{\text{см}^2}$. Это показывает, что в середине стержня значения σ_T будет 2,157 раза больше чем в ее концах. Упругий составляющий напряжения $\sigma_x(x)$ в первой и последней 1/3 части стержня ведет себя сжимающим, а в серединной 1/3 части растягивающей. При этом $\sigma_x(x=0) = \sigma_x(x=L) = -9027 \frac{\kappa\Gamma}{\text{см}^2}$, $\sigma_x(x = \frac{L}{2}) = 6597 \frac{\kappa\Gamma}{\text{см}^2}$. Закон распределения перемещения сечений стержня приводится на рисунке-7. Из этого рисунка видно, что все сечений левой половины стержня перемещаются против направления оси ox , а сечения правой половины перемещаются в направлении оси ox . Этот процесс обусловлена наличием на поверхности серединной 1/3 части исследуемого стержня подведенного теплового потока. В то время амплитуда перемещения сечения координата которого $x=7,5$ см будет равна $U(x=7.5) = -0.01627$ см. Аналогично имеем, что $U(x=22.5) = 0.01627$ см. это показывает, что исследуемый процесс является строго симметричной относительно середины стержня.

5. Заключение

Разработанный метод на основе фундаментальных законов сохранения энергии, вычислительный алгоритм и прикладная программа на Python позволяют автоматизировать построение разрешающих систем уравнений с учетом естественных граничных условий, для стержней ограниченной длины находящихся под воздействием разнородных видов источников тепла. Разработанная система также позволяет определить закон распределения температуры, всех составляющих деформаций, напряжений и перемещения. В связи с этим можно констатировать, что разработанный метод, алгоритм и прикладная программа Python являются относительно универсальными в исследовании установившегося термо-напряженно-деформированного состояния несущих стержневых элементов стратегических сооружений, находящихся под воздействием разнообразных локальных видов источников тепла. При этом полученные численные результаты будут отличаться высокой точностью, так как они удовлетворяют фундаментальным законам сохранения энергии.

Литература

1. Timoshenko S., Goodyear J.N. Theory of Elasticity. – McGRAW-Hill. Book. Company. Inc., 1987. – 567 p.
2. Khalid S. SHIBIB, Mohammed A. MINSHID, and Nebras E. Alattar: Thermal And Stress Analysis In Nd:Yag Laser Rod With Different Double End Pumping Methods. Thermal Science, Year 2011, Vol. 15, Suppl. 2, pp. S399-S407 [3] Daryl L. Logan. A First Course in the Finite Element Method, CENGAGE Learning 2012. P. 727-764
3. Yusuf Orçan. Thermal stresses in a heat generating elastic-plastic cylinder with free ends. International Journal of Engineering Science. Volume 32, Issue 6, June 1994, Pages 883-898
4. V.I. Andreev & R.A. Turusov. Nonlinear modeling of the kinetics of thermal stresses in polymer rods. Advanced Materials and Structural Engineering – Hu (Ed.) 2016 Taylor & Francis Group, London
5. Y.I. Nyashin, V.Y. Kiryukhin, F. Ziegler CONTROL OF THERMAL STRESS AND STRAIN. Journal of Thermal Stresses. Volume 23, Issue 4, 30 Nov 2000, Pages 309-326
6. K.K. Gornostaev, A.V. Kovalev, Y.V. Malygina, Stress-strain state in an elastoplastic pipe taking into account the temperature and compressibility of the material. Journal of Physics: Conf. Series 973 (2018)
7. L. Segerlind, *Applied Finite Element Analysis* (John Wiley & Sons, Inc., New York, Chichester, 1985).
8. Dr. Ibrahim Assakkaf. *RODS: THERMAL STRESS AND STRESS CONCENTRATION*. Department of Civil & Environmental Engineering
9. B.F. Shorr, Thermal Integrity in Mechanics and Engineering, Foundations of Thermoelasticity, Pages 33-55, Springer-Verlag Berlin Heidelberg 2015

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ В ПРОЦЕССЕ ТРАНСГРАНИЧНОГО ОБМЕНА

Бияшев Р.Г., Бегимбаева Е.Е., Рог О.А.
e-mail: brg@ipic.kz, enlik_89@mail.ru, olga@ipic.kz

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Аннотация. *Предлагается модель автоматизированной системы защищенного трансграничного информационного обмена. Исследованы возможные конфликтные ситуации и причины их возникновения. Предложен алгоритм действий по разрешению возможных конфликтных ситуаций. Осуществлен аналитический обзор основных моделей и методов разграничения доступа. Рассмотрены различные виды моделей АВАС и разрабатываемая модель типизированного атрибутного разграничения доступа (ТАРД)*

Введение

В современных условиях увеличение количества информации при международном электронном обмене привело к повышению актуальности задач создания автоматизированной системы защищенного электронного информационного обмена. Защита информации в автоматизированных системах (АС) является особо важной задачей, в рамках которой необходимо обеспечить ее достоверность, сохранность и конфиденциальность [1]

Государственный Стандарт Республики Казахстан «Информационная технология. Комплекс стандартов на автоматизированные системы. Автоматизированные системы термины и определения СТ РК 34.014-2002» [2] определяет автоматизированную систему как систему состоящую из персонала и комплекса средств автоматизации его деятельности, реализующая информационную технологию выполнения установленных функций.

Среди понятий теории защиты информации одними из базовых являются термины «безопасность информации» и «защищенные АС». Информация в АС находится в безопасности, если все компоненты этой системы находятся в таком состоянии, при котором информация защищена от возможных угроз на требуемом уровне. Автоматизированные системы, обеспечивающие безопасность, информации, называются защищенными [3].

Для обеспечения информационной безопасности в автоматизированной системе создается комплекс подсистем защиты информации: шифрования и электронной цифровой подписи (ЭЦП); разграничения доступа к хранимой информации в созданных базах данных; разрешения возможных конфликтных ситуаций.

Один из вариантов разработанной модели автоматизированной системы защищенного трансграничного информационного обмена представлен на рисунке 1 [4].

Разработка системы разрешения возможных конфликтных ситуаций в автоматизированной системе

Независимо от формы представления электронной информации, в автоматизированной системе могут возникать различные конфликтные ситуации. Обеспечение неконфликтного взаимодействия является важной задачей для обеспечения комплексной безопасности системы. В число таких задач входят исследование информационных объектов АС и анализ уровня защищенности передаваемой информации в случайных и преднамеренных конфликтах. В связи с этим возникает необходимость определения и исследования возможных видов конфликтных ситуаций и способов их разрешения

Под конфликтными ситуациями понимается такое состояние АС ЗТИО при котором исключается возможность корректного выполнения хотя бы одной задачи системы (контроль доступа, передача информации, шифрование, формирование и проверка подписи) в силу внешнего воздействия, либо внутренних сбоев, ошибок или отказов программного, или алгоритмического обеспечения.

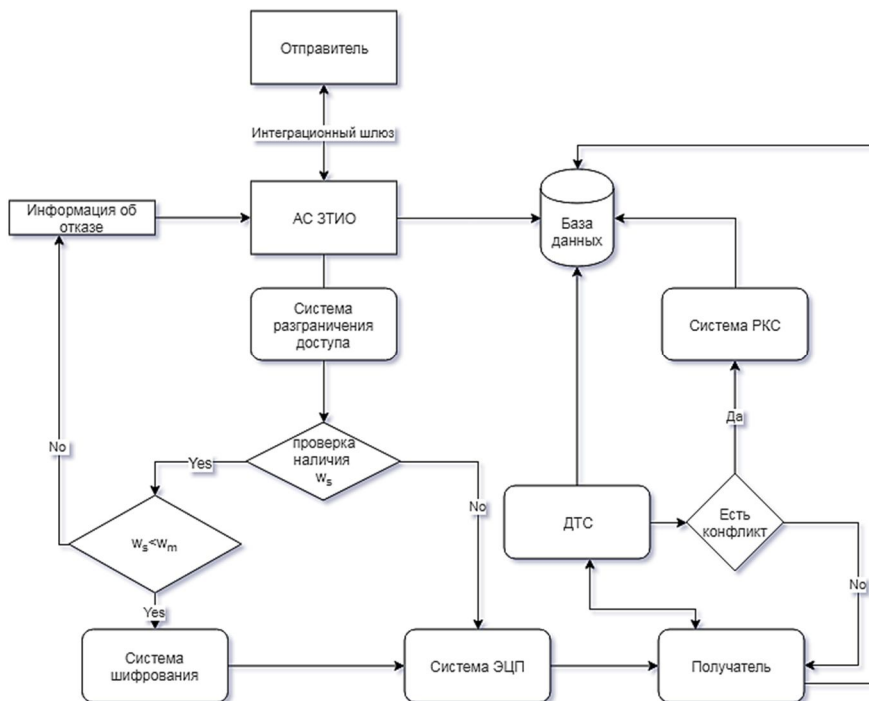


Рисунок 1 – Автоматизированная система ЗТИО

Условно возникновение конфликтных ситуаций можно разделить на три категории вследствие:

- искажения входного (полученного) потока данных (документов);
- изменения функционирования системы посредством внешнего вмешательства (несанкционированный доступ, умышленное изменение программного или алгоритмического обеспечения системы);
- искажения функций системы при изменении входного потока данных (вирусы, троянские кони, ошибки и закладки в программном обеспечении).

Для разрешения конфликтных ситуаций (РКС) при информационном взаимодействии субъектов двух сторон используется информация, поступающая непосредственно от интеграционного шлюза к субъектам одной стороны и обратно в открытом виде.

На рисунке 2 показана структурная схема системы разрешения конфликтных ситуаций (РКС). В состав РКС входят [5-7]:

– программный агент, который обеспечивает сбор необходимых данных для последующего анализа. Источником данных будет являться «журнал событий и/или журнал учета отправки/получения документов», который фиксирует все процессы автоматизированной системы;

– механизм РКС, который обрабатывает информацию из базы данных и управляет программным агентом. Результатом работы модуля является решение о внесении изменений в настройку системы защищенного ТИО для обеспечения неконфликтного взаимодействия, основанного на правилах, которые хранятся в базе данных.

– СУБД, которые используются для хранения информации о конфликтах и правил разрешения конфликтных ситуаций.



Рисунок 2 – Структурная схема системы РКС

Поскольку АС состоит из нескольких связанных подсистем, то в каждой подсистеме могут возникать конфликтные ситуации. Ниже приведены возможные конфликтные ситуации в АС ЗТИО.

Спор между субъектами информационного обмена в отношении:

1) факта отправления и/или получения электронного документа;

- отказ Стороны от электронного документа (Сторона утверждает, что ее абонент не подписывал принятый другой Стороной электронный документ, а другая Сторона утверждает обратное);

- отказ Стороны от факта получения электронного документа (Сторона утверждает, что посланный ею электронный документ был принят другой Стороной, а другая Сторона это отрицает).

- у Стороны нет факта получения электронного документа (Сторона утверждает, что посланный ею электронный документ был отправлен, но не был принят другой Стороной).

2) времени отправления и/или получения электронного документа;

3) содержания отправленного/полученного электронного документа;

4) целостности электронного документа.

Конфликт связанный с:

1) поступлением документа с ЭЦП с отрицательным результатом проверки;

2) поступлением документа с ЭЦП с несоответствующим уровнем допуска.

За период создания системы защищенного трансграничного обмена в области разработки системы разрешения возможных конфликтных ситуаций были проведены следующие работы:

- предложен один из вариантов алгоритма действий разрешения конфликтных ситуаций;

- продолжаются работы по классификации возможных конфликтных ситуаций;
 - сформулированы требования к системе разрешения конфликтных ситуаций.
- Система РКС должна реализовывать следующие функции:
- распознавать происхождение конфликтов, вырабатывать корректирующее воздействие путем внесения изменения в конфигурацию подсистем;
 - выдавать рекомендации пользователю по устранению конфликтных ситуаций;
 - вести аудит системных событий с целью исключения элементов, вызывающих конфликт. Полученные результаты работ будут использованы при реализации модели автоматизированной системы ЗТИО и будут способствовать развитию отечественной системы информационной безопасности при таком обмене.

Разработка моделей и средств разграничения доступа к информации в процессе трансграничного информационного обмена.

За период создания системы защищенного трансграничного обмена в области разработки разграничения доступа к информации были проведены следующие работы.

Осуществлен аналитический обзор основных моделей и методов разграничения доступа, начиная от традиционных (DAC, MAC, RBAC), до последних разработок—многочисленных моделей, реализующих атрибутное разграничение доступа (ABAC) [8-11].

Отмечены следующие недостатки традиционных моделей:

- идентификация сущностей уникальными именами;
- избыточность прав доступа («coarse-grained access control»);
- трудность управления большим числом пользователей;
- работа в замкнутых средах;
- невозможность использования интегрированных политик безопасности;
- отсутствие встроенных средств администрирования.

Сформулированы требования к вновь создаваемым методам и моделям разграничения доступа. Было установлено, что для обеспечения безопасного совместного использования информационных ресурсов как в глобальных так и локальных вычислительных средах модели разграничения доступа должны удовлетворять требованиям универсальности, гибкости и удобства администрирования, выполняя при этом задачи:

- идентификации сущностей наборами признаков для точного разграничения доступа (fine-grained access control);
- конструирования и использования нескольких политик разграничения доступа для выполнения парадигмы «множественной политики» и адаптации системы к работе в различных средах;
- администрирования как средства для динамического моделирования политик и удобного управления привилегиями большого числа пользователей.

Рассмотрены различные виды моделей ABAC. Описаны их достоинства и недостатки [12].

Преимуществами ABAC моделей являются:

- идентификация сущностей наборами атрибутов;
- точность разграничения доступа («fine-grained access control»);
- гибкость и выразительность языков спецификации моделей. Возможность создания новых и моделирования традиционных методов разграничения доступа;

- относительная простота администрирования. Управление привилегиями групп пользователей.

Основным недостатком АВАС является сложность вычисления значений атрибутов.

Авторами ведутся работы по созданию новых моделей разграничения доступа, обеспечивающих надежное совместное использование информационных ресурсов и отвечающих требованиям универсальности и гибкости управления безопасностью.

Рассмотрена разрабатываемая в настоящее время модель типизированного атрибутного разграничения доступа (ТАРД) [12-17].

Показано, что модели ТАРД отвечают поставленным требованиям и обеспечивают:

- точное разграничение доступа («fine-grained access control») путем идентификации сущностей множествами типизированных атрибутов;
- снижение сложности и увеличение скорости вычислений;
- управление привилегиями иерархических групп субъектов и объектов;
- динамическое конструирование политик;
- многокритериальное разграничение доступа.

Данные модели призваны решать задачи:

- точной идентификации сущностей,
- обеспечения возможностей динамического конструирования политик и
- использования множественных политик в рамках одной системы.

Создано описание и формальное представление разрабатываемого метода и модели типизированного атрибутного разграничения доступа (ТАРД).

Модель ТАРД принадлежит классу моделей АВАС, но, в отличие от АВАС, атрибутам безопасности сущностей ТАРД приписываются определенные типы. Решение о возможности доступа принимается на основе обработки однотипных атрибутов пары субъект-объект. Модель ТАРД, определяется на нескольких уровнях с помощью языков логической и алгебраической спецификации.

Описываемая модель может быть отнесена к разряду моделей разграничения доступа общего назначения, основанной на логических формулах. В ее состав входят:

- средства определения возможности доступа субъектов к объектам в соответствии с их полномочиями;
- средства двухступенчатого администрирования – для конструирования политик авторизации и для управления идентификацией сущностей в процессе разграничения доступа.

В соответствии с определением типа Т, модель ТАРД имеет многоуровневое представление, которое позволяет конструировать типы, представляющие собой различные политики безопасности, включая традиционные DAC, MAC, RBAC, реализуя таким образом парадигму «множественной политики».

Идентификация сущностей ТАРД осуществляется путем типизации. В результате сущности приобретают метки безопасности, содержащие структурированные множества типизированных атрибутов, которые определяют порядок доступа субъектов к объектам.

Предоставляется возможность множественной типизации сущностей, каждая из которых может иметь ряд атрибутов различных типов.

Тип Т определяется как математический объект, являющийся механизмом разграничения доступа и предназначенный для реализации определенной политики. Он содержит домен, имеющий вид конечного полного частично

упорядоченного множества всевозможных значений атрибутов, структурированного отношением предшествования, с определенными на нем операциями типизации, представленными непрерывными монотонными функциями, а также операцией (предикатом) доступа.

Модель ТАРД обладает следующими характеристиками:

- принцип обработки атрибутов одинаковых типов является предпосылкой обеспечения скорости вычисления их значений;
- имеет возможность формального доказательства правильности решений о предоставлении доступа путем использования дедуктивного аппарата логической спецификации модели;
- обеспечивает наглядность и контроль процесса администрирования;
- способна динамически конструировать новые модели разграничения доступа вместе с возможностью моделировать традиционные DAC, MAC, RBAC;
- модель непосредственно реализуема на языках функционального и логического программирования с использованием аппарата программирования в ограничениях.

Перечисленные особенности позволяют использовать системы типизированного атрибутного разграничения доступа в качестве средств защиты информации в казахстанском сегменте трансграничного обмена.

Литература

1. Иванов К.К., Юрченко Р.Н., Ярмонов А.С. Защита информации в автоматизированных системах // Молодой ученый. — 2016. — №29. — С. 22-24.
2. Государственный стандарт Республики Казахстан Информационная технология. Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Термины и определения. СТ РК 34.014-2002 // https://online.zakon.kz/Document/?doc_id=31056334 (дата обращения: 11.03.2019).
3. Зегжда Д.П. и др. Теория и практика обеспечения информационной безопасности. — М.: Яхтсмен, 1996
4. R.G. Biyashev, S.E. Nyssanbayeva, and Ye.Y. Begimbayeva The Development of a Structural Scheme of National Segment in a Protected Cross-Border Space // Proceedings of the International Conference on Wireless Communications, Network Security and Signal Processing International Conference on Advanced Material Science and Environmental Engineering. – 2016. – P. 250-252.
5. Бегимбаева Е.Е. Структурная схема и механизм разрешения конфликтов в трансграничном информационном обмене // III Междунар. науч. конф. «Информатика и прикладная математика», посв. 80-летию юбилею проф. Бияшева Р.Г. и 70-летию проф. Айдарханова М.Б. - Алматы, 2018. – С.176-178.
6. Ding, Y., Guo, X., Su, H., Wang, Z., Zhao, S. Method and system for avoidance of software conflict. US Patent 20070180441 A1, December 22, 2006.
7. М.А. Поляничко. Архитектура системы автоматизированного обнаружения и разрешения конфликтов программных средств защиты информации // ISSN 1815-588X. Известия ПГУПС, №1. – 2013. – 39-45.
8. Sandhu R.S., Samarati P. Access control: principle and practice. Communications Magazine, IEEE, 32(9):40–48, 1994.
9. Гайдамакин Н.А. Разграничение доступа к информации в компьютерных системах / – Екатеринбург: Изд-во Урал. ун-та, 2003. – 328 с.
10. Hosmer H. 1993. The multipolicy paradigm for trusted systems. In Proceedings on the 1992-1993 workshop on New security paradigms (NSPW '92-93), J. Bret Michael,

Victoria Ashby, and Catherine Meadows (Eds.). ACM, New York, NY, USA, 19-32. DOI=<http://dx.doi.org/10.1145/283751.283768>

11. Калимолдаев М.Н., Бияшев Р.Г., Рог О.А. Анализ атрибутивных методов разграничения доступа // Принята к печати в журнал ПДМ (Прикладная дискретная математика), г. Томск, РФ. 11.12.2018.

12. Karp A., Hauray H., Davis M. (2010). From ABAC to ZBAC: The evolution of access control models. ISSA (Information Systems Security Association) Journal. 8. 22-30.

13. Бияшев Р.Г., Калимолдаев М.Н., Рог О.А. Моделирование семантики типизированного атрибутивного разграничения доступа // Журнал Проблемы информатики, - 2017, № 1. С. 25-37.

14. Калимолдаев М.Н., Бияшев Р.Г., Рог О.А. Применение логики для построения моделей разграничения доступа к информации // «Доклады Национальной Академии Наук Республики Казахстан» 2017, №3. С. 48-54.

15. Калимолдаев М.Н., Бияшев Р.Г., Рог О.А. Основы архитектуры программных систем для осуществления типизированного атрибутивного разграничения доступа // Современные проблемы информатики и вычислительных технологий: Мат. науч. конф. (29-30 июня 2017 г). – Алматы, 2017, – С. 88-95.

16. Калимолдаев М.Н., Бияшев Р.Г., Рог О.А. О применении типизированного атрибутивного разграничения доступа в глобальных вычислительных средах // Известия Научно-Технического Общества «КАХАК» Алматы, 2017 г., № 3 (58), – С. 30-36.

17. Калимолдаев М.Н., Бияшев Р.Г., Рог О.А. Многоуровневое представление модели типизированного атрибутивного разграничения доступа // Современные проблемы информатики и вычислительных технологий: (Мат. Науч. Конф. 2-5 июля 2018 г.) – Алматы: ИИВТ МОН РК. 2018. – С. 111-120.

КРАТКИЙ ОБЗОР ОСНОВНЫХ ПРИНЦИПОВ РАЗРАБОТКИ СИСТЕМ УПРАВЛЕНИЯ КРИПТОГРАФИЧЕСКИМИ КЛЮЧАМИ

Варенников А.В.

e-mail: avarennikov@gmail.com

*Институт информационных и вычислительных технологий КН МОН РК,
Республика Казахстан*

Аннотация. В данной статье описываются требования и вопросы, которые необходимо рассматривать и учитывать при разработке систем управления криптографическими ключами.

Введение

Современные информационные системы и содержащаяся в них информация считаются критически важными активами, требующими защиты. Эта информация часто содержится в компьютерных системах, использующих общие сети, в том числе Интернет. Поскольку Интернет совместно используется различными, часто конкурирующими между собой организациями, информационные системы должны

защищать себя и содержащуюся в них информацию от несанкционированного раскрытия, изменения и использования. Кроме того, отказ в обслуживании законных пользователей также рассматривается как серьезная угроза безопасности. Информация, используемая этими системами, требует защиты, когда она находится на хранении, в процессе обработки, а также когда она передаётся от одного объекта другому.

Для защиты информации от несанкционированного раскрытия, обнаружения несанкционированной модификации и подтверждения подлинности источника данных чаще всего используется криптография. Криптографические методы используют криптографические ключи, которыми управляет и которые защищает система управления криптографическими ключами СКМС (Cryptographic Key Management System).

1 Криптографические алгоритмы

Существует три основных класса криптографических алгоритмов: хеш-функции, алгоритмы на основе симметричных ключей и алгоритмы на основе асимметричных ключей. Эти классы определяются количеством криптографических ключей, используемых конкретным алгоритмом.

Криптографические хеш-функции для выполнения своих операций не требуют ключей. Хеш-функции генерируют из исходных данных произвольной длины сравнительно небольшой дайджест (хеш-значение) фиксированной длины способом, который является принципиально односторонним (то есть найти исходные данные, в результате обработки которых будет получено заданное хеш-значение, практически невозможно). В механизмах управления ключами хеш-функции используются в качестве строительных блоков, например, для обеспечения аутентификации источника и целостности данных с использованием кодов аутентификации сообщений, для сжатия сообщений при генерации и проверке цифровой подписи, для генерации псевдослучайных чисел, а также в процессе распределения ключей.

Алгоритмы на основе симметричных ключей преобразуют данные таким образом, что их практически невозможно восстановить без знания секретного ключа. Ключ называется «симметричным», потому что один и тот же ключ используется как для криптографической операции, так и для её инверсии (например, зашифрования и расшифрования). Симметричные ключи часто бывают известны нескольким лицам, однако ключ не должен раскрываться объектам, у которых нет прав доступа к данным, защищенным этим алгоритмом и ключом. Алгоритмы на основе симметричных ключей используются для обеспечения конфиденциальности данных, обеспечения аутентификации источника и целостности данных, как часть процесса распределения криптографических ключей, а также для генерации псевдослучайных чисел.

Алгоритмы на основе асимметричных ключей (обычно известные как алгоритмы с открытым ключом) используют для выполнения своих функций два взаимосвязанных ключа (то есть пару ключей): открытый (публичный) ключ и закрытый (личный) ключ. Открытый ключ может быть известен любому; закрытый ключ должен находиться под исключительным контролем объекта, который «владеет» этой парой ключей. Несмотря на то, что открытый и закрытый ключи одной и той же пары ключей связаны между собой, знание открытого ключа невозможно использовать для определения закрытого ключа. Алгоритмы на основе асимметричных ключей используются для формирования и проверки цифровых подписей, а также для распределения ключевого материала.

2 Криптографические ключи и метаданные

Криптографический ключ – это параметр, используемый в сочетании с криптографическим алгоритмом, операции которого определены таким образом, что объект со знанием ключа может воспроизвести эту операцию, выполнить обратную операцию или проверить результат выполнения этой операции, тогда как объект без знания ключа сделать этого не может [1].

Все ключи должны быть защищены от несанкционированной замены, модификации, а также несанкционированного использования. Управление ключами обеспечивает основу для безопасной генерации, хранения, распределения и последующего уничтожения ключей. При разработке СКМС криптографические методы, используемые для защиты ключей, должны обеспечивать уровень защиты, который делает практически невозможными любые попытки её обойти со стороны потенциального злоумышленника. Единственный способ расшифровать ранее зашифрованные данные (без знания правильного ключа) – это проверить все возможные ключи, пока, в конце концов, не будет использован правильный ключ. Таким образом, одним из факторов, влияющих на стойкость криптографического алгоритма, является общее количество возможных ключей.

По своим свойствам криптографические ключи классифицируются как симметричные (секретные) и асимметричные. Асимметричные ключи образуют пары, состоящие из открытого (несекретного) и закрытого (секретного) ключа. Эти два ключа обладают свойством, которое заключается в том, что, зная открытый ключ, определить закрытый ключ, с вычислительной точки зрения, практически невозможно. Кроме того, ключи могут обладать статическими (долгосрочными) или эфемерными (краткосрочными) свойствами. Эфемерные ключи используются только для одного сеанса связи или одной транзакции, после чего немедленно уничтожаются. Способы применения ключей включают в себя создание и проверку цифровой подписи, аутентификацию (подтверждение подлинности), шифрование данных, шифрование ключей, генерацию случайных чисел, формирование производных ключей, распределение ключей (передача/согласование) и авторизация (предоставление разрешения на выполнение определенной функции).

Каждый ключ может быть связан с одним или несколькими элементами данных, называемых метаданными, которые определяют тип и размер ключа, каким образом он был сгенерирован, когда он был сгенерирован, идентификатор его владельца, алгоритм, для которого он предназначен, и период его действия. Во многих системах существует необходимость отличать один ключ от другого с использованием одного или нескольких явно записанных элементов метаданных. Как и ключи, метаданные должны быть защищены от несанкционированного изменения и, при необходимости, могут быть защищены от раскрытия.

Каждый ключ, начиная с его создания и заканчивая уничтожением, проходит через ряд состояний. От текущего состояния ключа зависит способ его использования. Переходы между состояниями часто требуют создания записей об этих событиях. Подходящим местом для таких записей является журнал аудита. Ниже приведены возможные состояний ключей [1]:

1) Состояние ключа, ожидающего ввода в действие. Данное состояние означает, что ключ уже создан, но его применение ещё не разрешено. Например, ключ может находиться в этом состоянии, если ещё не наступила дата начала его криптопериода (промежутка времени, в течение которого определённый ключ является действующим).

2) Состояние действующего ключа. Такой ключ может использоваться для криптографической защиты информации (например, для зашифрования открытого

текста или формирования цифровой подписи), для криптографической обработки ранее защищенной информации (например, для расшифрования зашифрованного текста или проверки цифровой подписи) или для того и другого.

3) Состояние ключа, приостановившего своё действие. Одной из причин приостановки может быть возможная компрометация ключа. Другая причина может заключаться в том, что объект, которому принадлежит данный ключ, в настоящее время недоступен.

4) Состояние скомпрометированного ключа. Как правило, ключи становятся скомпрометированными, когда их значение стало известным неавторизованному объекту. Скомпрометированный ключ не должен использоваться для применения криптографической защиты информации, однако в некоторых случаях разрешается обработка информации, криптографически защищенной скомпрометированным ключом (например, для проверки цифровой подписи при условии, что эта подпись была находилась под физической защитой всё время, предшествующее компрометации).

5) Состояние ключа, прекратившего свое действие. Такие ключи не должны использоваться для применения криптографической защиты, но в некоторых случаях могут использоваться для обработки криптографически защищенной информации при условии, что этот ключ не был деактивирован по причинам, связанным с компрометацией.

6) Состояние уничтоженного ключа. Несмотря на то, что ключи в этом состоянии больше не существуют, некоторые метаданные этих ключей (например, история перехода из одного состояния в другое, идентификатор ключа, его тип и криптопериод) могут быть сохранены для целей аудита.

3 Этапы управления криптографическими ключами

Процесс управления криптографическими ключами можно разделить на отдельные этапы. На каждом этапе ключи находятся в определенных состояниях, и на каждом этапе выполняются определенные функции управления ключами.

Существует четыре этапа управления ключами [1]:

1) Подготовительный этап. Ключевой материал (ключи и метаданные) пока ещё недоступен для обычных криптографических операций. Возможно, ключи ещё не были сгенерированы или находятся в состоянии ожидания ввода в действие.

2) Операционный этап. Ключевой материал доступен и используется в обычном режиме. Ключи находятся в действующем или приостановленном состоянии. Ключи в действующем состоянии могут использоваться либо только для защиты, либо только для обработки защищённых данных, либо для того и другого. Ключи в приостановленном состоянии могут быть использованы только для обработки.

3) Завершающий этап. Ключевой материал больше не используется в обычном режиме, но доступ к этому ключевому материалу по-прежнему возможен, и он может использоваться для обработки защищённой информации. Ключи находятся либо в состоянии, соответствующем ключам, прекратившим своё действие, либо в скомпрометированном состоянии. Ключи, которые на данном этапе не используются для обработки защищённых данных, могут быть помещены в архив.

4) Этап уничтожения ключей. Ключи больше не будут доступны. Записи об их существовании могут быть удалены или сохранены. Хотя сами ключи уничтожены, метаданные этих ключей (например, идентификатор ключа, его тип, криптопериод и период использования) могут быть сохранены.

4 Функции управления ключами и метаданными

Функции управления ключами и метаданными выполняются СКМС для ключей или метаданных в целях их рационального использования. СКМС должна обеспечивать создание, изменение, замену и уничтожение ключей и их метаданных. В зависимости от конкретной функции, к входным и/или выходным данным могут применяться средства, обеспечивающие их целостность, аутентификацию источника и/или конфиденциальность.

Ниже приведён стандартный список функций СКМС [2]. Система управления криптографическими ключами не обязательно должна реализовывать все перечисленные функции, поскольку некоторые функции могут быть неприменимы.

1) Генерация криптографических ключей.
2) Регистрация владельцев ключей (физических лиц, организаций, устройств или процессов).

3) Активация ключей. Эта функция обеспечивает перевод криптографического ключа из состояния ожидания ввода в действие в действующее состояние.

4) Деактивация ключей. Эта функция прекращает действие ключа путём перевода этого ключа в соответствующее состояние.

5) Аннулирование ключей. Эта функция используется в тех случаях, когда использование ключа должно быть прекращено до окончания установленного криптопериода этого ключа.

6) Приостановка действия и повторная активация ключей.

7) Продление срока действия открытых ключей.

8) Формирование ключей на основе разделяемого секрета. Данная функция часто используется в протоколах согласования ключей с целью получения общего ключа.

9) Уничтожение ключей и их метаданных.

10) Связывание ключей с метаданными. В зависимости от характера информации, хранящейся в элементе метаданных, элемент метаданных может потребовать защиты конфиденциальности, защиты целостности и аутентификации источника.

11) Изменение метаданных, связанных с ключом.

12) Удаление метаданных, связанных с ключом. Элементы метаданных могут быть удалены все сразу, как отдельные элементы или как заданное подмножество элементов.

13) Получение списка метаданных ключа. Эта функция позволяет объекту получить список элементов метаданных заданного ключа при условии, что данный объект имеет на это право.

14) Хранение оперативных ключей и их метаданных.

15) Резервное копирование ключей и метаданных. Резервные копии ключей и метаданных могут быть расположены в том же месте, что и оперативные ключи/метаданные, или в ином месте для того, чтобы гарантировать, что эти ключи и метаданные при необходимости смогут быть восстановлены.

16) Архивирование ключей и их метаданных – включает в себя помещение ключей и метаданных в безопасное долговременное хранилище с тем, чтобы при необходимости их можно было восстановить.

17) Восстановление ключей и их метаданных – включает в себя получение копии ключей и их метаданных, которые были предварительно сохранены в оперативном хранилище, резервной копии или архиве.

18) Безопасное распределение ключей между двумя или более объектами. Ключ может быть передан от одного объекта другому (передача ключей), или же ключ может быть получен из информации, предоставляемой каждым из этих объектов (согласование ключей).

19) Загрузка ключей и их метаданных в криптографический модуль (набор аппаратного, программного или программно-аппаратного обеспечения, реализующий необходимые функции обеспечения информационной безопасности).

20) Выгрузка ключей и их метаданных из криптографического модуля.

21) Проверка допустимости параметров, используемых некоторыми криптографическими алгоритмами.

22) Проверка допустимости открытых ключей для того, чтобы убедиться в его правильности с арифметической точки зрения. Эти проверки обычно зависят от алгоритма, для которого предназначен данный открытый ключ, но не зависят от знания соответствующего закрытого ключа.

23) Проверка пути сертификации открытых ключей (также известного как цепочка сертификатов).

24) Проверка допустимости симметричных ключей и их метаданных. Например, эти проверки могут включать проверку целостности, правильности длины и формата ключа.

25) Проверка допустимости закрытых ключей для того, чтобы убедиться в его соответствии заданным спецификациям. Эта проверка может быть выполнена только владельцем закрытого ключа или доверенной третьей стороной, действующей от имени владельца закрытого ключа. Данная проверка может также включать в себя проверку ключевой пары на согласованность, которая должна установить, что функция, применяемая к закрытому ключу, является дополняющей по отношению к функции, применяемой к открытому ключу.

26) Проверка обладания закрытым ключом. Эта функция используется объектом, который получает открытый ключ и хочет убедиться в том, что заявленный владелец открытого ключа действительно обладает соответствующим закрытым ключом и, следовательно, является владельцем ключевой пары.

27) Выполнение криптографических функций с использованием ключа. Эти функции могут включать в себя формирование цифровой подписи, проверку цифровой подписи, зашифрование данных, расшифрование данных, зашифрование ключей, расшифрование ключей, генерацию и проверку кодов аутентификации сообщений, используемых для обнаружения как случайных, так и преднамеренных изменений данных.

28) Управление сертификатами доверенных открытых ключей, используемых для установления доверия к другим открытым ключам. Доверие к таким открытым ключам устанавливается путём проверки всех подписей в цепочке сертификатов открытых ключей, начиная с ключа, которому доверяет проверяющий объект.

Заключение

Управление криптографическими ключами является сложной задачей, затрагивающей вопросы криптографии, протоколов и управления. В настоящее время лабораторией информационной безопасности Института информационных и вычислительных технологий КН МОН РК выполняются научно-исследовательские работы по созданию системы управления криптографическими ключами с использованием криптографических алгоритмов собственной разработки. На

данный момент разработана функциональная структура системы управления криптографическими ключами, логическая структура хранилища криптографических ключей, а также схема криптографической защиты сообщений с использованием шифрования и электронной цифровой подписи на основе протокола OpenPGP [3].

Работы ведутся за счет средств грантового финансирования научных исследований на 2018-2020 годы по проекту AP05132568 «Разработка системы управления криптографическими ключами».

Список литературы

1. Barker E. Recommendation for Key Management – Part 1: General // NIST Special Publication 800-57, Revision 4. – 2016. – 160 p.
2. Barker E., Smid M., Branstad D., Chokhani S. A Framework for Designing Cryptographic Key Management Systems // NIST Special Publication 800-130, Revision 4. – 2013. – 120 p.
3. Callas, J., Donnerhackle L., Finney H., Shaw D., Thayer R., OpenPGP Message Format // Request for Comments: 4880. – 2007. – 90 p.

КРАЕВЫЕ ЗАДАЧИ ДЛЯ СИНГУЛЯРНО ВОЗМУЩЕННЫХ ИНТЕГРО-ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Дауылбаев М.К., Атахан Н.
e-mail: mdauylbayev@gmail.com

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Аннотация. *Предложены и обоснованы методы качественного исследования асимптотического поведения решений краевых задач с одним и двумя пограничными слоями для сингулярно возмущенных линейных интегро-дифференциальных уравнений. Доказаны предельные теоремы об аппроксимации решений сингулярно возмущенных задач. Установлены, что решения данных краевых задач обладают явлениями начальных и граничных скачков.*

1. Введение. Сингулярно возмущенными дифференциальными уравнениями называют уравнения, содержащие малый параметр при старшей производной. Такие уравнения имеют большую прикладную значимость. Они выступают в качестве математических моделей при исследовании разнообразных процессов в физике, химии, биологии и технике. В работах В. Вазова [1], М.И. Вишика, Л.А. Люстерника [2] и А.Б. Васильевой [3], посвященных краевым задачам для квазилинейных дифференциальных уравнений с малым параметром при старшей производной, установлены, что эти краевые задачи приводятся к задачам с сингулярными начальными условиями. Для нелинейного обыкновенного дифференциального уравнения второго порядка, содержащего малый параметр, задачи с неограниченными начальными условиями изучены в работе М.И. Вишика, Л.А. Люстерника [4] и К.А. Касымова [5]. Им установлено, что решение исходной

начальной задачи при стремлении малого параметра к нулю стремится к решению вырожденного уравнения с измененными начальными условиями. Такие задачи стали называться задачами Коши с начальным скачком. Наиболее общие случаи задачи Коши с начальным скачком для сингулярно возмущенных нелинейных систем обыкновенных дифференциальных и интегро-дифференциальных уравнений, а также для систем и уравнений в частных производных гиперболического типа исследованы К.А. Касымовым [6-8]. Краевые задачи для сингулярно возмущенных обыкновенных дифференциальных уравнений, обладающие начальным скачком решения рассмотрены в [9].

В данной работе для сингулярно возмущенных линейных интегро-дифференциальных уравнений высших порядков предложены методы качественного исследования асимптотического поведения решений краевых задач с одним пограничным слоем на случаи нелокальных и интегральных краевых условий. Тем самым, будет расширен класс сингулярно возмущенных интегро-дифференциальных уравнений, обладающих явлениями начальных скачков, для которых удастся построить асимптотическое разложение решений с любой степенью точности по малому параметру. Будут исследованы также краевые задачи с двумя пограничными слоями, обладающие явлениями граничных скачков, когда быстрая переменная решения становится неограниченной на обеих границах. Новизна этих исследований заключается в том, что в случае интегро-дифференциальных уравнений скачок примет не только решение, но и интегральный член. Решение этих задач позволяет расширить класс сингулярно возмущенных интегро-дифференциальных уравнений, обладающих явлениями начальных скачков, уравнениями с граничными скачками.

Задача сингулярных возмущений для нейронных сетей рассматривалась в статьях [10-14]. Сингулярно возмущенные интегро-дифференциальные уравнения рассматривались в статье [10], чтобы доказать существование локально единственного медленного волнового импульса, распространяющегося вдоль сети таламокортикальных и таламических ретикулярных нейронов. Эти результаты показывают применимость техники сингулярных возмущений к нейронным сетям.

1. Краевые задачи с одним пограничным слоем. Рассмотрим на отрезке $0 \leq t \leq 1$ следующее линейное сингулярно возмущенное интегро-дифференциальное уравнение:

$$L_\varepsilon y \equiv \varepsilon y^{(n)} + A_1(t)y^{(n-1)} + \dots + A_n(t)y = F(t) + \int_0^1 \sum_{i=0}^{m+1} H_i(t,x)y^{(i)}(x,\varepsilon)dx \quad (1)$$

с краевыми условиями:

$$h_i y(t,\varepsilon) \equiv \sum_{j=0}^m \alpha_{ij} y^{(j)}(0,\varepsilon) + \sum_{j=0}^r \beta_{ij} y^{(j)}(1,\varepsilon) = a_i, \quad i = \overline{1,n}, \quad m < n-1, r < n-1, \quad (2)$$

где $\varepsilon > 0$ – малый параметр, $\alpha_{ij}, \beta_{ij}, a_i$ – некоторые известные постоянные, не зависящие от ε . Пусть выполнены условия:

1. Функции $A_i(t), F(t), i = \overline{1,n}$ являются достаточно гладкими на отрезке $[0,1]$, а $H_i(t,x), i = \overline{0,m+1}$ – в области $D = (0 \leq t \leq 1, 0 \leq x \leq 1)$

$$2. A_1(t) \geq \gamma = \text{const} > 0, \quad 0 \leq t \leq 1$$

$$3. \bar{\Delta} = \begin{vmatrix} h_1 y_{10}(t) & \dots & h_1 y_{n-1,0}(t) & \alpha_{1m} \\ \dots & \dots & \dots & \dots \\ h_n y_{10}(t) & \dots & h_n y_{n-1,0}(t) & \alpha_{nm} \end{vmatrix} \neq 0$$

4. Число $\lambda = 1$ не является собственным значением ядра

$$H(t, s, \varepsilon) = \frac{1}{\varepsilon} \int_s^1 \sum_{i=0}^{m_1+1} H_i(t, x) K^{(i)}(x, s, \varepsilon) dx.$$

Теорема 1. Пусть выполнены условия 1-4. Тогда решение $y(t, \varepsilon)$ краевой задачи (1), (2) на отрезке $0 \leq t \leq 1$ представимо в виде:

$$y(t, \varepsilon) = \sum_{i=1}^n C_i(\varepsilon) Q_i(t, \varepsilon) + P(t, \varepsilon), \quad (3)$$

где

$$Q_i(t, \varepsilon) = \Phi_i(t, \varepsilon) + \frac{1}{\varepsilon} \int_0^t K(t, s, \varepsilon) \bar{\varphi}_i(s, \varepsilon) ds, \quad P(t, \varepsilon) = \frac{1}{\varepsilon} \int_0^t K(t, s, \varepsilon) \bar{F}(s, \varepsilon) ds, \quad (4)$$

а $C_i(\varepsilon), i = 1, \dots, n$ – решение системы алгебраических уравнений

$$\begin{cases} C_1 (1 + d_{11}(\varepsilon)) + C_2 d_{12}(\varepsilon) + \dots + C_n d_{1n}(\varepsilon) = a_1 - e_1(\varepsilon), \\ \dots & \dots & \dots & \dots \\ C_1 d_{n1}(\varepsilon) + C_2 d_{n2}(\varepsilon) + \dots + C_n (1 + d_{nn}(\varepsilon)) = a_n - e_n(\varepsilon), \end{cases} \quad (5)$$

где

$$d_{ik}(\varepsilon) = \sum_{j=0}^r \frac{\beta_{ij}}{\varepsilon} \int_0^1 K^{(j)}(1, s, \varepsilon) \bar{\varphi}_k(s, \varepsilon) ds, \quad i, k = 1, \dots, n,$$

$$e_i(\varepsilon) = \sum_{j=0}^r \frac{\beta_{ij}}{\varepsilon} \int_0^1 K^{(j)}(1, s, \varepsilon) \bar{F}(s, \varepsilon) ds, \quad i = 1, \dots, n.$$

$K(t, s, \varepsilon)$ - функция Коши, являющаяся решением задачи:

$$L_\varepsilon K(t, s, \varepsilon) = 0, \quad K^{(j)}(s, s, \varepsilon) = 0, \quad j = \overline{0, n-2}, \quad K^{(n-1)}(s, s, \varepsilon) = 1,$$

функции $\Phi_i(t, \varepsilon), i = \overline{1, n}$ называются граничными функциями краевой задачи (1), (2) и являются решениями следующей задачи:

$$L_\varepsilon \Phi_i(t, \varepsilon) = 0, \quad i = \overline{1, n}, \quad h_k \Phi_i(t, \varepsilon) = \begin{cases} 1, & i = k, \\ 0, & i \neq k. \end{cases}$$

Для функций Коши $K(t, s, \varepsilon)$ справедливы следующие асимптотические при $\varepsilon \rightarrow 0$ представления:

$$K^{(j)}(t, s, \varepsilon) = -\varepsilon \frac{\overline{W}_{n-1}^{(j)}(t, s)}{\mu(s)\overline{W}(s)} + \varepsilon^{n-1-j} \exp\left(\frac{1}{\varepsilon} \int_s^t \mu(x) dx\right) \frac{\mu^j(t) y_{n0}(t)}{\mu^{n-1}(s) y_{n0}(s)} + \quad (6)$$

$$+ O\left(\varepsilon^2 + \varepsilon^{n-j} \exp\left(\frac{1}{\varepsilon} \int_s^t \mu(x) dx\right)\right), j = \overline{0, n-1}, \quad \mu(t) = -A_1(t),$$

а для граничных функций $\Phi_i(t, \varepsilon), i = \overline{1, n}$ справедливы следующие асимптотические при $\varepsilon \rightarrow 0$ представления:

$$\Phi_1^{(j)}(t, \varepsilon) = -\varepsilon^{m_1-m_2} \frac{\alpha_{2, m_2} \overline{\Delta}_1^{(j)}(t)}{\mu^{m_1-m_2}(0) \alpha_{1, m_1} \overline{\Delta}} + \varepsilon^{m_1-j} \cdot \frac{y_{n0}(t) \mu^j(t)}{\mu^{m_1}(0) \alpha_{1, m_1}} \exp\left(\frac{1}{\varepsilon} \int_0^t \mu(x) dx\right) + \quad (7)$$

$$+ O\left(\varepsilon^{m_1-m_2+1} + \varepsilon^{m_1+1-j} \exp\left(\frac{1}{\varepsilon} \int_0^t \mu(x) dx\right)\right), j = \overline{0, n-1},$$

$$\Phi_i^{(j)}(t, \varepsilon) = \frac{\overline{\Delta}_{i-1}^{(j)}(t)}{\overline{\Delta}} + (-1)^{i-1} \varepsilon^{m_1-j} \cdot \frac{y_{n0}(t) \mu^j(t) \overline{\Delta}_i}{\mu^{m_1}(0) \alpha_{1, m_1} \overline{\Delta}} \exp\left(\frac{1}{\varepsilon} \int_0^t \mu(x) dx\right) +$$

$$+ O\left(\varepsilon + \varepsilon^{m_1+1-j} \exp\left(\frac{1}{\varepsilon} \int_0^t \mu(x) dx\right)\right), j = \overline{0, n-1}; i = \overline{2, n},$$

где $\overline{\Delta}_i^{(j)}(t)$ - определитель, полученный из $\overline{\Delta}$ заменой его i -ой строки фундаментальной системой решений $\overline{y}_1^{(j)}(t), \dots, \overline{y}_{n-1}^{(j)}(t)$ невозмущенного однородного уравнения.

Из формулы (3) с учетом (4) - (7) получаем следующую теорему

Теорема 2. Пусть выполнены условия I-VII. Тогда для решения $y(t, \varepsilon)$ краевой задачи (1), (2) и его производных на отрезке $0 \leq t \leq 1$ справедливы следующие асимптотические при $\varepsilon \rightarrow 0$ оценки:

$$\left| y^{(q)}(t, \varepsilon) \right| \leq C \left(1 + \max_{0 \leq t \leq 1} |F(t)| + \max_{0 \leq t \leq 1} |H_{m+1}(t, 0)| \right) + C \varepsilon^{m-q} \exp\left(-\gamma \frac{t}{\varepsilon}\right),$$

$$q = 0, \dots, n-2, 0 \leq t \leq 1,$$

где $C > 0, \gamma > 0$ - постоянные, не зависящие от ε .

Из теоремы 2 следует, что решение краевой задачи (1), (2) в точке $t=0$ имеет следующий порядок роста:

$$y^{(j)}(0, \varepsilon) = O(1), j = \overline{0, m}, y^{(m+j)}(0, \varepsilon) = O\left(\frac{1}{\varepsilon^j}\right), j = \overline{1, n-1-k}, \varepsilon \rightarrow 0.$$

Это означает, что решение краевой задачи (1), (2) в точке $t=0$ обладает явлением начального скачка m -го порядка.

2. Краевые задачи с двумя пограничными слоями. В предыдущем разделе рассмотрена краевая задача с одним пограничным слоем, для которой имеет место условие устойчивости 2. В этом разделе рассмотрим краевые задачи вида:

$$L_\varepsilon y \equiv \varepsilon^2 y''' + \varepsilon A_0(t)y'' + A_1(t)y' + A_2(t)y = F(t) + \int_0^1 \sum_{i=0}^1 H_i(t,x)y^{(i)}(x,\varepsilon)dx \quad (8)$$

с краевыми условиями

$$h_1 y \equiv y(0,\varepsilon) = \alpha, \quad h_2 y \equiv y'(0,\varepsilon) = \beta, \quad h_3 y \equiv y(1,\varepsilon) = \gamma \quad (9)$$

Пусть выполнены условия:

I. Функции $A_i(t), i = \overline{0,2}$, $F(t)$ на отрезке $0 \leq t \leq 1$, а $H_0(t,x), H_1(t,x)$ - в области $D = \{0 \leq t \leq 1, 0 \leq x \leq 1\}$ являются непрерывно дифференцируемыми.

II. $A_1(t) \neq 0, 0 \leq t \leq 1$.

III. Корни уравнения $\mu^2 + A_0(t)\mu + A_1(t) = 0$ удовлетворяют условиям $\mu_1(t) < 0, \mu_2(t) > 0$.

Строятся вспомогательные функции

$$K_0(t,s,\varepsilon) = \frac{P_0(t,s,\varepsilon)}{W(s,\varepsilon)}; \quad K_1(t,s,\varepsilon) = \frac{P_1(t,s,\varepsilon)}{W(s,\varepsilon)}, \quad (10)$$

где $W(s,\varepsilon)$ – вронскиан ФСР сингулярно возмущенного однородного уравнения, соответствующее уравнению (8), $P_0(t,s,\varepsilon), P_1(t,s,\varepsilon)$ – определители, получаемые из $W(s,\varepsilon)$ заменой третьей строки строками $y_1(t,\varepsilon), 0, y_3(t,\varepsilon)$ и $0, y_2(t,\varepsilon), 0$ соответственно. Из (10) для этих функций справедливы асимптотические представления:

$$K_0^{(q)}(t,s,\varepsilon) = \varepsilon^2 \left(\frac{y_{30}^{(q)}(t)}{y_{30}(s)\mu_1(s)\mu_2(s)} - \frac{\mu_1^{(q)}(t)y_{10}(t)}{\varepsilon^q y_{10}(s)\mu_1(s)(\mu_2(s) - \mu_1(s))} \exp\left(\frac{1}{\varepsilon} \int_s^t \mu_1(x)dx\right) + O(\varepsilon) \right),$$

$$K_1^{(q)}(t,s,\varepsilon) = \varepsilon^2 \left(\frac{\mu_2(t)y_{20}(t)}{\varepsilon^q y_{20}(s)\mu_2(s)(\mu_2(s) - \mu_1(s))} \exp\left(-\frac{1}{\varepsilon} \int_t^s \mu_2(x)dx\right) + O(\varepsilon) \right),$$

$$s \leq t, \quad t \leq s, \quad q = \overline{0,2}. \quad (11)$$

Функции $\Phi_i(t,\varepsilon), i = 1,2,3$, являющиеся решениями задачи

$$L_\varepsilon \Phi_i(t,\varepsilon) = 0, \quad h_k \Phi_i(t,\varepsilon) = \delta_{ki}, \quad i, k = 1,2,3$$

называются граничными функциями и для них справедливы асимптотические представления:

$$\begin{aligned} \Phi_1^{(i)}(t, \varepsilon) &= y_{30}^{(i)}(t) - \frac{\mu_1^i(t)y_{10}(t)y'_{30}(0)}{\varepsilon^{i-1}\mu_1(0)} \exp\left(\frac{1}{\varepsilon} \int_0^t \mu_1(x) dx\right) + \frac{\mu_2^i(t)y_{20}(t)y_{30}(1)}{\varepsilon^i y_{20}(1)} \times \\ &\times \exp\left(-\frac{1}{\varepsilon} \int_t^1 \mu_2(x) dx\right) + \\ &+ O\left(\varepsilon + \frac{1}{\varepsilon^{i-2}} \exp\left(\frac{1}{\varepsilon} \int_0^t \mu_1(x) dx\right) + \frac{1}{\varepsilon^{i-1}} \exp\left(-\frac{1}{\varepsilon} \int_t^1 \mu_2(x) dx\right)\right), i = 0, 1, 2, \end{aligned} \quad (12)$$

$$\begin{aligned} \Phi_2^{(i)}(t, \varepsilon) &= -\varepsilon \frac{y_{30}^{(i)}(t)}{\mu_1(0)} + \frac{\mu_1^i(t)y_{10}(t)}{\varepsilon^{i-1} \cdot \mu_1(0)} \exp\left(\frac{1}{\varepsilon} \int_0^t \mu_1(x) dx\right) + \\ &+ \frac{\mu_2^i(t)y_{20}(t)y_{30}(1)}{\varepsilon^{i-1} \cdot \mu_1(0)y_{20}(1)} \exp\left(-\frac{1}{\varepsilon} \int_t^1 \mu_2(x) dx\right) + \end{aligned}$$

$$+ O\left(\varepsilon^2 + \frac{1}{\varepsilon^{i-2}} \exp\left(\frac{1}{\varepsilon} \int_0^t \mu_1(x) dx\right) + \frac{1}{\varepsilon^{i-2}} \exp\left(-\frac{1}{\varepsilon} \int_t^1 \mu_2(x) dx\right)\right), i = 0, 1, 2,$$

$$\Phi_3^{(i)}(t, \varepsilon) = \frac{\mu_2^i(t)y_{20}(t)}{\varepsilon^i y_{20}(1)} \cdot \exp\left(-\frac{1}{\varepsilon} \int_t^1 \mu_2(x) dx\right) + O\left(\frac{1}{\varepsilon^{i-1}} \exp\left(-\frac{1}{\varepsilon} \int_t^1 \mu_2(x) dx\right)\right), i = 0, 1, 2$$

IV. Число 1 не является собственным значением ядра

$$H(t, s, \varepsilon) = \frac{1}{\varepsilon^2} \int \sum_{s=0}^1 H_i(t, x) K_0^{(i)}(x, s, \varepsilon) dx - \frac{1}{\varepsilon^2} \int \sum_{s=0}^1 H_i(t, x) K_1^{(i)}(x, s, \varepsilon) dx$$

Тогда из (11), (12) получаем справедливость следующей теоремы

Теорема. Если выполнены условия I-IV, то для решения краевой задачи (8), (9) при $\varepsilon \rightarrow 0$ справедливы следующие асимптотические оценки:

$$\begin{aligned} |y(t, \varepsilon)| &\leq C(\max_{0 \leq t \leq 1} |F(t) - \alpha H_1(t, 0)| + |\beta| + \varepsilon |\gamma|) + C e^{-\gamma_1 \frac{t}{\varepsilon}} (|\alpha| + |\beta| + \varepsilon |\gamma| + \\ &+ \varepsilon \max_{0 \leq t \leq 1} |F(t)|) + C \varepsilon e^{-\gamma_2 \frac{1-t}{\varepsilon}} (|\alpha| + |\beta| + |\gamma| + \max_{0 \leq t \leq 1} |F(t)|), \end{aligned}$$

$$\begin{aligned}
|y'(t, \varepsilon)| &\leq C(\max_{0 \leq t \leq 1} |F(t) - \alpha H_1(t, 0)| + |\beta| + \varepsilon |\gamma|) + \frac{C}{\varepsilon} e^{-\gamma_1 \frac{t}{\varepsilon}} (|\alpha| + |\beta| + \varepsilon |\gamma| + \\
&+ \varepsilon \max_{0 \leq t \leq 1} |F(t)|) + C e^{-\gamma_2 \frac{1-t}{\varepsilon}} (|\alpha| + |\beta| + |\gamma| + \max_{0 \leq t \leq 1} |F(t)|), \\
|y''(t, \varepsilon)| &\leq C(\max_{0 \leq t \leq 1} |F(t) - \alpha H_1(t, 0)| + |\beta| + \varepsilon |\gamma|) + \frac{C}{\varepsilon^2} e^{-\gamma_1 \frac{t}{\varepsilon}} (|\alpha| + |\beta| + \varepsilon |\gamma| + \\
&+ \varepsilon \max_{0 \leq t \leq 1} |F(t)|) + \frac{C}{\varepsilon} e^{-\gamma_2 \frac{1-t}{\varepsilon}} (|\alpha| + |\beta| + |\gamma| + \max_{0 \leq t \leq 1} |F(t)|),
\end{aligned} \tag{13}$$

где $C > 0, \gamma_i > 0, i = 1, 2$ – постоянные, не зависящие от ε .

Из оценок (13) получаем для решения задачи (8), (9) справедливо следующий порядок роста на границах рассматриваемого отрезка

$$y''(0, \varepsilon) = O\left(\frac{1}{\varepsilon}\right), y'(1, \varepsilon) = O\left(\frac{1}{\varepsilon}\right), y''(1, \varepsilon) = O\left(\frac{1}{\varepsilon^2}\right), \varepsilon \rightarrow 0$$

Отсюда, получаем, что решение данной краевой задачи в точке $t=0$ имеет скачок первого порядка, а в точке $t=1$ - нулевого порядка. В таком случае говорят, что решение данной краевой задачи обладает явлениями граничных скачков.

Литература

1. W. Wasow Singular perturbations of boundary value problems for nonlinear differential equations of the second order. *Comm. On Pure and Appl. Math.* 9. 1956. P. 93-113.
2. М.И. Вишик, Л.А. Люстерник Об асимптотике решения краевых задач для квазилинейных дифференциальных уравнений. // *ДАН СССР.* 1958. 121. № 5. С. 778-781.
3. А.Б. Васильева Асимптотика решений некоторых краевых задач для квазилинейных уравнений с малым параметром при старшей производной. // *ДАН СССР.* 1959. 123. №4. С. 583-586.
4. М.И. Вишик, Л.А. Люстерник О начальном скачке для нелинейных дифференциальных уравнений, содержащих малый параметр. // *ДАН СССР,* 1960, 132. № 6, С. 1242-1245.
5. К.А. Касымов Об асимптотике решения задачи Коши с большими начальными условиями для нелинейного обыкновенного дифференциального уравнения, содержащего малый параметр. // *УМН,* 1962, № 17. вып. 5, С. 187-188.
6. К.А. Касымов О задаче с начальным скачком для нелинейных систем дифференциальных уравнений, содержащих малый параметр. // *ДАН СССР.* 1968. № 179. Вып. 2. С. 275-278.
7. К.А. Касымов Асимптотика решения задачи Коши с начальным скачком для нелинейных систем интегро-дифференциальных уравнений, содержащих малый параметр при старшей производной. // *Известия АН КазССР. Серия физико-математическая.* 1968. № 5. С. 69-72.
8. К.А. Касымов Асимптотика решения задачи с начальными скачками для системы дифференциальных уравнений гиперболического типа с малым параметром при производной. // *ДАН СССР.* 1971. № 196. Вып. 2. С. 274-277.

9. K.A. Kasymov, D. Nurgabyl Asymptotic Estimates of Solution of a Singularly Perturbed Boundary Value Problem with an Initial Jump for Linear Differential Equations // *Differential Equations*, Vol.40, No.5, 2004, pp. 641-651.
10. Glyzin S. D., Kolesov A. Yu., Rozov N. Kh. On a Method for Mathematical Modeling of Chemical Synapses // *Differential Equations*. - 2013. - Vol. 49, No. 10. - P. 1193–1210.
11. Jalics J. Slow waves in mutually inhibitory neuronal networks // *Physica D*. - 2004. - № 192. –P. 95–122.
12. Han X., Xie W.-F., Fu Z., Luo W. Nonlinear systems identification using dynamic multi-time scale neural networks // *Neurocomputing*.- 2011. - №74. – P. 3428-3439.
13. Cao Y. and Wu J. Dynamics of Projective Adaptive Resonance Theory Model: The Foundation of PART Algorithm // *IEEE Transactions on Neural Networks*. - 2004. - Vol. 15, no. 2. - P. 245-260.
14. FitzHugh R.A. Impulses and physiological states in theoretical models of nerve membrane // *Biophysical J*. - 1961. - Vol. 1. - P. 445–466.

РАЗРАБОТКА ЛИНГВИСТИЧЕСКОГО И АЛГОРИТМИЧЕСКОГО ОБЕСПЕЧЕНИЯ ТЕХНОЛОГИЙ ИНФОРМАЦИОННОГО ПОИСКА

Еримбетова А.С., Мурзин Ф.А.
e-mail: aigerian@mail.ru, murzin@academ.org

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Институт систем информатики имени А.П.Ершова СО РАН, Россия

Аннотация. Создание новых методов семантического анализа текстов, в том числе и на агглютинативных языках, откроет новые возможности и позволит существенно продвинуться в решении многих задач компьютерной лингвистики, таких как машинный перевод, информационный поиск, автореферирование, классификация текстов и других, в которых также важна автоматизация тех или иных аспектов референции. Не менее актуальна разработка новых инструментов, позволяющих автоматизировать анализ текстов, и кастомизация имеющихся инструментов (для агглютинативных языков (казахского и турецкого)). Целью работы является разработка нового лингвистического и алгоритмического обеспечения технологий информационного поиска с учетом синтаксиса и элементов семантики, в том числе, для тюркоязычных текстов.

Методы автоматического анализа естественного языка (ЕЯ) – это наиболее динамично развивающиеся информационные технологии, востребованные в поисковых системах, системах реферирования, системах классификации и кластеризации текстов [1]. Тем не менее, в этих областях остаются задачи, до сих пор не нашедшие общепризнанного решения, и среди них семантический анализ, тесно связанный с задачей понимания текста машиной, самый сложный этап

автоматической обработки текста [2]. Текстовые документы и, соответственно, специальные знания, содержащиеся в них, обычно являются слабо структурированными. Это обстоятельство вызывает существенные трудности при формальном описании и автоматизации процесса извлечения и обработки знаний.

Существует много толкований термина «смысл текста» и его понимания машиной. Например, Д. А. Поспелов [3], говорит о том, что с точки зрения эксперта (или группы экспертов), факт того, что система понимает введенный в нее текст, означат, что она правильно отвечает на вопросы, связанные с информацией, содержащейся в тексте. При этом акцент делается на определении скрытого смысла, который вводит автор, а не о простом получении фактов, которые явно присутствуют в тексте. Д. А. Поспелов также выделяет некоторые уровни понимания текста с точки зрения сложности вопроса, который интеллектуальная система должна иметь возможность ответить.

В системах АОТ традиционно выделяют пять основных этапов анализа: графематический, морфологический, фрагментационный, синтаксический и семантический. Исследования ориентированы на казахский и русский языки, для охвата как агглютинативного, так и флективного типа языков. Было замечено, что при обработке текстов на казахском языке бывает сложно четко разделить этапы морфологического, синтаксического и семантического анализа, наблюдается их связность. Это объясняется особенностями словообразования в языках такого строя. Кроме того, именно эти три этапа вызывают наибольшие трудности при автоматизации независимо от языка.

Грамматика связей и анализ семантических отношений. Были проведены исследования по формальному описанию синтаксических связей казахского и турецкого языка с использованием функций синтаксического анализатора Link Grammar Parser (LGP), проведена работа по кастомизации LGP для указанных языков [4]. Выделены основные синтаксические связи для казахского и турецкого языков: AS – определение при подлежащем, АО – определение при дополнении, Е – обстоятельство при сказуемом, J – соединяет послелог с глаголом, OV – прямое дополнение при сказуемом, OJV – косвенное дополнение при сказуемом, S – соединяет подлежащее и сказуемое. При автоматическом разборе предложений LGP выделяет одновременно, как морфологические, так и синтаксические связи. Например, при разборе предложения *Адамдар алма жеді.* (Люди съели яблоко.) анализатор выделил две синтаксические (S3p, OV) и две морфологические (Np, Va3p) связи (Рис.1).

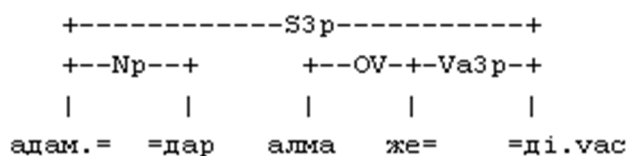


Рис. 1 Пример предложения с прямым дополнением

Для работы системы LGP создается словарь, в котором для каждого слова указывается разъем – тип связи, который может быть создан с использованием данного слова. Также были рассмотрены методы заполнения словарей LGP с помощью машинного обучения. В работе [5] исследованы возможностей алгоритма машинного обучения с внедрением технологии Word2Vec для определения грамматических категорий. С целью определения синтаксиса и семантических

связей в казахских текстах разработан алгоритм автоматизированного заполнения словарей LGP с использованием алгоритмов машинного обучения, таких как: LogReg, DTC, SVM, Random Forest, K-neighbors на примере модели KazKorpus. Для автоматической разметки частей речи была использована библиотека NLTK с использованием множественного двоичного классного подхода.

Реализация программы для определения релевантности текста поисковому запросу. Разработаны функциональная часть и визуальный модуль программы для определения релевантности текста поисковому запросу (Рис.2). Алгоритм написан на C++ с использованием LGP и сторонних стеммеров. Интерфейс реализован на языке со статической типизацией TypeScript с использованием платформы для разработки веб-приложений Angular 5, библиотеки реактивного программирования Rx.js, библиотеки компонентов Material Angular, фреймворка для построения нативных графических приложений для настольных операционных систем с помощью веб-технологий Electron. Из библиотеки Rx.js активно используется класс Observable для работы с асинхронными данными.

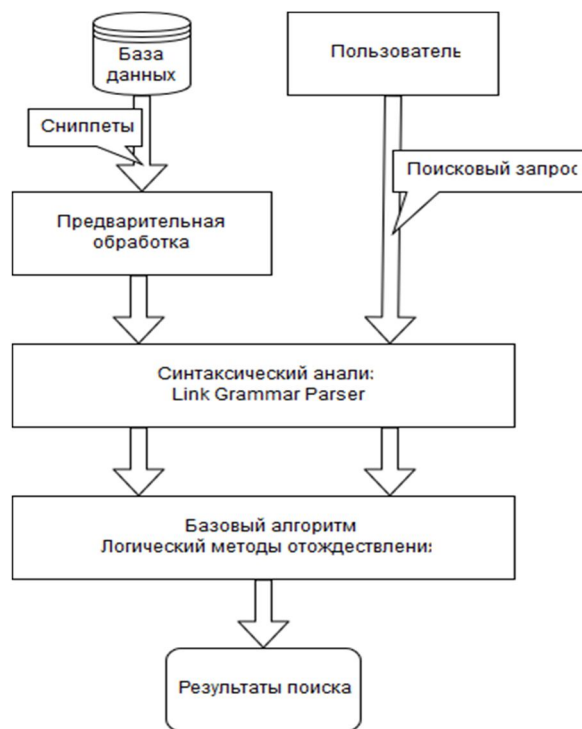


Рис. 2 Схема работы программной системы

Алгоритм описан отдельно от интерфейса. После того, как получены данные о релевантности, интерфейс отображает их в виде таблице. В процессе решения поставленных задач были разработаны способы представления семантико-синтаксических отношений между смысловыми единицами предложения на основе диаграмм, выдаваемых в процессе работы программной системой LGP. Также предложены способы вычисления степени совпадения естественно-языковых конструкций.

Определены функциональные требования к программному продукту, требования к интерфейсу, взаимодействию с другими системами. Для выполнения

пункта о лицензионных соглашениях нефункциональных требований было принято решение о работе над проектом в рамках лицензии MIT.

В качестве среды разработки использовались инструменты от компании JetBrains, а именно CLion и WebStorm. Использовалась студенческая лицензия, позволяющая использовать инструменты для некоммерческого использования. Основным инструментом для контроля версий был выбран Git, являющийся наиболее популярным инструментом для контроля версий. Для хранения кода в открытом доступе был выбран сервис GitHub. Для сборки функциональной части используется система сборки CMake, позволяющая описать сборку исходного кода, не делая предположений об операционной системе, компиляторе, линтере и других инструментах.

Выделены классы:

- src
 - app
 - dialog
 - DialogComponent
 - material
 - MaterialModule
 - Services
 - DialogService
 - ExecService
 - SnippetService
 - TabsService
 - table
 - TableComponent
 - tabs
 - TabsComponent
 - toolbar
 - ToolbarComponent
 - types
 - Snippet
 - Tab
 - AppModule
 - AppComponent

Классы с постфиксом Component и Module являются компонентом и модулем в смысле Angular. Модули и компоненты – классы в смысле TypeScript, аннотированные как @NgModule и @Component соответственно [6, 7].

AppModule – корневой модуль интерфейса. Объявляет необходимые описанные ниже компоненты, сервисы и модуль MaterialModule. AppComponent – корневой компонент интерфейса. DialogComponent – компонент диалога создания нового запроса. В AppModule объявлен как entryComponents [8] и создается динамически в DialogService при нажатии на кнопку New Request. После заполнения формы отправляет данные в TabsService.

MaterialModule – модуль, в котором объявлены используемые компоненты из библиотеки компонентов Material Angular. Подключается в модуле AppModule. Используется только для того, чтобы не перечислять компоненты в корневом модуле и, соответственно, его упрощения.

DialogService – сервис, ответственный за отображение диалога создания нового запроса, запрос сниппетов у сервиса SnippetService и передачи полученных данных в TabsService для создания вкладки с новым запросом. ExecService – сервис, ответственных за передачу сниппетов для обработки функциональной части. Здесь используется функция spawn, предоставляемая библиотекой Electron, для создания дочернего процесса. SnippetService – сервис, отвечающий за получение сниппетов из Google и передачу их алгоритму через сервис ExecService для анализа на релевантность поисковому запросу. Имеет один публичный метод getSnippets(text: string), возвращающий Observable<Snippet[]>. TabsService – сервис, отвечающий за хранение информации о вкладках типа Tab, их создание и удаление. Для создания вкладки принимает поисковый запрос и другие необходимые данные, запрашивает сниппеты и данные о релевантности из SnippetService в виде Observable<Snippet[]>.

TableComponent – компонент, ответственные за вывод результатов анализа на релевантность в виде таблицы.

TabsComponent – компонент, ответственный за отображение вкладок, переключения между ними и их закрытие.

ToobarComponent – компонент, отображающий некоторые элементы управления программой. При нажатии на кнопку New Request, компонент вызывает функцию openDialog из сервиса DialogService.

Snippet – класс с полями title, link, snippet типа string.

Tab – класс с полями title типа string и data типа Observable <Snippet[]>.

Для управления зависимостями в визуальном модуле выбран NPM, как стандартный инструмент в NodeJS. Для сборки визуальной части используется tsc (TypeScript compiler) и команды NPM, описанные в файле со списком зависимостей package.json.

Результаты анализа выводятся в виде таблицы и отсортированы по убыванию коэффициента релевантности.

Таким образом, разработанное программное обеспечение позволяет получать от пользователя поисковый запрос и проводить его предварительную обработку (экранирование некоторых символов, удаление символов переноса строки, невидимых символов и т.д.). Получать от поисковой системы результаты поиска по запросу пользователя, с использованием WebAPI поисковой системы или разбора полученных гиперссылок и текста, а также проводить их предварительную обработку (удаление символов переноса строки, невидимых символов и т.д.). Выводить результаты анализа в таблицу и сортировать их по убыванию коэффициента релевантности.

Для демонстрации эффективности работы системы были произведены тестовые испытания. Ниже даны примеры простых запросов на русском языке из области ИТ. Результаты тестирования базового показаны ниже (Таблица 1).

По каждому запросу загружались списки адресов с их описанием, которые поисковики обычно выдают пользователю. По этим коротким описаниям (сниппетам; англ. snippet) производилась оценка ресурса. Для сравнения с поисковой системой (а именно с системой Google) была составлена статистика. Система оставляла релевантные ссылки, отбрасывая нерелевантные по ее мнению. В итоге выяснилось, что на проведенных тестах в среднем из 100 ссылок, полученных из поискового сервиса Google, система выделяла 5–15 качественных релевантных ссылок, до 5 ссылок система ошибочно принимала за релевантные и остальные отбрасывала как нерелевантные, что соответствовало действительности. Это показывает, что данная система смогла произвести фильтрацию на хорошем уровне.

Отметим, что компании, предоставляющие поисковые сервисы, не заинтересованы, чтобы качественно осуществлять фильтрацию информации. Это позволяет им собрать информацию о пользователе, предложить рекламу и т.д., но это относится к области технической политики компаний.

В таблице 1, приведенной ниже, использованы следующие обозначения:

n_0 – общее количество фрагментов, полученных от Google;

n_1 – количество релевантных ссылок, одобренных системой;

n_2 – количество релевантных ссылок, пропущенных системой;

n_3 – количество не релевантных ссылок, одобренных системой.

Таблица 1. Результаты тестовых испытаний

Текст поискового запроса	n_0	n_1	n_2	n_3
Основные понятия теории информации Шеннона и понятие энтропии	140	20	0	0
Коды, исправляющие ошибки, их связь с избыточностью и энтропией	100	14	3	0
Статистические модели источников сообщений	160	0	0	0
Приложения статистики в лингвистике, исследовании музыки и в генетике	112	15	0	8
Элементы криптографии, коды открытого ключа и метод эллиптических кривых	120	19	0	2
Алгоритмические основы систем символьных преобразований, сличение с образцом	140	16	0	9
Источники и типы изображений, классификация алгоритмов обработки изображений	110	16	0	7

Тестирование системы для казахского и турецкого языков также показало хорошие результаты. Однако в настоящий момент здесь имеются меньшие возможности, т.к. соответствующие словари содержат недостаточно лексики. Было также обнаружено, что метод, учитывающий перефразирования предложений, позволяет улучшить работу системы, но часто незначительно по сравнению с базовым алгоритмом.

Качество работы системы в значительной мере зависит от полноты словарей. Объем словарей казахского и турецкого языков в настоящее время составляет около 500 слов и 100 аффиксов. Для данных языков также было проведено аналогичное тестирование. При этом, анализировались не короткие сниппеты, а довольно большие тексты, объемом 11-27 килобайт, чтобы «захватить» в анализируемых текстах достаточное количество лексики, представленной в словарях. Статистика, полученная в результате тестирования, примерно такая же, как для английского и русского языков.

В настоящее время поиск и анализ текстовой информации является важной задачей. Появились сложные теории о структуре самого текста и предлагающие развивать алгоритмы в направлении семантического анализа. В работе мы попытались найти золотую середину, ограничившись учетом определенных грамматических связей в тексте.

При решении задачи построения алгоритмов оценки релевантности текста поисковому запросу, были использованы семантико-синтаксические отношения между словами предложения (в том числе на казахском и турецком языках), получаемых на выходе программной системы LGR. Проведенное тестирование показало высокую эффективность предложенного алгоритма в решении задачи

определения релевантности текста поисковому запросу. Дальнейшее развитие предложенного метода представляется весьма затруднительным и легко не приводит к существенным улучшениям имеющихся результатов. Одной из причин является то, что на данном этапе возможности анализатора LGP почти полностью исчерпаны, он вынуждает оставаться на уровне синтаксиса с небольшим охватом семантики. Поэтому, чтобы получить существенное продвижение, необходимо перейти на более высокий уровень – к инженерии знаний. Были рассмотрены логические методы сопоставления предложений с использованием логики предикатов первого порядка.

Для решения сопутствующей проблемы информационного поиска, когда перестановка слов в предложении при определении тем текстов может существенно менять его смысл и приводит к некорректной работе алгоритмов, оперирующих отдельными ключевыми словами или их частотами и т. д. В работе использована идея Нирадж Кумара, использующую размытую логику Заде, и также позволяющую учесть порядок слов. Было предложено обобщение указанного метода с дополнительным учетом синтаксических отношений, получаемых на выходе системы LGP. При таком подходе основной задачей также является определение релевантности, а на основе анализатора LGP эта задача решается очень высококачественно.

Работа выполнена в рамках грантового финансирования по научным и (или) научно-техническим проектам на 2018-2020 годы МОН РК № AP05133550 и проект РФФИ № 18-07-01457-А.

Литература

- 1 Баранов, А.Н. Введение в прикладную лингвистику / А.Н. Баранов. – Москва.: Издательство ЛКИ, 2007. – 360 с.
- 2 Леонтьева, Н.Н. Стрoение семантического компонента в информационной модели автоматического понимания текста: автореф. ... д.т.н.: 05.13.17. – Москва.: 1990. – 46 с.
- 3 Поспелов, Д. А. Десять горячих точек в исследованиях по искусственному интеллекту // Интеллектуальные системы. – МГУ, 1996. – Т. 1, – № 1-4. – С. 47–56.
- 4 Batura, T.V., Tazhibayeva, S.Zh., Murzin, F.A., Yerimbetova, A.S., Sagnayeva S.K., Bakiyeva, A.M. Link Grammar Parser for turkic languages and algorithms for estimation the relevance // The 10th International Conference on Application of Information and Communication Technologies. – Baku, Azerbaijan, 2016. – P.104-107.
- 5 Tussupova, M.J., Murzin, F.A., Yerimbetova, A.S. Filling up Link Grammar Parser dictionaries by using Word2vec techniques // Совм. вып. по матер. Междунар. конф. «Вычислительные и информационные технологии в науке, технике и образовании» (CITech-2018), Вестник ВКГТУ им. Д. Серикбаева, Вычислительные технологии. – Усть-Каменогорск – Новосибирск, 2018. – Т. 1, – 4.3. – С.169-176.
- 6 Angular - Introduction to modules // URL: <https://angular.io/guide/architecture-modules>: 13.05.2018
- 7 Angular - Introduction to components // URL: <https://angular.io/guide/architecture-components>: 13.05.2018
- 8 Angular - Entry Components // URL: <https://angular.io/guide/entry-components>: 13.05.2018

ОБЗОР ЛИТЕРАТУРЫ ПО ПРОБЛЕМЕ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ РЕЧНЫХ ПОТОКОВ

Зиятбекова Г.З.¹, Мазаков Т.Ж.¹, Кисала П.²

e-mail: ziyatbekova@mail.ru

¹*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

²*Люблинский технический университет, Польша*

Аннотация. В статье рассматриваются основные подходы к моделированию течения в реках и области их применения. Кроме того, обсуждаются общие вопросы гидродинамики рек, а также способы описания турбулентности и их применение к моделям речных течений. Задача статьи состоит в определении методов и средств мониторинга уровня воды в водоемах. Описанные методы основаны на использовании технических средств сбора и передачи информации о состоянии уровня воды в водоемах.

Ключевые слова: гидравлическое моделирование, волна прорыва, турбулентная гидродинамика, система мониторинга, прогнозирование последствий наводнений, защита территорий от затопления.

Введение. Поведение водного объекта оказывает существенное влияние на жизнь и деятельность человека. Ухудшение качества воды в реке в связи со сбросом сточных вод от промышленных предприятий и крупных городов, затопление прибрежных территорий во время локальных весенних паводков, деформации русла реки, угрожающие постройкам, значительно влияют на жизнь города. Экспериментальное исследование водоемов сопряжено с рядом экономических и технологических трудностей и требует множества измерений на протяженном отрезке времени для выявления тенденций, в то время как математическое моделирование позволяет получить картину течения и дает возможность оценить последствия хозяйственной деятельности. В настоящее время чистая пресная вода становится одним из дефицитных ресурсов. Большинство крупных городов и промышленных центров расположено на берегах рек.

Методы. Математические модели, сочетающие детальность описания течения с приемлемой для решения практических задач вычислительной сложностью, очень востребованы при изучении процессов, происходящих в окружающей среде.

Результаты. Математическое моделирование распространения загрязняющих веществ, попавших в реку в результате деятельности человека, позволяет решать задачи оптимизации расположения сбросов в реку сточных вод, оценки необходимой степени их очистки перед сбросом в водоем, а также определения участков накопления примесей в реке, где водоем не может использоваться в рекреационных целях. Намного менее изученной проблемой является применение математического моделирования для прогноза появления и оценки последствий локальных наводнений, зачастую вызываемых

загромождением льдинами речного русла во время ледохода. Особое внимание в данной задаче заслуживает построение модели движущихся в потоке ледяных частиц. Вопросы моделирования двухфазных течений жидкости с твердыми частицами возникают во многих задачах, связанных с моделированием течений в окружающей среде. Большой вклад в моделирование гидродинамики естественных водоемов внесли А.В. Караушев, О.Ф. Васильев, В.А. Шлычков, В.В. Беликов, А.Т. Зиновьев, И.И. Потапов, Е.А. Цветова, В.И. Квон, П.Ю. Пушистов, С.С. Храпов, А.Л. Чикин, В.М. Белолипецкий и другие отечественные исследователи. А также зарубежные ученые X. Wang, H.T. Shen, W. Rodi, M.-E. Vasques-Cendon, L. Yu, S. Beltaos, S. Kang, A.N. Sukhodolov, W. Wu. При построении моделей следует учитывать, что адекватное моделирование течения в водоеме невозможно без корректного описания турбулентного переноса и перемешивания. Модель турбулентности при этом должна сочетать относительную простоту, позволяющую применять ее для расчетов течений с подвижной границей в областях с нерегулярной геометрией и учитывать наиболее существенные факторы, определяющие структуру потока.

Расчет естественных водоемов требует учета гораздо большего числа факторов, чем расчет технических сооружений и требует материалов исследований (карт, гидрологических данных о реке и подстилающей поверхности русла) для верификации и калибровки численных моделей. Исследование гидрологических процессов в руслах рек с помощью экспериментальных установок (физических моделей) было начато еще во времена О. Рейнольдса, который исследовал участок реки Мерси и ее эстуарий (затопляемое устье реки, расширяющееся в сторону моря) с помощью лабораторной модели и представил результаты в докладе «On certain laws relating to there gime rivers and estuaries and on the possibility of experiment on small scale» [1]. С начала XX века во многих странах стали создаваться гидрологические лаборатории для исследования различных явлений механики жидкости, появляющихся в природе [2]. Наиболее крупными и активными из них являются лаборатория Дельфта (Нидерланды), экспериментальная станция водных путей корпуса армейских инженеров США (Виксбург), Русловая лаборатория Государственного гидрологического института (Москва), гидравлическая лаборатория Касима (Цукуба, Япония). Оснащение ведущих лабораторий позволяет проводить исследования режимов течения рек, перенос концентрации потоком, деформации русел, волновых процессов в эстуариях, распространение паводковых волн с целью их прогнозирования и проектирования защитных сооружений, и многие другие задачи. В середине XX века появилось и развивалось множество лабораторий в Германии, Франции, США, СССР, Японии, Китае [2]. Наиболее полный обзор экспериментальных исследований русловых потоков дан в [2]. В этой работе проанализированы различные подходы к экспериментальному и гидравлическому моделированию речных потоков, представленные в работах советских гидрологов. В настоящее время экспериментальные исследования турбулентной гидродинамики активно продолжаются как для течений в открытых каналах [3,4,5,6], так и для речных течений и направлены на анализ структуры потока [3,7], переноса взвешенных и донных наносов в реках [8,9,10,11], изменения структуры потока при образовании препятствий [12,13]. На основе обзора экспериментальных исследований авторы [3] делают вывод о том, что турбулентность оказывает существенное влияние на все гидрологические процессы, происходящие в водоеме.

Однако экспериментальное исследование водоемов связано с рядом экономических и технологических трудностей, которые можно избежать, применяя

методы математического моделирования. Развитие моделей и методов расчета природных течений в XX в. связано с активным развитием промышленности и необходимостью прогнозирования поведения водного объекта в случае изменения его режима (строительства плотины, проведение дноуглубительных работ, разработка месторождений гравия и др.).

В середине XX в., когда исследование течений методами механики сплошных сред только начинались, было предложено несколько моделей для течения в русле реки. В своей работе [14] А.В. Караушев на основе диффузионной теории турбулентности представил комплексное исследование и приемы расчетов речных потоков и водохранилищ с помощью гидродинамических, гидравлических и экспериментальных методов [14]. Исследование включает технику расчета задачи о транспортировке наносов, переносе тепла и примеси в потоках, течения под ледяным покровом, а также русловых деформаций. Для расчета течения в речном русле предложена двумерная модель с одной горизонтальной и одной вертикальной координатой. В рамках используемого подхода «фоновые деформации» русла и транспорт наносов на протяженных участках рассчитываются с помощью простейшего уравнения баланса наносов.

Одной из первых попыток моделирования крупномасштабных течений можно назвать модель речного бассейна Stanford Watershed Model (SWM), разработанную в Стэнфордском университете, которая была впервые опубликована в 1962, а затем в 1966 [15]. С помощью SWM модели исследовались такие проблемы как: перенос радионуклидов, перенос наносов, накопление и таяние снега, моделирование качества воды и др. Модель относится к классу имитационных моделей, для вычисления характеристик течения используются эмпирические соотношения, полученные на основе экспериментальных данных.

По мере развития математических моделей и численных методов со второй половины XX в. математическое моделирование все шире применяется для исследования различных процессов, протекающих в естественных водоемах. Наиболее часто возникающие на практике задачи, связанные с необходимостью моделирования течения в русле реки с определенной точностью – это моделирование деформаций русла под воздействием потока и хозяйственной деятельности человека, моделирование распространения примеси, попавшей в водоем в результате аварийного сброса или постоянно поступающей в результате деятельности промышленных предприятий, и моделирование затопления прибрежных территорий в результате прорыва плотин и чрезвычайных ситуаций на технологических сооружениях. Намного менее изученной проблемой является применение математического моделирования для прогноза возникновения и оценки последствий локальных наводнений, возникающих во время ледохода.

Причиной возникновения водной угрозы может стать как неконтролируемое изменение климатических условий, так и деятельность человека. Большая доля природных катастроф, источниками которых являются водоемы, вызвана чрезмерным увеличением уровня воды. Защитные сооружения неспособны противостоять силе, которая превосходит их характеристики в несколько раз. В то же время слишком низкий уровень воды в реках препятствует развитию речного судоходства. Изменению уровня воды в озерах предшествуют различные факторы. Засушливая погода впоследствии может стать причиной уменьшения уровня воды в озерах. Увеличение количества осадков, как следствие, вызывает пополнение объема воды. Реки подвержены еще большему количеству влияний на их уровень. Например, сброс воды с плотины может сказаться на уровне воды в реке на далеком расстоянии от плотины. Некоторые реки протекают на территории разных

государств, вследствие чего обычный сброс воды в реку может стать поводом для политического скандала или вовсе оказаться причиной стихийного бедствия. На незначительные изменения уровня воды в реке способен повлиять даже ветер, в зависимости от его направления относительно течения реки. Встречный ветер может замедлять течение, следовательно ниже по реке уровень воды будет уменьшаться, а вверх по реке - подниматься. Для своевременного принятия мер по недопущению прорыва гидротехнических сооружений и естественных водоемов и рек необходим постоянный мониторинг уровня воды в контролируемом водоеме. Несмотря на наличие множества методов измерения уровня воды в водоемах, от самых примитивных до высокотехнологичных, – идея создания системы мониторинга является актуальной, так как она предоставляет автоматизированные средства регистрации уровня воды.

Мониторинговые системы должны обеспечивать постоянное наблюдение за явлениями и процессами, происходящими в природе и техносфере, для предвидения нарастающих угроз для человека и среды его обитания. Главной целью мониторинга является предоставление данных для точного и достоверного прогноза чрезвычайных ситуаций на основе объединения интеллектуальных, информационных и технологических возможностей различных ведомств и организаций, занимающихся наблюдением за отдельными видами опасностей. Мониторинговая информация служит основой для прогнозирования. Система мониторинга может быть связана с другими автоматизированными системами, например, системами забора и сброса воды на плотинах гидроэлектростанций, системами тревожного оповещения, и прочими функционально взаимосвязанными системами. Внедрение системы мониторинга уровня воды позволит осуществлять профилактику возникновения чрезвычайных ситуаций.

Попытки разрешить конфликт между необходимостью использования пойменных и прибрежных земель и убытками от возможных наводнений предпринимались неоднократно многими специалистами. Но досих пор этот конфликт неразрешен. Для решения задачи о возможности использования прибрежных земель необходимо проводить анализ возможного ущерба при наводнениях, вызванных в частности разрушением плотин. В сельскохозяйственных районах особенно велики убытки, вызванные затоплением сельскохозяйственных угодий. Так наводнения наносят ущерб сельскохозяйственным культурам, повреждают различные инженерные и мелиоративные системы.

Оценить качество и точность моделирования распространения волны прорыва поруслуреки является непростой задачей. Истинный результат в таких исследованиях никому не известен. Также не существует единственно верного способа, позволяющего получить достоверные результаты. При использовании любого из методов получают только ориентировочные данные о зонах затопления, получившихся в результате прохождения волны прорыва. Точность полученных результатов оценить сложно. Для оценки экологической безопасности региона необходим анализ больших объемов разнородной информации, несогласованность целей различных государственных органов. Решение таких задач невозможно без использования современных информационных систем поддержки принятия решений. В последнее время широкое распространение получили системы мониторинга, основанные на ГИС-технологии, которые помогают ускорить и повысить эффективность процедуры принятия решений, обеспечивают представление результатов анализа в наглядном и удобном для восприятия виде.

Для практического применения разрабатываются системы мониторинга состояния водоемов в реальном времени, основанные на автоматизации процесса сбора и обработки информации. В основном автоматизированные системы мониторинга применяют следующие датчики: инклометрические; течи; деформации; температуры; давления на грунт; уровня воды [16-18].

Для информационных систем мониторинга нужно обеспечить сбор данных в реальном режиме времени. Для решения данной проблемы можно использовать системы, которые в комплексе с измерительными датчиками могут собирать и передавать данные. Примером таких систем могут быть «Автоматизированная подсистема энергоснабжения спутниковых систем передачи данных» и «Унифицированная система передачи данных».

Обсуждение. Обобщив все обязательные условия, можно прийти к выводу, что решение задачи мониторинга уровня воды в озерах, реках, морях и гидротехнических сооружениях находится в применении систем «Автоматизированная подсистема энергоснабжения спутниковых систем передачи данных» и «Унифицированная система передачи данных» в комплексе с измерительными средствами. Унифицированная система передачи данных предоставляет средства и инструменты для организации передачи данных посредством спутниковых систем связи. Автономная подсистема энергоснабжения для спутниковых систем передачи данных включает в себя комплекс оборудования, функции которого заключаются в выработке и накоплении энергии для последующего его использования в целях обеспечения электропитания оборудования передачи данных. Основные концепции и полная структура указанных систем описываются в других статьях, посвященных этим системам.

Работа выполнена за счет средств грантового финансирования научных исследований на 2018-2020 годы по проекту АР05132044 «Разработка аппаратно-медицинского комплекса оценки психофизиологических параметров человека».

Литература

1. Рейнольдс О. Динамическая теория движения несжимаемой вязкой жидкости и определение критерия // Проблемы турбулентности. – М.-Л.: ОНТИ НКТП СССР, 1936. – С. 185-228.
2. Клавен А.Б. Экспериментальные исследования и гидравлическое моделирование речных потоков и руслового процесса / А.Б. Клавен, З.Д. Копалиани. – СПб: Нестор-История, 2011. – 504 с.
3. Blanckaert K. Mean flow and turbulence in open-channel bend / K. Blanckaert, W.H. Graf // Journal of hydraulic engineering. – 2001. – Vol. 127, № 10. – P. 835-847.
4. Hou J. A robust well-balanced model on unstructured grids for shallow water flows with wetting and drying over complex topography / J. Hou [et. al.] // Computer methods in applied mechanics and engineering. – 2013. – №. 257. – P. 126-149.
5. Pu J.H. Numerical and experimental turbulence studies on shallow open channel flows / J.H. Pu, S.D. Shao, Y.F. Huang // Journal of Hydro-environment Research. – 2014. – №. 8. – P. 9-19.
6. Pu J.H. Turbulence modelling of shallow water flows using Kolmogorov approach // Computers & Fluids. – 2015. – Vol. 115. – P. 66-74.
7. Roy A.G. Size, shape and dynamics of large-scale turbulent flow structures in a gravel-bed river / A.G. Roy [et. al.] // Journal of fluid mechanics. – 2004. – Vol. 500. – P. 1-27.

8. Paiement-Paradis G. Effects of turbulence on the transport of individual particles as bedload in a gravel-bed river / G. Paiement-Paradis, G. Marquis, A. Roy // Earth surface processes and landforms. – 2011. – Vol. 36. – P. 107-116.
9. Schmeeckle M. W. Forces on stationary particles in near bed turbulent flows / M. W. Schmeeckle, J. M. Nelson, R. L. Shreve // Journal of Geophysical Research. – Vol. 112, № F2. – P. F02003.
10. Sumer B.M. Influence of turbulence on bed load sediment transport / B.M. Sumer [et. al.] // Journal of Hydraulic Engineering. – 2003. – Vol. 129. – P. 585-596. 147
11. Nelson J.M. Role of near-bed turbulence structure in bed load transport and bed forms mechanics / J.M. Nelson [et. al.] // Water Resources Research. – 1995. – Vol. 31, № 8. – P. 2071-2086.
12. Sukhodolov A.N. Shallow wake behind exposed wood-induced bar in a gravel-bed river / A.N. Sukhodolov, T.A. Sukhodolov // Environmental Fluid Mechanics. – 2014. – Vol. 14, № 5. – P. 1071-1083.
13. Sukhodolov A.N. Dynamics of shallow lateral shear layers: Experimental study in a river with a sandy bed / A.N. Sukhodolov, I. Schnauder, W.S.J. Uijtewaal // Water Resources Research. – 2010. – Vol. 46, № 11. – P. W11519.
14. Караушев А.В. Проблемы динамики естественных водных потоков / А.В. Караушев. – Ленинград: Гидрометеорологическое издательство, 1960. – 392 с.
15. Crawford N. H. Digital Simulation in Hydrology: The Stanford Watershed Model IV: Technical Report / N. H. Crawford, R. S. Linsley. – Palo Alto: Department of Civil Engineering, Stanford University, 1966. – 210 p.
16. Котюк А.Ф. Датчики в современных измерениях. – М.: «Радио и связь», 2006. – 96 с.
17. Фрайден Дж. Современные датчики. – М.: «Техносфера», 2005. – 592 с.
18. Алейников А.Ф. Гридчин В.А. Цапенко М.П. Датчики (перспективные направления развития). – Новосибирск: НГТУ, 2001. – 176 с.

ПРОГРАММА РАСЧЕТА ПАРАМЕТРОВ НЕПОЗИЦИОННОЙ СИСТЕМЫ СЧИСЛЕНИЯ

Калимолдаев М.Н., Тынымбаев С.Т., Магзом М.М.,

Тананова Д.Д.

e-mail: magzomxzn@gmail.com

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** В данном докладе приведены промежуточные результаты выполнения задачи проекта грантового финансирования КН МОН РК «Разработка программно-аппаратных средств для криптосистем на базе непозиционной системы счисления». Идея данного проекта заключается в разработке различных вариантов умножителей полиномов по модулю неприводимых полиномов и их программно-аппаратная реализация.*

Введение

Основной целью данного проекта является разработка программно-аппаратных средств, выполняющих различные вычисления в непозиционной системе счисления для эффективной реализации алгоритмов криптографической защиты информации. Первой задачей ставится разработка программ для выполнения вычислений параметров полиномиальной непозиционной системы счисления (НПС).

В результате поиска путей повышения эффективности программно-аппаратных вычислений, методов обнаружения и исправления ошибок и создания высоконадежных компьютерных систем проводятся исследования в области непозиционных систем обозначений, таких как система остаточных классов (СОК) [1, 2]. В классической позиционной системе счисления значение каждой цифры в обозначении номера зависит от ее положения. В непозиционных системах счисления, напротив, обозначение чисел основано на других принципах. В СОК многозначное целое число в позиционной нотации представляется как последовательность нескольких позиционных чисел небольшой разрядности. Эти числа являются остатками от деления исходного числа на модули СОК. Повышение эффективности при аппаратной реализации данных систем обеспечивается за счет правил полиномиальной НПС, в которой все арифметические операции могут выполняться параллельно по модулям оснований НПС. Особенности НПС дают значительные преимущества перед позиционной системой счисления при выполнении модульных операций сложения, вычитания и умножения. Особенно это актуально, если в качестве операндов выступают многоразрядные числа [3].

Работа криптосистемы на основе полиномиальной НПС предусматривает выполнение предварительных расчётов параметров непозиционной полиномиальной системы. Данные расчёты включают в себя следующие шаги:

- создание базы данных неприводимых полиномов с коэффициентами над $GF(2)$;
- определение необходимой длины N блока и ключа шифрования;
- формирование в соответствии с длиной блока системы оснований. Для формирования полиномиальной НПС при обработке блока длиной N бит из множества всех неприводимых многочленов степени не выше значения N выбираются рабочие основания.
- подбор и проверка содержания системы оснований. Все выбираемые основания должны отличаться друг от друга, даже если они являются неприводимыми полиномами одной степени. Тогда в этой системе любой многочлен степени меньше суммы степеней всех выбранных рабочих оснований имеет единственное представление в виде последовательности остатков (вычетов) от деления его на данные основания.
- предварительная интерпретация входного блока данных в виде остатков от деления на выбранную систему оснований. При этом сгенерированная ключевая последовательность так же может быть представлена в виде последовательностей вычетов, а для расшифрования необходимо вычислить обратный многочлен к каждому из этих вычетов по модулю соответствующего основания.

Процесс шифрования и расшифрования с помощью нетрадиционного алгоритма шифрования на базе полиномиальной СОК описаны в [4, 5]. Общая схема работы программно-аппаратной криптосистемы на основе НПС показана на рисунке 1.

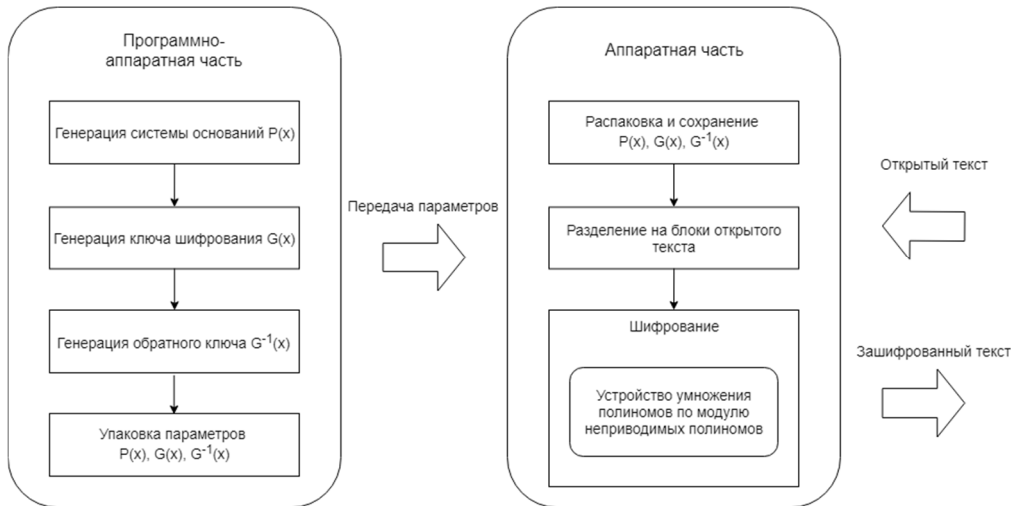


Рис. 1. Работа программно-аппаратной криптосистемы на основе НПС

Основная часть расчетов параметров НПС включает в себя формирование системы модулей полиномиальной системы остаточных классов. Рабочее основание системы формируется из предварительно рассчитанной базы данных неприводимых многочленов. Эти полиномы выбираются случайным образом в соответствии с предоставленными степенями для заполнения длины блока входных данных, как описано в [6].

Разработка программы расчёта параметров полиномиальной НПС

При разработке структуры программы расчёта применена объектно-ориентированная и доменно-ориентированные подходы. Был создан отдельный пакет, включающий в себя описание интерфейсов для выполнения указанных выше операций с полиномами и их реализацию в виде классов.

Классы описывают структуру полинома с бинарными коэффициентами. Основным классом является класс `BinaryPolynomial`, реализующий интерфейс `Polynomial`. Бинарные коэффициенты полинома хранятся в классе `BitSet` из программного пакета `Java`.

Полиномы представлены в виде последовательности битов, где каждый разряд соответствует позиции члена полинома, например, полином $x^6+x^5+x^2+1$ будет представлен в виде последовательности 1100101.

Для выполнения арифметических операций с полиномами с двоичными коэффициентами реализован класс `BinaryPolynomialMath`.

Сложение в $GF(2)$ следует правилам арифметики по модулю 2 без переноса. Это побитовая двоичная операция исключающего ИЛИ, пример которой показан на рисунке 2.

$$\begin{array}{r}
 x^6+x^5+x^2+1 \\
 + \quad x^4+x^3+x+1 \\
 \hline
 x^6+x^5+x^4+x^3+x^2+x
 \end{array}$$

Рис. 2. Операция сложения

Вычитание в GF(2) идентично сложению и следует правилам арифметики по модулю 2 без переноса. Пример вычитания показан на рисунке 3.

$$\begin{array}{r}
 x^6+x^5+x^2+1 \\
 - \quad \quad \quad x^4+x+1 \\
 \hline
 x^6+x^5+x^4+x^2+x
 \end{array}$$

Рис. 3. Операция вычитания

Операция умножения не отличается от таковой для десятичных чисел, за исключением того, что при вычислении конечной суммы необходимо использовать приведение по модулю 2, что показано на рисунке 4.

$$\begin{array}{r}
 x^3+x^2+1 \\
 * \quad \quad \quad x^3+1 \\
 \hline
 x^6+x^5+x^2+1
 \end{array}$$

Рис. 4. Операция умножения

Таким же образом выполняется деление с учетом остатка, согласно правилам арифметики в GF(2), как показано на рисунке 5.

$$\begin{array}{r}
 x^6+x^5+x^2+1 \\
 \hline
 x^4+x^3+x+1
 \end{array} = x^2, \text{ Остаток: } x^3+1$$

Рис. 5. Операция деления с остатком

Эта операция используется для приведения по модулю неприводимого полинома. Полином, который не имеет других факторов, кроме 1 и его самого, называется неприводимым многочленом. Проверка на неприводимость полинома выполняется с применением теста Рабина [7], алгоритм которого показан на рисунке 6.

```

Algorithm Rabin Irreducibility Test
Input: A monic polynomial  $f$  in  $F_q[x]$  of degree  $n$ ,
        $p_1, \dots, p_k$  all distinct prime divisors of  $n$ .
Output: Either " $f$  is irreducible" or " $f$  is reducible".
Begin
  for  $j = 1$  to  $k$  do
     $n_j = n/p_j$ ;
    for  $i = 1$  to  $k$  do
       $h := x^{q^{n_i}} - x \text{ mod } f$ ;
       $g := \text{gcd}(f, h)$ ;
      if  $g \neq 1$ , then return ' $f$  is reducible' and STOP;
    end for;
   $g := x^{q^n} - x \text{ mod } f$ ;
  if  $g = \emptyset$ , then return " $f$  is irreducible",
  else return " $f$  is reducible"
end.

```

Рис. 6. Алгоритм теста Рабина на неприводимость

В состав BinaryPolynomialMath также входит метод выполняющий расширенный алгоритм Евклида. Этот метод используется для вычисления обратного полинома, необходимого для расшифрования [5]. Обратный полином $\beta_i^{-1}(x)$ определяется из сравнения:

$$\beta_i(x)\beta_i^{-1}(x) \equiv 1 \pmod{p_i(x)}, \quad i = \overline{1, S} \quad (1)$$

Полученный полином $G^{-1}(x) = (\beta_1^{-1}(x), \beta_2^{-1}(x), \dots, \beta_S^{-1}(x))$ является обратным к полиному $G(x)$.

Генерация псевдослучайной последовательности выполняется с помощью генератора проверенного класса javax.crypto.KeyGenerator платформы Java. Этот функционал инкапсулируется в разработанном классе NPNKeyGenerator. Также в него входят классы для генерации так называемого полного ключа шифрования, использующие различные источники энтропии операционной системы. Подробное описание формирования полных ключей шифрования в НПСС рассматривается в [8].

Пользовательский интерфейс разработанной программы показан на рисунке 7. Разработанная программа работает в операционных системах Windows, Linux и MacOS.

При разработке части программы, реализующей умножение полиномов по модулю системы оснований из неприводимых полиномов, были использованы два программных подхода. Умножение и деление на каждое основание из системы оснований можно выполнять последовательно в программном цикле, либо с применением параллельных потоков для коллекций, появившихся в Java 8 [9].

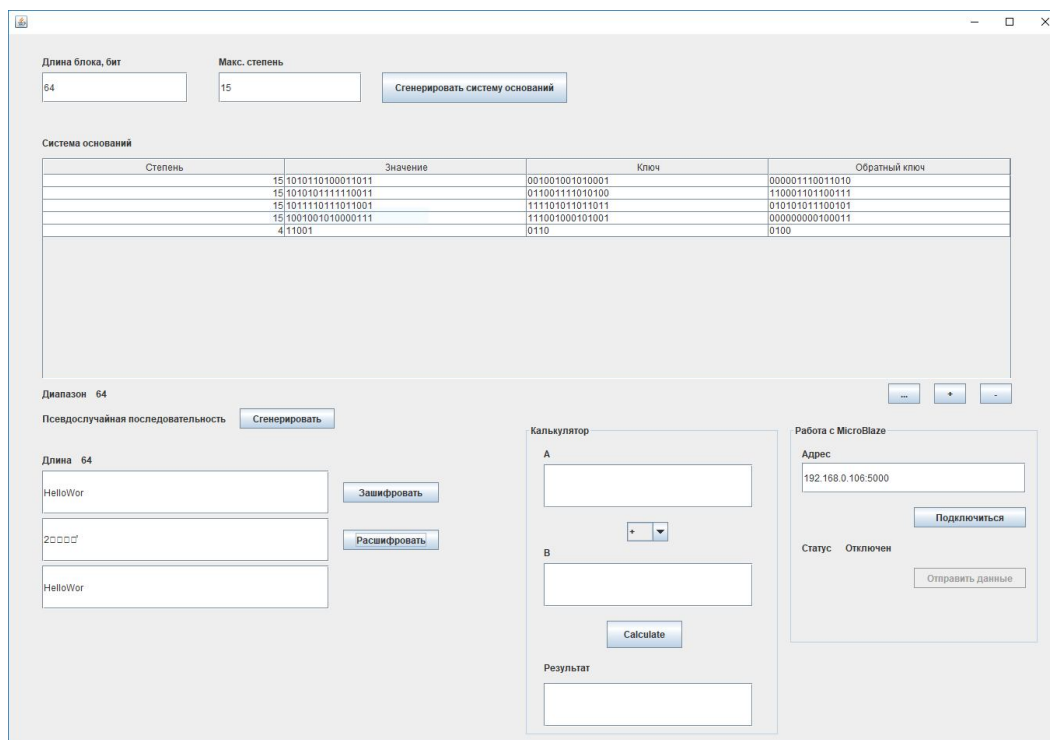


Рис. 7. Основное окно разработанной программы

В качестве теста проводилось умножение серии из 64 пар полиномов по модулю системы оснований, состоящей из фиксированного количества неприводимых полиномов. В таблицах 1 и 2 показано время, затраченное на умножение по модулю в последовательном и параллельном режимах.

Как видно, для работы с полиномами большой степени, и, следовательно, с двоичными последовательностями большой длины, требуется значительное время. В связи с этим актуально выполнение подобных ресурсоемких операций в программно-аппаратном и аппаратном видах.

Таблица 1. Умножение полиномов в последовательном режиме

Максимальная степень полинома в системе оснований	Количество многочленов в системе оснований	Время выполнения, мс
24	3	2912
24	4	2515
22	6	1947
20	5	2142
17	8	1567
14	9	1309
12	12	1147
9	13	1071
7	15	1020
7	14	955

Таблица 2. Умножение полиномов в параллельном режиме

Максимальная степень полинома в системе оснований	Количество многочленов в системе оснований	Время выполнения, мс
24	3	1646
24	4	1689
22	6	1430
20	5	1337
17	8	1110
14	9	930
12	12	847
9	13	800
7	15	791
7	14	725

Заключение

Разработана программа расчета параметров полиномиальной НПС на основе Java с функционалом формирования системы оснований НПС, выполнения арифметических операций над GF(2) и проверки шифрования с помощью базового алгоритма шифрования в непозиционной полиномиальной системы счисления. Разработанный функционал служит для выполнения рутинного вычисления, необходимого для аппаратной реализации алгоритма симметричного шифрования данных на основе полиномиальной НПС. Дальнейшая работа заключается в синтезе разрабатываемых схем на основе ПЛИС и испытание их работы в составе криптосистем на основе полиномиальной НПС.

Литература

- 1 Svoboda A. Valach M. Operatorove obvody // Stroje Na Zpracovani Informaci – 1955. – Vol 3. – P. 247-295.
- 2 Акушский И.Я., Юдицкий Д.И. Машинная арифметика в остаточных классах. – М.: Советское радио, 1968.
- 3 Бияшев Р.Г., Нысанбаева С.Е. Алгоритм формирования электронной цифровой подписи с возможностью обнаружения и исправления ошибки // Кибернетика и системный анализ. – 2012. – Т. 48, № 4. – С. 14-23.
- 4 Бияшев Р.Г. Разработка и исследование методов сквозного повышения достоверности в системах обмена данными распределенных АСУ: дисс. докт. тех. наук: 05.13.06: защищена 09.10. 1985: утв. 28.03.1986. - М., 1985. - 328 с.
- 5 Нысанбаев Р.К. Криптографический метод на основе полиномиальных оснований // Вестник Мин-ва науки и высшего образования и Нац. акад. наук Республики Казахстан – Алматы: Гылым, 1999. – № 5. – С. 63-65.
- 6 Biyashev R., Nyssanbayeva S., Begimbayeva Ye., Magzom M. Building modified modular cryptographic systems // International Journal of Applied Mathematics and Informatic. – 2015. – Vol. 9. – P. 103-109.
- 7 Rabin M. Probabilistic algorithms in finite fields // SIAM Journal on Computing. – 1980. – №2. – P. 273–280.
- 8 Бияшев Р.Г., Нысанбаева С.Е., Капалова Н.А. Секретные ключи для непозиционных криптосистем. Разработка, исследование и применение. –Lambert Academic Publishing, 2014.– 136 с.

9 HaiTao Mei, Ian Gray, Andy Wellings. Integrating Java 8 Streams with The Real-Time Specification for Java // Proceedings of the 13th International Workshop on Java Technologies for Real-time and Embedded Systems. – Paris, France, 2015. – Article No. 10. – P. 43-53.

МОДЕЛЬ УПРАВЛЕНИЯ КЛЮЧАМИ АЛГОРИТМОВ ШИФРОВАНИЯ ДАННЫХ

Капалова Н., Хаумен А.

e-mail: kapalova@ipic.kz, haumen.armanbek@gmail.com

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** В данной работе предложена модель управления криптографическими ключами, которая предназначена для применения в ранее разработанном алгоритме шифрования ВС-2. Данная схема основана на создании единой базы ключей для всех пользователей. База ключей заполняется ключами, т.е. псевдослучайными последовательностями определенной длины, которые генерируются с помощью алгоритма ГПСП. Предложенную схему можно использовать и для других алгоритмов шифрования.*

В июне 2017 года Правительством РК была утверждена Концепция кибербезопасности («Киберщит Казахстана»). Она определяет основные направления реализации государственной политики в сфере защиты электронных информационных ресурсов, информационных систем и сетей телекоммуникаций, обеспечения безопасного использования информационно-коммуникационных технологий [1, 2].

Одним из приоритетных направлений концепции является создание отечественных алгоритмов криптографической защиты данных, применяемых в информационных системах. В лаборатории информационной безопасности Института информационных и вычислительных технологий разрабатываются различные модели алгоритмов шифрования данных, в том числе и алгоритм блочного симметричного шифрования ВС-2 [3-6]. В данное время исследования продолжаются, проводится тестирование разработанных алгоритмов и осуществляются реализации моделей в различных средах компьютерной системы.

Как известно, помимо выбора подходящей для конкретной информационной системы средств криптографической защиты информации, важной проблемой является управление ключами. Как бы ни была сложна и надежна сама криптосистема, она основана на использовании ключей. Если для обеспечения конфиденциального обмена информацией между двумя пользователями процесс обмена ключами тривиален, то в информационной системе, где количество пользователей составляет сотни и тысячи, управление ключами – серьезная проблема [7].

Следующим этапом работы является организация управления ключами для созданных алгоритмов шифрования. В данной работе предлагается одна модель управления ключами для симметричных криптографических алгоритмов.

Модель управления ключами для алгоритмов шифрования

Перед тем, как пользователи начнут использовать алгоритм шифрования для передачи сообщения, каждому участнику необходимо создать базу ключей, так называемую «мастер-ключей». Для этого один из пользователей берет на себя функцию генерации «первичных ключей» и выбирает три целых числа (A , B , C). Где первые 2 числа (A , B) служат начальными значениями для генерации псевдослучайных последовательностей, а третье число (C) указывает длину генерируемых последовательностей.

Считается, что (A , B , C) безопасно распределяются остальным пользователям (для этого можно использовать асимметричные криптосистемы или физическое распределение при встрече, рис. 1).

Общая схема генерации мастер-ключей показана на рисунке 2. Теперь подробно опишем процесс генерации мастер-ключей.

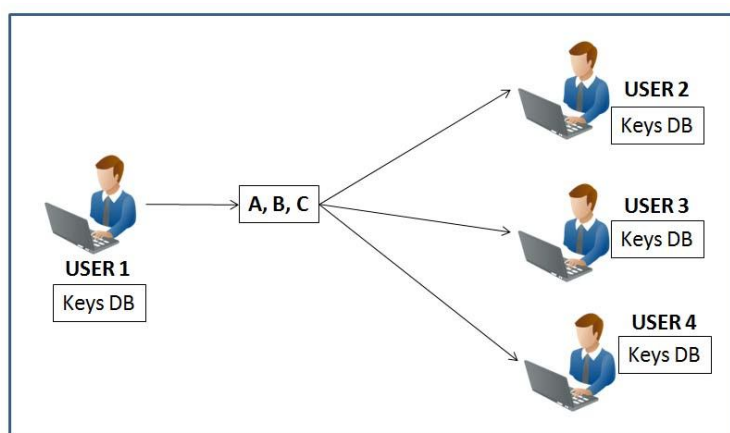


Рис. 1. Общая схема передачи первичных ключей

На основе чисел A , B генерируется псевдослучайная последовательность длиной C : $k_1, k_2, k_3, \dots, k_C$. Для генерации используется генератор псевдослучайных последовательности (ГПСП), разработанный ранее [8].

На следующем этапе из генерированных последовательностей вырезается последовательность длиной, равной длине блока алгоритма шифрования. Полученная последовательность преобразуется в мастер-ключ (PRIV_KEYS) с помощью преобразования KSMB (Рис. 2), состоящего из следующих преобразований: табличная замена байтов (S-Block), линейное преобразование MB (MixBlock) и сложение с ключевым словом по модулю 2. Табличная замена и преобразование MB взяты из алгоритма BC-2.

Для преобразования KSMB ключом преобразования служит ключевое слово пользователя. Пользователь придумывает ключевое слово любой длины, затем с помощью функции хэширования SHA получает хэш своего ключевого слова, который будет использоваться в качестве ключа для преобразования KSMB.

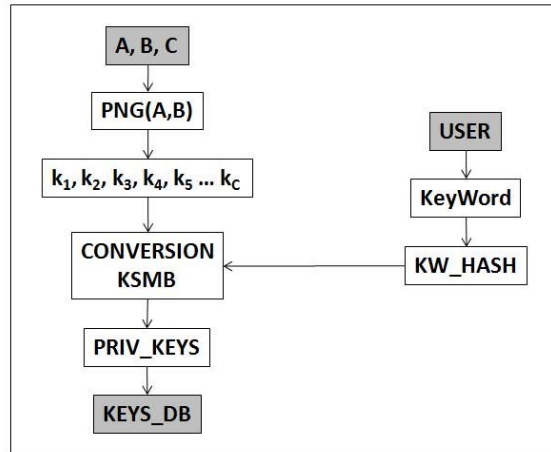


Рис. 2. Генерация мастер-ключей для алгоритмов шифрования

Полученные числа (PRIV_KEYS) в результате преобразования KSMB, записываются в базу ключей. И таким образом из последовательности псевдослучайных последовательностей вырезается следующая порция последовательности и формируется очередной мастер-ключ.

В итоге у каждого пользователя, который участвует в обмене данных с помощью алгоритма шифрования, сформировывается единая база мастер-ключей, зашифрованных хэш значением его ключевого слова (Рис. 3).

База мастер-ключей состоит из одной таблицы, имеющей следующие поля (столбцы): ID_KEYS – идентификационный номер ключа, PRIV_KEYS – зашифрованный ключ для алгоритма шифрования, USED – метка использования ключа, CHECK_KEYS – проверочный ключ (для аутентификации пользователя).

ID_KEYS	PRIV_KEYS	USED	CHECK_KEYS
1	15 45 0 2 56 89 125 7 2 53...	YES	7815469745632669501...
2	77 12 89 205 59 64 31 39...	NO	236598451258748596...

Рис. 3. Структура таблицы для мастер-ключей

Для получения ключа алгоритма шифрования пользователь должен пройти аутентификацию. Для этого необходимо вести свое ключевое слово. Из этих ключевых слов формируется хэш значение (KW_HASH). Далее из базы выбирается неиспользованный ключ (PRIV_KEYS) по метке USED. Полученные данные KW_HASH и PRIV_KEYS преобразуются в SMB_KEYS с помощью преобразования SMB (Рис. 4). Затем сгенерированные числа SMB_KEYS сравниваются со значениями из поля CHECK_KEYS.

Если значения совпадают, то из PRIV_KEYS получают ключи для алгоритма шифрования с помощью обратных преобразований INV_KSMB. В противном случае программа выдает сообщение об ошибке при аутентификации.

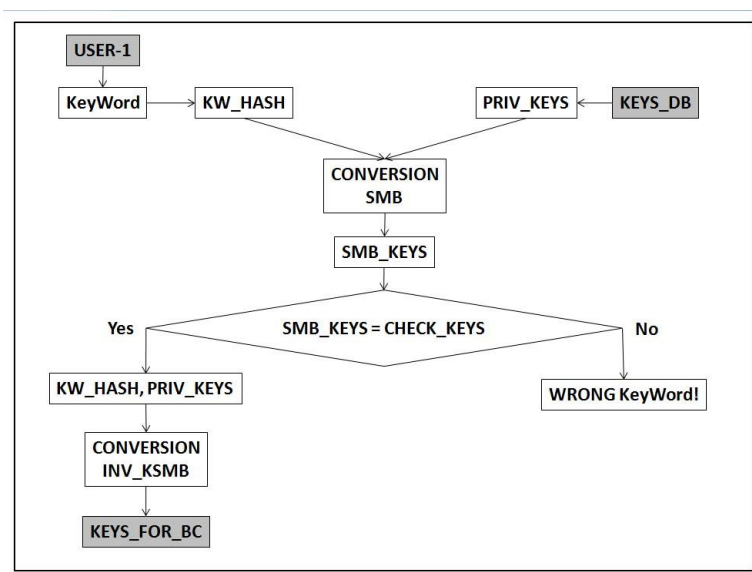


Рис. 4. Аутентификация пользователя и получение мастер-ключа

Проверочный ключ CHECK_KEYS генерируется с помощью преобразования SMB и записывается в таблицу при первоначальном создании самой базы (Рис. 5). В преобразовании SMB – S-табличная замена байтов (S-Block), MB-линейное преобразование (MixBlock).

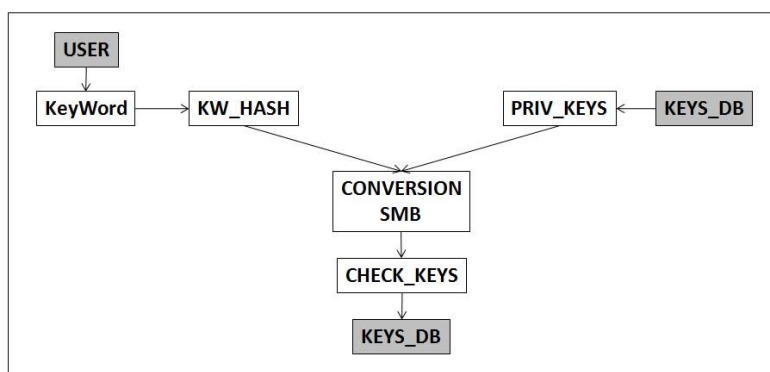


Рис. 5. Генерация проверочных ключей

После успешной аутентификации и получения ключа для шифрования в базу по ID_KEYS ключа ставится метка в поле USED, которая означает, что данный ключ уже использован, в дальнейшем этот ключ не будет выбран для шифрования. Но для расшифровки сообщения эти же ключи могут использоваться независимо от того, что ключ использован или не использован для шифрования.

Предложенная модель управления ключами будет использоваться в рамках работ по созданию моделей алгоритмов шифрования данных для компьютерных систем, в том числе для алгоритма шифрования BC-2.

Следующими этапами работ являются исследование надежности предложенной модели и ее программная реализация. Будут также проводиться работы по тестированию и внедрению модели управления ключами в криптографические системы.

Проводимые исследования финансируются Министерством образования и науки Республики Казахстан по теме гранта № AP05132568 «Разработка системы управления криптографическими ключами».

Литература

1. Концепция кибербезопасности ("Киберцит Казахстана") / Утверждена постановлением Правительства Республики Казахстан от 30 июня 2017 года № 407.
2. Каримова М. «Киберцит Казахстана»: первые итоги реализации. – <https://strategy2050.kz/ru/news/52085/> (дата просмотра: 26.05.2019)
3. Амербаев В.М., Бияшев Р.Г., Нысанбаева С.Е. Применение непозиционных систем счисления при криптографической защите // Известия Национальной академии наук Республики Казахстан. Серия физико-математическая. - Алматы: Гылым, 2005. - № 3. - С. 84-89.
4. Kapalova N., Naumen A., The model of encryption algorithm based on non-positional polynomial notations and constructed on an SP-network // Open Engineering – 2018. – Volume 8, Issue 1. – P. 140-146.
5. Biyashev R.G., Kalimoldayev M.N., Nyssanbayeva S.E., Kapalova N.A., Dyusenbayev D.S., Algazy K.T., Development and analysis of the encryption algorithm in nonpositional polynomial notations // Eurasian Journal of Mathematical and Computer Applications. – 2018. - № 6(2). - С.19-33.
6. Отчет о научно-исследовательской работе «Разработка программных и программно-аппаратных средств для криптографической защиты информации при ее передаче и хранении в инфокоммуникационных системах и сетях общего назначения», 2018 г., № гос. регистрации 0118РК01064.
7. Фомина И.А. Управления ключами в криптографических системах. Вестник нижегородского университета им. Н.И.Лобачевского. 2010, №4 (1), стр. 165-169.
8. Капалова Н.А., Нысанбаева С.Е. Разработка алгоритма и оценка качества генератора псевдослучайных последовательностей // Новости науки Казахстана: Науч.-техн. сб. – Алматы: НЦ НТИ, 2007. – Вып. 4. - С. 105-111.

КОМПЕТЕНТНОСТНАЯ МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ И КОНЦЕПЦИЯ SMART-КОНТРАКТА В ПРОЕКТИРОВАНИИ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Кубеков Б.С., Утегенова А.У., Науменко В.В., Аленова Р.А.

e-mail: b.kubekov@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** В статье представлено решение проблемы перехода на новую информационную базу обучения, основанную на онтологическом инжиниринге и компетентностной модели представления знаний. Методика проектирования знаниевой архитектуры образовательной программы основана на описании сценария реализации знаниевых компонент, smart-контракты которых позволяют конфигурировать знаниевые компоненты в соответствии с целями обучения. На*

материале известной монографии Хассан Гома «UML. Проектирование систем реального времени, параллельных и распределенных приложений», для этапа Conceive- проектно-ориентированной технологии обучения, рассмотрены формализмы отображения знаний через онтологии опорных понятий и их спецификации в виде выражений знаний. На примере профессиональных компетенций этапа Conceive, проекта "Банковская система типа клиент-сервер", продемонстрирован механизм конфигурирования знаниевых компонент параметрами smart-контракта.

***Ключевые слова:** онтология опорного понятия, инженерия знаний, CDIO, проектный метод обучения, компетентностная модель, smart-контракт, знаниевый компонент.*

Введение

Современный бум развития информационно-коммуникационных технологий и инженерии знаний позволяет существенным образом влиять на организацию высшего профессионального образования. Данные технологии привели не только к созданию так называемых смарт-технологий, призванных повысить результативность и эффективность профессионального обучения, но и к появлению возможности индивидуализации образовательных траекторий и дифференциации обучения.

В этой связи становятся актуальными исследования по формированию новой парадигмы информационной базы обучения и диверсификации методик электронной педагогики и ассортимента образовательных услуг, а также по освоению новых образовательных Smart-технологий.

Новые потребности студентов бросают вызовы университетам, поэтому инженерия знаний должна стать основой модернизации образовательных программ. С другой стороны, большую роль начинает играть неформальное обучение, когда студент сам определяет цели обучения, находит материал в открытых образовательных ресурсах, моделирует себе курс и, таким образом, формирует свою знаниевую архитектуру специализации. Гибкость, приспособляемость, качественные показатели, инновации - этим требованиям должны соответствовать smart-университеты, чтобы успевать за происходящими изменениями и растущими запросами студентов. Собственно, само понятие Smart подразумевает более быстрый ответ на требования экономики и мира.

Smart-технологии, несомненно, могут повысить производительность труда педагогов и организаторов, однако без «умного» образовательного контента, такие технологии окажутся очередной дорогостоящей игрушкой. Успешность использования образовательного контента возможна лишь за счет управления академическими знаниями, когда знаниевые компоненты могут быть идентифицированы и описаны соответствующими структурами понятий семантического контекста областей знаний. Необходимый и достаточный набор подобных повторно используемых знаниевых компонент позволит удовлетворять потребности каждого обучающегося. Помимо этого, возможность переноса репозитория знаниевых компонент в сферу Интернет, создает предпосылки взаимосвязи образовательных сообществ, обновления репозитория наиболее актуальными знаниями и совершенствования самого процесса обмена знаниями.

Как известно, большое преимущество получают те системы знаний, которые способны адаптироваться к изменениям и редуцировать сложность реальности до приемлемого уровня. Этим условиям в полной мере отвечает Smart-обучение, которое подразумевает гибкое обучение с точки зрения предпочтений и

индивидуальных возможностей обучающихся; поддержку индивидуальных траекторий для личностного развития обучающихся; создание и использование различных мотивационных моделей обучения и, пожалуй, самое главное, обеспечивает постоянную взаимосвязь между требованиями работодателей и содержанием образования. Условиями достижения последнего является внедрение проектно-ориентированной технологии обучения, инициированной MIT под названием CDIO (Conceive, Design, Implement, Operate), формирование компетентностных моделей этапов CDIO и их представление знаниевыми компонентами, каждый из которых идентифицируется и описывается соответствующими онтологиями опорных понятий.

В данной статье рассматриваются концепция компетентностной модели представления знаний и применение smart-контракта для формализации процесса проектирования знаниевой архитектуры образовательной программы. Введенные с этой целью формализмы отображения знаний, параметры smart-контракта дают возможность, во-первых, конфигурировать знаниевые компоненты, в соответствии с потребностями обучающихся, и, во-вторых, подвергать формальной верификации компетентностные модели и сценарии траектории обучения, с целью доказательства их корректности компетентностной модели выпускника.

1 Формализмы отображения знаний

В наших исследованиях, в качестве базовой модели отображения знаний предметной области, используется модель онтологии: $Om = \langle C, R, F \rangle$, где C – множество понятий предметной области; R – множество отношений между понятиями; F – функции интерпретации, определения которых заданы на отношениях между понятиями в онтологии.

Под семантическим контекстом предметной области будем понимать совокупность базовых абстракций, называемых опорными понятиями. Для каждого опорного понятия строится онтология, представляющая собой иерархическую структуру понятий, корневой вершиной которой является опорное понятие, первый уровень – идентифицирующие понятия опорного понятия, второй уровень – конкретизирующие понятия каждого из идентифицирующих понятий онтологии. Основное назначение идентифицирующего понятия – определить семантические и отличительные свойства опорного понятия, относительно других опорных понятий предметной области. Роль конкретизирующих понятий заключается в конфигурировании семантических и отличительных свойств опорного понятия, посредством типичных, либо всевозможных сочетаний дочерних понятий, обладающих ясным и недвусмысленным описанием опорного понятия.

Между понятиями онтологии опорного понятия определены отношения «композиция», «агрегация», «альтернативный выбор», помеченные символами ‘*’, ‘+’, ‘~’. В соответствии с введенными отношениями, понятия в онтологии наделяются свойствами общности, или обязательности присутствия, и изменчивости, то есть необязательности или альтернативности их присутствия в отношении [1-5]. Следует отметить, что понятия выражения знания, вместе с определенными над ними отношениями, имплицитно истинность заключения об опорном понятии. Введенные таким образом типы отношений над понятиями и их свойства, обеспечивают релевантное моделирование семантического контекста предметной области, представляемого онтологиями опорных понятий.

2 Формализмы спецификации опорных понятий

Формализмы спецификации опорных понятий покажем на примере опорного понятия - C_1 - ICONIX - метода проектирования систем в архитектуре MDA.

Данное опорное понятие идентифицируется понятиями первого уровня онтологии: $C_{1.1}$, $C_{1.2}$ и $C_{1.3}$, с помощью которых определяются *семантические* и *отличительные свойства* опорного понятия C_1 , относительно других понятий домена обучения. Идентифицирующие понятия, таким образом, связаны с опорным понятием онтологии *ассоциативно - семантическими* признаками, что подчеркивает уникальность опорного понятия в масштабах понятий домена обучения:

* $C_{1.1}$ - MDA архитектура. Данное идентифицирующее понятие является семантическим контекстом второго уровня онтологии, включающего следующие конкретизирующие понятия: * C_1 - PIM модель; + C_2 - PSM модель.

* $C_{1.2}$ - классы анализа. Данное идентифицирующее понятие является семантическим контекстом второго уровня онтологии, включающего следующие конкретизирующие понятия: * C_1 - классы-сущности (entity); * C_2 - граничные классы (boundary); * C_3 - классы управления (control).

* $C_{1.3}$ - анализ пригодности. Данное идентифицирующее понятие является семантическим контекстом второго уровня онтологии, включающего следующие конкретизирующие понятия: * C_1 - свойства и поведение доменных объектов предметной области; * C_2 - правила построения диаграмм пригодности; * C_3 - диаграмма последовательности; * C_4 - эволюционирующая объектная модель.

Далее, с помощью выражения знания специфицируем онтологию опорного понятия последовательностью из идентифицирующих и конкретизирующих понятий, с определенными над понятиями отношениями «композиция», «агрегация» и «альтернативный выбор». Выражение знания, таким образом, имплицитно истинность заключения об опорном понятии, то есть, если «*посылка* – выражение знания» верно, то это же верно и для «*заключение* – опорное понятие».

Выражение знания опорного понятия C_1 - ICONIX - метод проектирования систем в архитектуре MDA, будет иметь следующий вид:

$$C_1 \Leftarrow *C_{1.1} (*C_1 + C_2) *C_{1.2} (*C_1 * C_2 * C_3) *C_{1.3} (*C_1 * C_2 * C_3 * C_4); \quad (1)$$

а визуальный образ онтологии - в виде следующего реляционного графа:

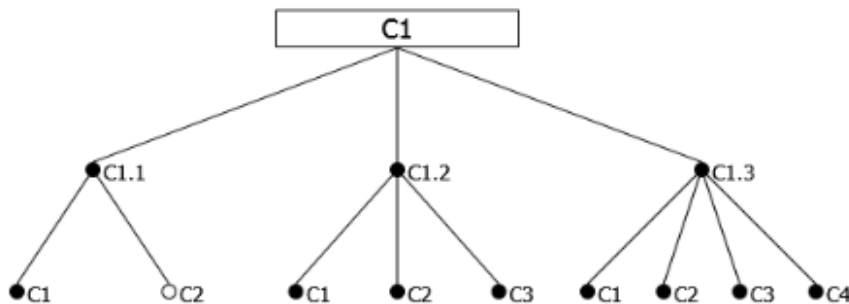


Рисунок 1 – Реляционный граф опорного понятия C_1 - ICONIX - метода проектирования систем в архитектуре MDA

Принятая в наших исследованиях методика моделирования и отображения знаний основана на концепции компетентностного подхода, этапах инициативы

CDIO - проектно-ориентированной технологии обучения [7-9]. Введенные правила и свойства описания понятий позволяют релевантно моделировать знания предметной области в виде онтологий опорных понятий и специфицировать их выражениями знаний. Выражение знания, таким образом, является оптимальной спецификацией, как с точки зрения компактного и релевантного отображения, так и с точки зрения хранения и редактирования знаний.

Дальнейшее использование выражений знаний будет связано с их композицией в виде знаниевых компонент, и последующим их применением для проектирования знаниевой архитектуры образовательной программы.

Для этих целей, введены определения следующих понятий: знаниевый компонент, компонентная модель и компонентная среда.

Знаниевый компонент (КС, Knowledge Component) – композиция выражений знаний, представляющий собой структурный элемент среды и конфигурационного управления, и обладающий четко определенным smart-контрактом, который фактически описывает правила реализации знаниевого компонента в образовательной программе. Другими словами, знаниевый компонент представляет собой выделенную область специальных знаний, требуемых для решения определенных практических задач и взаимодействующий с другими знаниевыми компонентами среды, посредством следующих параметров smart-контракта [10]:

$$\{P\} < KC, Col, Lev > \{Q\} \quad (2)$$

Параметр P - предусловие, то есть утверждение в виде сигнатуры требуемых компетентностей, необходимых для успешности обучения, предусмотренных знаниевым компонентом КС. Предусловие всегда должно выполняться при обращении к знаниевому компоненту КС, иначе корректность результата обучения не гарантируется.

Параметр Q – постусловие, то есть утверждение в виде сигнатуры компетентностей, гарантируемых обучающемуся, при выполнении предусловия и успешном освоении знаний, предоставляемых знаниевым компонентом КС.

Постусловие Q определяет результаты реализации знаниевого компонента КС, то есть корректность требуемых компетентностей, получаемых обучающимся. Что касается вызываемых знаниевых компонент, выполнение их постусловий должно гарантироваться ими самими, поэтому вызывающий знаниевый компонент КС может на них опираться в своей реализации.

При реализации smart-контракта знаниевого компонента, его предусловия могут ослабляться, а постусловия – только усиливаться. Это означает, что знаниевый компонент может реализовать свои обязанности для более широкого диапазона ограничений, чем это требуется его предусловием, и, в результате своей реализации, выполнять более строгие ограничения, чем это требуется его постусловием.

Таким образом, знаниевый компонент можно считать единицей развертывания знаниевой архитектуры, который может присоединяться в образовательную программу, если в ней присутствуют все знаниевые компоненты, от которых он зависит, либо удален из образовательной программы и тогда могут перестать работать те, которые зависят от него.

Следующими параметрами smart-контракта являются: Lev-параметр, задающий ограничения на уровень сложности компетентностей, предусматриваемых в пред- и постусловиях знаниевого компонента; Col-параметр, задающий ссылки на знаниевые компоненты образовательной среды, которые

задействуются при реализации данного знаниевого компонента. Параметры Lev и Col, таким образом, определяют инфраструктуру, позволяющую учитывать индивидуальные особенности обучающегося, а знаниевым компонентам, взаимодействовать по определенным правилам.

Под компонентной моделью понимается smart-контракт знаниевого компонента, правила его конфигурирования и реализации, в соответствии со сценарием взаимодействия с другими знаниевыми компонентами образовательной среды. Можно сказать, что компонентная модель определяет «язык», на котором знаниевые компоненты взаимодействуют друг с другом по определенному сценарию.

Для работы знаниевых компонент необходим некоторый набор базовых служб и сервисов образовательной среды, обеспечивающих функционирование знаниевых компонент. Такой набор служб и сервисов, вместе с поддерживаемой с их помощью компонентной моделью, называется компонентной средой.

3 Концепция сценариев проектирования знаниевой архитектуры

В наших исследованиях концепция проектирования знаниевой архитектуры дисциплин образовательной программы основана на так называемых smart-контрактах, с помощью которых моделируются сценарии знаниевой архитектуры.

Методика формирования выражений знаний основана на концепциях проектно-ориентированной технологии обучения, называемой инициативой CDIO, компетентностной модели этапов CDIO и онтологического инжиниринга предметной области. В результате формируется репозиторий образовательной среды из выражений знаний, на базе которого конструируются знаниевые компоненты, как композиции выражений знаний, взаимодействующие со своим окружением через smart-контракт.

Процесс построения знаниевой архитектуры дисциплин образовательной программы, начинается с выбора обучающимся одного из проектов среды, каждый из которых определяет концепцию образовательной программы обучающегося, и соответствующие компетентности, которые нужны ему в профессиональной деятельности.

Далее, обучающийся выбирает необходимый этап CDIO выполнения проекта и уровень сложности компетентностей этого этапа, подразделяемый на *профессиональные, базовые* либо *дополнительные* компетентности.

Дальнейшее уточнение знаниевой архитектуры связано с моделированием сценариев обучения, путем выбора вариантов структуризации знаниевых компонент и их конфигурирования с помощью параметров smart-контракта.

Понятно, что к наиболее сложному уровню относится уровень *профессиональных* компетентностей, который определяет знания, необходимые обучающемуся в условиях реального производства сложных и достаточно надежных программных систем, и техниках, используемых в современной индустрии программирования. Уровень *базовых* компетентностей представляет собой композицию компетентностей в областях знаний SWEBOOK и предметно-ориентированных *soft_skills*, которые обязательны для усвоения профессиональных компетентностей. И, наконец, уровень *дополнительных* компетентностей, представляет собой агрегацию компетентностей, которые обеспечивают усвоение *базовых* компетентностей.

Конфигурирование уровня сложности компетентностей предусматривается в пред- и постусловиях знаниевого компонента и реализуется параметром Lev, а, с

помощью параметра Col, конфигурируются ссылки на знаниевые компоненты, которые должны использоваться в данном знаниевом компоненте.

Таким образом, введенные параметры smart-контракта знаниевого компонента, позволяют рассматривать знаниевый компонент как элемент конфигурационного управления процесса проектирования знаниевой архитектуры образовательной программы. Концептуально, архитектура образовательной программы представляется в виде сценария взаимодействия знаниевых компонент, реализация которых задается параметрами smart-контракта.

Введенная концепция сценариев позволяет, во-первых, использовать smart-контракт для моделирования знаниевых компонент дисциплин и прототипа образовательной программы; во-вторых, проводить уточнение и верификацию знаниевой архитектуры, на предмет соответствия требованиям обучающихся; в-третьих, производить оценку выражений знаний репозитория среды; в-четвертых, производить мониторинг процесса обучения и успеваемости обучающихся; в-пятых, проводить анализ характеристик знаниевой архитектуры с целью оценки ее пригодности для обучения; и, наконец, находить проектные решения по формированию знаниевых компонент повторного использования.

4 Компетентностные модели этапов CDIO

В наших построениях, компетентностный подход подразумевает разработку компетентностных моделей этапов CDIO, являющихся композициями *профессиональных, базовых и дополнительных* компетенций, где каждая из компетенций определяется наборами онтологий опорных понятий.

Для спецификации *профессиональных, базовых и дополнительных* компетенций введем так называемое выражение компетенции (ес-expression of competence).

Допустим, что для успешной реализации ключевых вопросов, возникающих в процессе разработки проекта, на каждом из этапов CDIO, требуется усвоение профессиональных - PrC, базовых - BaC и дополнительных - AdC компетенций.

Продемонстрируем примеры выражений компетенций этапов Conceive, Design, Implement и Operate, следующим образом:

$$\begin{aligned} \text{ес.PrC.Conceive} &\leq *PrC_1 (*P_1*P_2)*PrC_2*PrC_3; \\ \text{ес.BaC.Conceive} &\leq *BaC_1 \sim +BaC_2*BaC_3 (*B_1\sim+B_2)+BaC_4; \\ \text{ес.AdC.Conceive} &\leq *AdC_1+AdC_2*AdC_3\sim+AdC_4; \end{aligned}$$

$$\begin{aligned} \text{ес.PrC.Design} &\leq *PrC_1*PrC_2*PrC_3; \\ \text{ес.BaC.Design} &\leq *BaC_1+BaC_2*BaC_3; \\ \text{ес.AdC.Design} &\leq *AdC_1\sim*AdC_2 (*A_1+A_2)\sim*AdC_3; \end{aligned}$$

$$\begin{aligned} \text{ес.PrC.Implement} &\leq *PrC_1*PrC_2*PrC_3*PrC_4 (*P_1*P_2)*PrC_5; \\ \text{ес.BaC.Implement} &\leq *BaC_1*BaC_2*BaC_3 (*B_1\sim+B_2)\sim+BaC_4; \\ \text{ес.AdC.Implement} &\leq *AdC_1*AdC_2(*A_1+A_2)*AdC_3+AdC_4; \end{aligned}$$

$$\begin{aligned} \text{ес.PrC.Operate} &\leq *PrC_1*PrC_2; \\ \text{ес.BaC.Operate} &\leq *BaC_1(*B_1*B_2)*BaC_2+BaC_3; \\ \text{ес.AdC.Operate} &\leq *AdC_1 (+A_1\sim+A_2)*AdC_2+AdC_3; \end{aligned}$$

Следует отметить, что в приведенных выражениях компетенций, все *профессиональные* компетенции являются обязательными для усвоения.

Выражения *базовых* и *дополнительных* компетенций, кроме обязательных, могут включать необязательные и альтернативные компетенции. Во всех выражениях компетенций может присутствовать второй уровень (в круглых скобках), из конкретизирующих компетенций, имеющих простой формат обозначений. В этом случае, компетенция первого уровня играет роль контекста для конкретизирующих компетенций.

Каждая из компетенций подвергается формализации выражением из опорных понятий, усвоение которых необходимо и достаточно для решения практических учебно-производственных задач соответствующего этапа CDIO разработки проекта.

Пусть, профессиональную компетенцию PrC_1 этапа *Conceive*, можно представить следующим выражением опорных понятий:

$$PrC_1 \leq *C_1 * C_2 (*C_1 * C_2) * C_3 * C_4; \quad (3)$$

В свою очередь, каждое из опорных понятий профессиональной компетенции PrC_1 представляется выражением знания, включающим обязательные, необязательные, либо альтернативные понятия первого и второго уровней конкретизации, на которых определены отношения «композиция», «агрегация» и «альтернативный выбор».

Например, опорное понятие C_1 профессиональной компетенции PrC_1 , можно идентифицировать некоторым набором понятий предметной области, и специфицировать следующим выражением знания:

$$C_1 \leq *C_{1.1} * C_{1.2} (*C_1 \sim + C_2) \sim *C_{1.3} * C_{1.4} (*C_1 * C_2 + C_3); \quad (4)$$

В завершении, рассмотрим компетентностные модели этапов CDIO, связанные с разработкой проекта "Банковская система типа клиент-сервер", и которые были использованы при проектировании онтологий опорных понятий учебного контента дисциплины "Технология разработки распределенных приложений", специальности "Вычислительная техника и программное обеспечение".

Согласно требований инициативы CDIO, любой технологический процесс выполнения проекта представляется последовательностью четырех обязательных этапов:

Conceive - проектный менеджмент разработки и производства продуктов;

Design - проектирование продуктов инженерной деятельности на дисциплинарной и междисциплинарной основе;

Implement - производство продуктов инженерной деятельности, в том числе аппаратуры и программного обеспечения, их интеграция, проверка, испытание и сертификация;

Operate - применение продуктов инженерной деятельности, управление их жизненным циклом и утилизацию.

Продemonстрируем пример представления компетентностной модели этапа *Conceive*, знания которой необходимы для успешного выполнения работ по проекту "Банковская система типа клиент-сервер". Напомним, что компетентностная модель является композицией выражений *профессиональных, базовых* и *дополнительных* компетенций.

Выражение профессиональных компетенций:

$$\text{ec.PrC.Conceive} \leq *PrC_1 * PrC_2 * PrC_3; \quad (5)$$

где PrC_1 – профессиональные компетенции, связанные с Инженерией предметной области, включающие два опорных понятия: C_1 - Инженерия требований, C_2 - Моделирование предметной области;

PrC_2 – профессиональные компетенции, связанные с Динамическим моделированием аспектов системы, включающие два опорных понятия: C_1 - Разбиение на объекты, C_2 - Динамическое моделирование системы;

$*PrC_3$ – профессиональные компетенции, связанные с Технологиями параллельных и распределенных систем, включающие три опорных понятия: C_1 – Поддержка исполнения в мультипрограммной и мультипроцессорных средах, C_2 - Планирование задач, C_3 - Технологии клиент-серверных и распределенных систем.

Выражение базовых компетенций:

$$\text{ec.VaC.Conceive} \leq *VaC_1 * VaC_2; \quad (6)$$

где VaC_1 - базовые компетенции, связанные с Моделированием видов деятельности, включающие два опорных понятия: C_1 - Анализ предметной области, C_2 - Моделирование видов деятельности;

VaC_2 - базовые компетенции, связанные с Моделью пользовательского интерфейса, включающие два опорных понятия: C_1 - Моделирование поведения объектов, C_2 - Концептуальная модель пользовательского интерфейса.

Выражение дополнительных компетенций:

$$\text{ec.AdC.Conceive} \leq *AdC_1 * AdC_2 + AdC_3; \quad (7)$$

где AdC_1 – дополнительные компетенции, связанные с Системным анализом бизнес-процессов, включающие два опорных понятия: C_1 - Анализ и реинжиниринг бизнес-процессов, C_2 - Системный анализ и проектирование;

AdC_2 - дополнительные компетенции, связанные с Аттестацией и управлением требованиями, включающие два опорных понятия: C_1 - Аттестация требований, C_2 - Управление требованиями;

AdC_3 - дополнительные компетенции, связанные с Оценкой стоимости программного проекта, включающие одно опорное понятие: C_1 - Оценка стоимости программного проекта.

5 Пример сценария для проекта «Банковская система типа клиент-сервер»

Допустим, обучающийся желает приобрести компетенции, необходимые ему для работы над проектами, связанными с технологиями разработки систем реального времени, параллельных и распределенных приложений.

В качестве предметной области, где рассматриваются соответствующие технологии, был использован контент монографии [11]. Для проекта «Банковская система типа клиент-сервер» и соответствующих этапов CDIO, разработаны компетентностные модели, в рамках каждой из которых определены опорные понятия *профессиональных, базовых и дополнительных* компетенций, построены онтологии опорных понятий и проведена их спецификация в виде выражений знаний.

Пусть обучающийся выбрал уровень профессиональных компетенций этапа *Conceive*, проекта «Банковская система типа клиент-сервер».

Как было отмечено, начальным этапом проектирования образовательной программы является разработка ее знаниевой архитектуры, в виде сценария последовательности взаимодействия знаниевых компонент. С учетом требований обучающегося и используя введенные формализмы спецификации знаниевых компонент, приведем пример Сценария в виде следующего выражения из четырех знаниевых компонент:

$$\text{Сценарий} \leq *KC_{10} * KC_{24} \sim *KC_3 + KC_8; \quad (8)$$

Выражение (8) имплицитно утверждает о том, что исполнение сценария должно обязательно включать реализацию знаниевых компонент KC_{10} и KC_{24} , при этом, реализация знаниевого компонента KC_3 является альтернативной относительно знаниевого компонента KC_{24} , а реализация знаниевого компонента KC_8 – дополнительной в сценарии обучения.

Далее, следует этап конфигурирования знаниевых компонент выражения (8), соответствующими параметрами *smart-контракта*.

Например, рассмотрим знаниевый компонент KC_{10} , для которого:

- предисловием P , являются утверждения, в виде сигнатур PrC_1, PrC_5, PrC_{12} – профессиональных компетенций, BaC_2, BaC_3 – базовых и, возможно, AdC_7, AdC_{10} и AdC_{11} – дополнительных компетенций, необходимых для успешного освоения знаниевого компонента KC_{10} ;

- постусловием Q – являются утверждения, в виде сигнатур PrC_{11}, PrC_{14} – профессиональных компетенций, которые гарантируются обучающему после успешного освоения материала знаниевого компонента KC_{10} ;

- Lev – требование обучающегося к уровню компетентности этапа, который он желает освоить. Например, обучающийся выбрал уровень профессиональных компетенций этапа *Conceive*, проекта «Банковская система типа клиент-сервер»;

- Col – параметр, задающий ссылки на знаниевые компоненты KC_1, KC_{23}, KC_{17} , образовательной среды, которые обязательно задействуются при реализации знаниевого компонента KC_{10} .

Таким образом, введенные формализмы представления знаний, в виде знаниевых компонент, и возможности по их конфигурированию, в соответствии с сценариями обучения с помощью параметров *smart-контракта*, позволяют создавать уникальный семантический контекст знаниевой архитектуры образовательной программы, удовлетворяющий профессиональным потребностям обучающихся.

Изложенные концепции и механизмы конфигурирования знаниевых компонент проектно-ориентированной технологии обучения, положены в основу разрабатываемой образовательной среды, в виде *Web-приложения*, с адаптивной кросс-браузерной версткой, базой данных и возможностью интеграции с другими системами.

Заключение

Таким образом, в данной статье, согласно задач этапов грантового проекта АР 05134973 «Исследование и разработка моделей и методики представления и организации знаний с применением онтологического подхода, и инструментальных средств *Smart-технологии*, при реализации образовательных программ и процессов», выполняемого под научным руководством к.т.н., доцента

ВАК РК Кубекова Б.С., приведены результаты исследований, научная новизна которых заключается в том, что впервые в образовательной практике:

- предложена и реализуется проектно-компетентностная парадигма и основанная на ней инновационная модель отображения знаний предметной области;

- на базе инновационной модели отображения знаний, разработана методика формирования и реализации знаниевых компонент, которую можно признать *образовательной технологией*, поскольку методика предусматривает использование:

- концептуальных основ проектно-ориентированной технологии обучения, на базе стандартов и принципов всемирной инициативы CDIO, эпистемологической функции компетентностного подхода уровневого инженерного образования, онтологического инжиниринга, компонентного подхода и инженерии предметной области;

- методов представления и формализации знаниевых компонент и использование их для проектирования знаниевой архитектуры образовательной программы, на основе smart-контракта;

- семантического анализа предметной области, с целью выделения ключевых понятий домена обучения и предоставления возможности конфигурирования знаниевых компонент, в соответствии с требуемым сценарием и результатами обучения, а также с целью воспроизводимости и повторного использования знаниевых компонент;

- навыков и рациональных приемов «сжатия» информации и ее когнитивно-графического представления;

- возможности предоставления преподавателям и квалифицированным ИТ-специалистам удобного и доступного инструментального средства для создания и визуального представления онтологий понятий предметной области, их редактирования и использования в простой и наглядной форме;

- возможности формирования повторно используемых знаниевых компонент, их конструирования и адаптации для **широкого спектра образовательных программ и услуг, что будет способствовать укреплению конкурентоспособности учебного заведения на рынке образовательных услуг.**

Практическая значимость исследований заключается в том, что представленные материалы исследований и методические разработки могут быть использованы в системе высшего образования, в условиях расширяющейся академической автономности как при разработке образовательных программ и практических рекомендаций Smart-обучения, так и в структуризации и оптимизации контента учебных планов и курсов, что будет способствовать повышению эффективности обучения, самообразования и саморазвития обучающихся.

Литература

1. Kubekov B., Utebergenov I. The use of multiparadigm approach to knowledge modeling.- // 7th International Conference on Education and New Learning Technologies".- Barcelona (Spain), 6th-8th of July 2015.- pp.136-141
2. Kubekov B., Kuandykova J., Utebergenov I., Utegenova A. Application of the conceptual model of knowledge for formalization of concepts of educational content /9th International Conference on Application of Information and Communication Technologies AICT2015.- Rostov-on-Don, 14-16 October 2015.- pp.294-306
3. Kubekov B., Utegenova A., Naumenko V. Applying of ontological engineering to represent knowledge and training sessions. 10th International Conference

on Application of Information and Communication Technologies -AICT2016, Baku (Azerbaijan), 12-14 October 2016, pp.115-118.

4. Kubekov B. Educational components formation technology for the planned CDIO SYLLABUS education. 9th annual International Conference of Education, Research and Innovation - ICERI2016, Seville (Spain), 14-16 November, 2016, pp.6139-6145

5. Kubekov B.S., Ditmar Beyer, Utegenova A.U., Zhaksybaeva N.N. Innovative paradigm of education of knowledge - competency form based on ontology. Journal of theoretical and applied information technology 15th November 2017, Vol.95. №21, 2005-ongoing JATIT@LLS (E-ISSN 1817-3195 / ISSN 1992-8645). pp. 5859-5868.

6. Software Engineering 2014: Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering. A Volume of the Computing Curricula Series. IEEE CS and ACM, 2015. 134p.

7. Crowley E.F, CDIO Program: Description of the aims and objectives of bachelor's engineering education, CDIO Report No. 1 ed. MIT, 2001. - Access mode: <http://www.cdio.org>

8. Чучалин А.И. Модернизация бакалавриата в области техники и технологий, с учетом международных стандартов инженерного образования // Высшее образование в России. 2011. №10.

9. Чучалин А.И. О применении подхода CDIO для проектирования уровневых программ инженерного образования// Высшее образование в России. 2016. №4(200). с. 17-32

10. Гома Хассан UML. Проектирование систем реального времени, параллельных и распределенных приложений: Пер. с англ.-М.: ДМК Пресс, 2014. - с.704.

РАЗРАБОТКА ЧИСЛЕННОГО МЕТОДА УЧЕТА НАЛИЧИЯ ЛОКАЛЬНЫХ ПОВЕРХНОСТНЫХ ТЕПЛООБМЕНОВ В СТЕРЖНЯХ ПЕРЕМЕННОГО СЕЧЕНИЯ

**Кудайкулов А.К.¹, Ташев А.А.¹, Аршидинова М.Т.^{1,2},
Бегалиева К.Б.^{1,2}**

e-mail: kudaykulov2006@mail.ru, azattash@mail.ru, mukaddas_arshidi@mail.ru

¹Институт информационных и вычислительных технологий КН МОН РК

²Казахский Национальный Университет имени аль-Фараби, Казахстан

Аннотация. Рассматривается установившееся термомеханическое состояние стержня ограниченной длины переменного сечения при одновременном наличии локальных поверхностных теплообменов, теплоизоляции и тепловых потоков. При исследовании стержня используются фундаментальные законы сохранения энергии и квадратичные сплайн функции в местной системе координат.

Ключевые слова. Удлинение, осевая сила, сечения, температура, перемещение, деформация, напряжение.

1. Введение

Рассматривается горизонтальный стержень ограниченной длины L [см] и переменного сечения $F(x)$ [см²]. Радиус сечения стержня меняется линейно по его длине, т.е. $r = ax + b$ [см], ($0 \leq x \leq L$), где a и b постоянные. Площадь поперечного сечения стержня меняется нелинейно по длине стержня следующим образом $F(x) = \pi r^2 = \pi(a^2 x^2 + 2abx + b^2)$ [см²], ($0 \leq x \leq L$).

Физико-механические свойства материала стержня характеризуется коэффициентами теплопроводности $K_{xx} \left[\frac{\text{watt}}{\text{см} \cdot ^\circ\text{K}} \right]$, и теплового расширения $\alpha \left[\frac{1}{^\circ\text{K}} \right]$. Модулем упругости материала стержня $E \left[\frac{\text{кг}}{\text{см}^2} \right]$. Предположим, что на площадь поперечного сечения левого конца стержня подведен тепловой поток постоянной интенсивности $q_1 \left[\frac{\text{watt}}{\text{см}^2} \right]$. В это время на площадь поперечного сечения правого конца стержня подводится тепловой поток интенсивностью $q_2 \left[\frac{\text{watt}}{\text{см}^2} \right]$. Локальные боковые поверхности ($0 \leq x \leq L_1$) и ($L_2 \leq x \leq L$) исследуемого стержня теплоизолированы. Через локальную боковую поверхность ($L_1 \leq x \leq L_2$) стержня происходит теплообмен с окружающей этой поверхности средой. При этом коэффициент теплообмена $h \left[\frac{\text{watt}}{\text{см}^2 \cdot ^\circ\text{K}} \right]$, а температура окружающей среды $T_{oc} [^\circ\text{K}]$, необходимо определить поле распределения температуры, трех составляющих деформаций, напряжений и перемещения. Расчетная схема задачи приводится на рисунке 1.

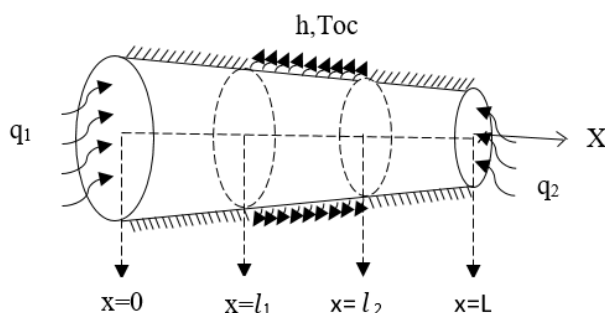


Рис.1. Расчетная схема задачи

2. Построение квадратичной сплайн функции в локальной системе координат

Исследуемый стержень переменного сечения и ограниченной длины дискретизируем элементами длиной $l=L/n$ [см] [1,4]. При этом n – число дискретных элементов в стержне. Здесь, естественно $l \ll L$. Рассмотрим поле распределения температуры по длине одного дискретного элемента длиной l [см], рисунок 2.

В локальной системе координат ($0 \leq x \leq l$) сечение с координатой $x=0$, обозначим через i . Аналогично, сечение с координатой $x=l/2$, обозначим через j .

Наконец, сечение с координатой $x=l$, обозначим через k . Значения температуры в узлах i, j, k обозначим соответственно T_i, T_j и T_k . Закон распределения температуры по длине одного дискретного элемента аппроксимируем полным полиномом второго порядка, т.е.

$$T(x) = C_1 x^2 + C_2 x + C_3, (0 \leq x \leq l) \quad (1)$$

здесь C_1, C_2 и C_3 постоянные.

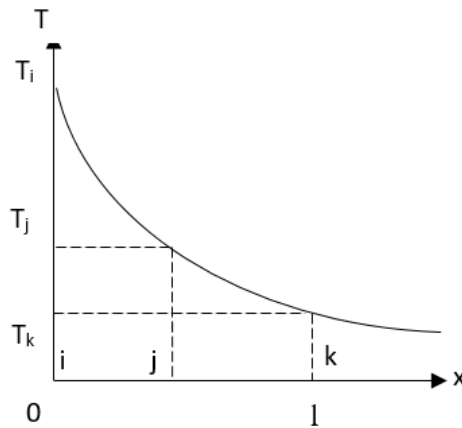


Рис.2. Поле распределения температуры по длине одного дискретного элемента ($0 \leq x \leq l$)

При вышепринятых обозначениях можно построить следующую систему для определения значений постоянных C_1, C_2 и C_3 .

$$\begin{cases} T(x=0) = C_1 \cdot 0 + C_2 \cdot 0 + C_3 = T_i \\ T(x = \frac{l}{2}) = C_1 \cdot \frac{l^2}{4} + C_2 \cdot \frac{l}{2} + C_3 = T_j \\ T(x=l) = C_1 l^2 + C_2 \cdot l + C_3 = T_k \end{cases} \quad (2)$$

Отсюда находим, что

$$C_3 = T_i, C_2 = \frac{4T_j - T_k - 3T_i}{l}, C_1 = \frac{2}{l^2} (T_i - 2T_j + T_k) \quad (3)$$

Далее подставляя (3) формулу в (1), получим

$$T(x) = \frac{2x^2 - 3lx + l^2}{l^2} T_i + \frac{4lx - 4x^2}{l^2} T_j + \frac{2x^2 - lx}{l^2} T_k, (0 \leq x \leq l) \quad (4)$$

Здесь введем следующие обозначения:

$$\varphi_i(x) = \frac{2x^2 - 3lx + l^2}{l^2}; \varphi_j(x) = \frac{4lx - 4x^2}{l^2}; \varphi_k(x) = \frac{2x^2 - lx}{l^2}; (0 \leq x \leq l) \quad (5)$$

Эти функции назовем квадратичными сплайн функциями в местной системе координат $(0 \leq x \leq l)$ [2].

$$T(x) = \varphi_i(x)T_i + \varphi_j(x)T_j + \varphi_k(x)T_k, (0 \leq x \leq l) \quad (6)$$

Выражение градиента температуры в пределах длины одного дискретного элемента имеет следующий вид:

$$\begin{aligned} \frac{\partial T}{\partial x} &= \frac{\partial \varphi_i(x)}{\partial x} T_i + \frac{\partial \varphi_j(x)}{\partial x} T_j + \frac{\partial \varphi_k(x)}{\partial x} T_k = \\ &= \frac{4x - 3l}{l^2} T_i + \frac{4l - 8x}{l^2} T_j + \frac{4x - l}{l^2} T_k, \end{aligned} \quad (0 \leq x \leq l) \quad (7)$$

Аналогично (8), поле перемещения в пределах длины одного дискретного элемента будет иметь следующий вид

$$U(x) = \varphi_i(x)U_i + \varphi_j(x)U_j + \varphi_k(x)U_k, (0 \leq x \leq l) \quad (8)$$

где $U_i, U_j, U_k \dots$ - перемещение сечений i, j, k .

3. Построение функционалов полной тепловой энергии для дискретных элементов стержня переменного сечения

Рассмотрим дискретные элементы исследуемого стержня переменного сечения с лева на право. Сначала рассмотрим первый дискретный элемент, с левого конца стержня (рисунок 3).

Для этого элемента функционал полной тепловой энергии имеет следующий вид [6-9]

$$J_1 = \int_{F(x=0)} q_1 T ds + \int_{V_1} \frac{K_{xx}}{2} \left(\frac{\partial T}{\partial x} \right)^2 dv, (0 \leq x \leq l) \quad (9)$$

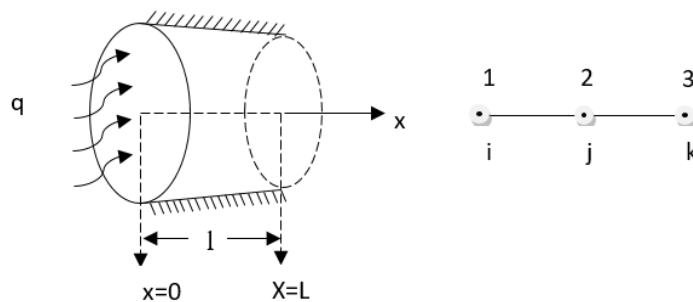


Рис.3. Первый дискретный элемент исследуемого стержня

Для первого дискретного элемента к локальным номерам i, j, k соответствуют глобальная нумерация узлов 1, 2 и 3.

Тогда, для первого дискретного элемента исследуемого стержня имеет место

$$T_i = T_1; T_j = T_2; T_k = T_3;$$

Теперь рассмотрим второй дискретный элемент, через боковую поверхность которого происходит теплообмен с окружающей средой (рисунок 4).

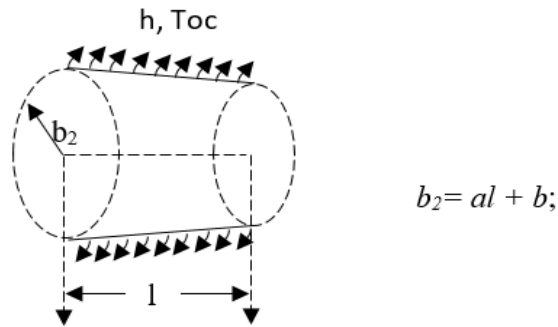


Рис.4. Второй дискретный элемент

Функционал полной тепловой энергии для второго дискретного элемента будет следующим [12]

$$J_2 = \int_{V_2} \frac{K_{xx}}{2} \left(\frac{\partial T}{\partial x} \right)^2 dv + \int_{S_{\text{бон}2}} \frac{h}{2} (T - T_{oc})^2 ds \quad (10)$$

где V_2 – объем второго дискретного элемента, $S_{\text{бон}2}$ – площадь боковой поверхности второго дискретного элемента.

Теперь переходим к третьему дискретному элементу. Боковая поверхность этого элемента теплоизолирована. На площадь поперечного сечения левого конца подведем тепловой поток интенсивностью q_2 (Рисунок 5).

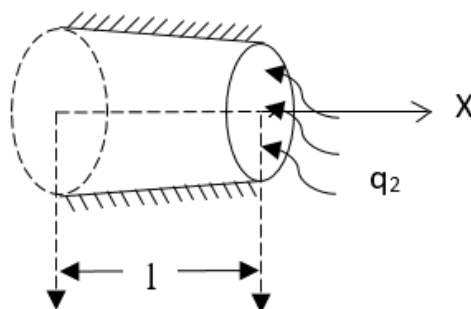


Рис.5. Третий дискретный элемент

Для третьего дискретного элемента функционал полной тепловой энергии имеет следующий вид

$$J_3 = \int_{V_3} \frac{K_{xx}}{2} \left(\frac{\partial T}{\partial x} \right)^2 dv + \int_{F(x=L)} q_2 T ds \quad (11)$$

где V_3 – объем второго дискретного элемента, $F(x=L)$ – площадь поперечного сечения правого конца стержня. $F(x=L) = F(x=3l) = \pi(al + b)^2 = \pi(3al + b)^2$.

Тогда функционал полной тепловой энергии для исследуемого стержня по узловым значениям температур T_i , ($i=1-7$), построим разрешающую систему линейных алгебраических уравнений с учетом естественных граничных условий.

$$J = J_1 + J_2 + J_3 \quad (12)$$

Далее минимизируя этот функционал по узловым значениям температур T_i , ($i=1-7$), построим разрешающую систему линейных алгебраических уравнений с учетом естественных граничных условий [10].

Решая эту систему, определяются узловые значения температуры T_i , ($i=1-7$). По ним строится закон распределения температуры по длине исследуемого стержня следующим образом:

Для первого участка стержня

$$T(x) = \varphi_i(x)T_1 + \varphi_j(x)T_2 + \varphi_k(x)T_3, \quad (0 \leq x \leq l) \quad (13)$$

Для второго участка стержня

$$T(x) = \varphi_i(x)T_3 + \varphi_j(x)T_4 + \varphi_k(x)T_5, \quad (0 \leq x \leq l) \quad (14)$$

Наконец для третьего участка стержня имеем

$$T(x) = \varphi_i(x)T_5 + \varphi_j(x)T_6 + \varphi_k(x)T_7, \quad (0 \leq x \leq l) \quad (15)$$

4. Определение поля напряжений и деформаций

Если один конец рассматриваемого горизонтально расположенного стержня переменного сечения жестко зашпемлен, то он из-за наличия поля температуры удлинится. Величина удлинения определяется в соответствии общих законов теплофизики

$$\Delta l_T = \int_0^L \alpha T(x) dx \quad (16)$$

где $\alpha \left[\frac{1}{^\circ K} \right]$ – коэффициент теплового расширения материала стержня. $T(x)$ – закон распределения температуры по длине стержня.

Если оба конца исследуемого стержня жестко зашпемлены, то в нем возникает термо-напряженно-деформированное состояние, а также осевая сжимающая сила R [кГ].

В случае зашпемления двух концов исследуемого стержня, он не удлинится и не укорачивается. Но возникает осевое сжимающее усилие R (Рисунок 6).

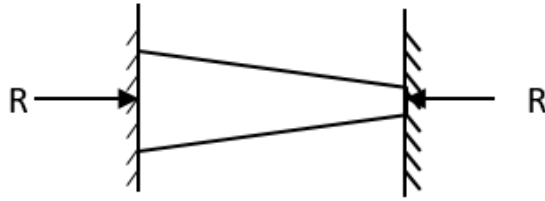


Рис.6. Стержень переменного сечения под воздействием осевого сжимающего усилия R

Для определения величины возникающего осевого усилия R , находим среднюю площадь поперечного сечения

$$F_{\text{ср}} = \frac{\int_0^L F(x) dx}{L} = \frac{\int_0^L (a^2 x^2 + 2abx + b^2) dx}{L} = \left(\frac{a^2 L^2}{3} + abL + b^2 \right) \quad (17)$$

где $F_{\text{ср}} [\text{см}^2]$

5. Определение поля перемещения

По длине исследуемого стержня переменного сечения обозначим $U(x)$. Поле распределения этого перемещения по длине одного дискретного элемента в локальной системе координат аппроксимируем квадратичными сплайн функциями.

$$U(x) = \varphi_i(x)U_i + \varphi_j(x)U_j + \varphi_k(x)U_k = \frac{2x^2 - 3lx + l^2}{l^2}U_i + \frac{4lx - 4x^2}{l^2}U_j + \frac{2x^2 - lx}{l^2}U_k, \quad (0 \leq x \leq l) \quad (18)$$

где в локальной системе координат $U_i = U(x=0)$; $U_j = U(x=\frac{l}{2})$; $U_k = U(x=l)$;

Тогда в этой локальной системе координат закон распределения упругой составляющей деформации $\varepsilon_x(x)$ определяется в соответствии соотношения Коши

$$\varepsilon_x(x) = \frac{\partial U}{\partial x} = \frac{4x - 3l}{l^2}U_i + \frac{4l - 8x}{l^2}U_j + \frac{4x - l}{l^2}U_k, \quad (0 \leq x \leq l) \quad (19)$$

6. Численное решение задачи

Для иллюстрации вышеописанного метода и вычислительного алгоритма рассмотрим задачу со следующими исходными данными:

$$L = 30 \text{ см}; \quad l = \frac{L}{3} = 10 \text{ см}; \quad a = -\frac{1}{30}; \quad b = 2 \text{ см}; \quad b_2 = \frac{5}{3} \text{ см}; \quad b_3 = \frac{4}{3} \text{ см};$$

$$K_{xx} = 100 \frac{\text{W}}{\text{см}^{\circ}\text{C}}; \quad \alpha = 125 \times 10^{-7} \frac{1}{^{\circ}\text{C}}; \quad E = 2 \cdot 10^6 \frac{\text{кГ}}{\text{см}^2};$$

$$q_1 = q_2 = -500 \frac{\text{W}}{\text{см}^2}; \quad h = 10 \frac{\text{W}}{\text{см}^2 \cdot ^{\circ}\text{C}}; \quad T_{oc} = 30^{\circ}\text{C};$$

При этих исходных данных видно, что исследуемый стержень дискретизируется тремя дискретными элементами одинаковой длины. При этом

боковые поверхности первого и третьего дискретных элементов полностью теплоизолированы. В то время через боковую поверхность второго элемента происходит теплообмен с окружающей средой. На площадь поперечных сечений двух концов стержня подведены тепловые потоки одинаковой интенсивности.

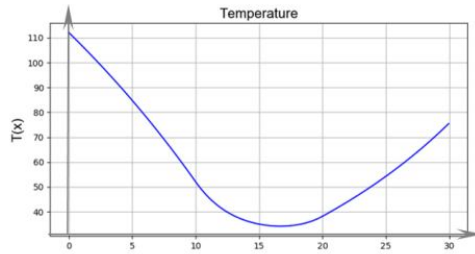


Рис 7. Зависимости температуры T по длине стержня

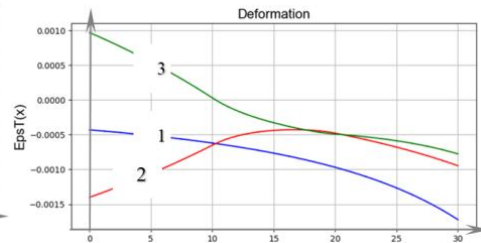


Рис 8. Зависимости деформации по длине стержня
1 – $\varepsilon(x)$; 3 – $\varepsilon_x(x)$; 2 – $\varepsilon_T(x)$

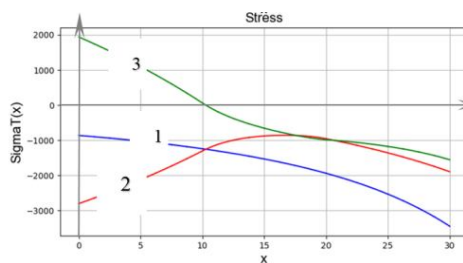


Рис 9. Зависимости напряжении по длине стержня
1 – $\sigma(x)$; 3 – $\sigma_x(x)$; 2 – $\sigma_T(x)$

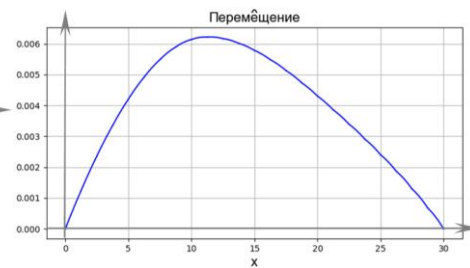


Рис10. Зависимости перемещении по длине стержня

7. Анализ полученных результатов

На рисунке 7 приводится закон распределения температуры по длине исследуемого стержня переменного сечения. При этом значения температуры на левом конце стержня $T(x=0) = 112,06 \text{ }^\circ\text{C}$. В то время на правом конце $T(x=30 \text{ см}) = 75,69 \text{ }^\circ\text{C}$. Эта разница мотивируется тем, что площадь поперечного сечения левого конца $F(x=0) = 4\pi \text{ см}^2$, а площадь на правом конце $F(x=30 \text{ см}) = \pi \text{ см}^2$. Таким образом, $F(x=0)/F(x=30 \text{ см}) = 4$. Это означает, что на левом конце $q \cdot F(x=0) = -500 \cdot 4\pi = -2000 \pi(w)$. В то время $q \cdot F(x=30 \text{ см}) = -500 \cdot \pi(w)$. Поэтому значение температуры на левом конце превышает в 1,49 раза значения температуры на правом конце исследуемого стержня переменного сечения. В следствии теплообмена с окружающей средой серединной части 1/3 части стержня значение температуры в сечении, координата которого $x=16,87 \text{ см}$ будет минимальным, то есть $T(x=16,87 \text{ см}) = 34,238 \text{ }^\circ\text{C}$. Кроме того, это обусловлена еще тем что радиус сечения стержня слева в право. Если левый конец стержня жестко защемлена, а правый конец свободен, то исследуемый стержень переменного сечения удлинится. Величина удлинения при принятых исходных данных будет

$$\Delta l_T = \int_0^L \alpha T(x) dx = 0,02214 \text{ см}$$

Если оба конца стержня жестко заземлены, то он не может удлиняться. В этом случае из-за теплового расширения материала возникает осевое сжимающее усилие $R[\text{кг}]$. При наших исходных данных значение этого усилия будет $R = -10820,8148 \text{ кг}$.

Тогда в этом случае в стержне возникает установившееся термо-напряженно-деформированное состояние. На рисунке 8 приводятся законы распределения трех составляющих деформации. Здесь кривая I – это закон распределения термоупругой составляющей деформаций. Она по всей длине стержня имеет сжимающий нелинейный характер. При этом значение $\varepsilon(x=0) = -0,00043$, а $\varepsilon(x=30\text{см}) = 0,001722$. Это означает, что значение $\varepsilon(x)$ на левом конце стержня 4 раза меньше, чем на правом. Этот процесс обусловлено тем, что площадь поперечного сечения левого конца стержня 4 раза больше, чем на правом. Закон распределения температурной составляющей температуры $\varepsilon_T(x)$ по всей длине исследуемого стержня переменного сечения имеет сжимающий и нелинейный характер. При этом его максимальное значение наблюдается на левом конце стержня $\varepsilon_T(x) = -0,0014$. Наблюдается монотонное уменьшение $\varepsilon_T(x)$ до сечения стержня координата которого $x=16,25$ см. Здесь $\varepsilon_T(x=16,25) = -0,00042854$. Это обусловлено теплообменом через боковой поверхности срединного участка стержня с окружающей средой. Далее $\varepsilon_T(x)$ имеет опять слабо возрастающую природу, и $\varepsilon_T(x=30\text{см}) = -0,000964619$. В законе распределения $\varepsilon_T(x)$ выявлено, что $\varepsilon_T(x=0)/\varepsilon_T(x=L=30\text{см}) = 1,4804$ раза. Это обусловлено тем, что площадь поперечного сечения левого конца исследуемого стержня 4 раза больше, чем правого. В отличие от других закон распределения упругих составляющих деформаций по длине исследуемого стержня переменного сечения имеет знакопеременный характер. В частности, $\varepsilon_x(x)$ на участке $0 \leq x \leq 10$ см стержня имеет растягивающий характер. На остальном $10 \leq x \leq L = 30$ см участке стержня она имеет так же сжимающий характер. В общем случае $\varepsilon_x(x)$ так же имеет нелинейный вид. На рисунке – 9 приводится закон распределения термоупругих $\sigma(x)$ и $\sigma_T(x)$ и упругих $\sigma_x(x)$ составляющих по длине исследуемого стержня переменного сечения.

Из рисунка 8-9 видно, что они пропорциональны к соответствующим составляющим деформации. При этом коэффициентом пропорциональности является модуль упругости материала стержня $E = 2 \cdot 10^6 \left[\frac{\text{кг}}{\text{см}^2} \right]$. Наконец на рисунке -10 приводится поле распределения перемещения сечений исследуемого стержня переменного сечения. Из рисунка видно, что все сечения перемещаются слева вправо по направлению оси Ox . Это обусловлено тем, что площадь поперечного сечения левого конца стержня куда подводится тепловой поток постоянной интенсивности q , 4 раза больше, чем правый. Наибольшая амплитуда перемещения соответствует к сечению стержня, координата которого $x=11,25$ см. Это обусловлено возникновением большой температуры $T(x=0) = 112,059 \text{ }^\circ\text{C}$ на левом конце исследуемого стержня переменного сечения.

8. Выводы

На основе фундаментальных законов сохранения энергии в сочетании квадратичных сплайн функций разработаны метод, вычислительный алгоритм и программное обеспечение на Python для полного исследования установившегося

термо-напряженно-деформированного состояния стержня переменного сечения при одновременном наличии локальных тепловых потоков, теплообмена и теплоизоляции. При конкретных исходных данных вычислены величины удлинения стержня в случае свободного одного конца. В случае заземления двух концов стержня определены законы распределения температуры, термоупругих температурных и упругих составляющих деформаций, напряжений, а также перемещения. Вычислена величина возникающего осевого усилия. Выявлены соответствующие закономерности, и они так же обусловлены физическими законами.

Все полученные решения удовлетворяют фундаментальные законы сохранения энергии, поэтому отличаются высокой точностью. Предполагается, что разработанные методы, вычислительный алгоритм и пакет прикладных программ на Python будут полезными при проектировании энергетических установок, двигателей внутреннего сгорания, реактивных двигателей.

Работа выполнена при поддержке Министерства образования и науки Казахстана в рамках грантового финансирования научных исследований на 2018-2020 годы по проекту № AP05131093 «Математическое и компьютерное моделирование нелинейных термо-физических процессов в стержнях переменного сечения»

Литература

1. Larry J. Segerlind. Applied Finite Element Analysis. - Jhon Willey and Sons, Inc. New York/London/Sydney/Toronto, 1976.- 392p.
2. Carslaw H.S., Jaeger J.C. Conduction of Heat in Solids. Oxford University Press / London, 1986 -526 p.
3. Timoshenko S., Goodyear J.N. Theory of Elasticity. – McGRAW-Hill. Book. Company. Inc., 1987. – 567 p.
4. Zienkiewicz O.C. The method in Engineering science. Butterworth-Heinemann / Oxford-Auckland-Boston-Johannesburg-Melburne-New Delhi, 2000. – 690p.
5. Maugin, G.A. The saga of internal variables of state in continuum thermo-mechanics // Mechanics Research Communications - 2015. - P.79-80.
6. Tashenova, Z. Nurlybaeva, E. Kudaykulov, A. Method preparation and solution algorithm for resolving stationary problem of a rod under thermo - stressed condition restrained at both ends affected by heat exchange and heat flows //2012 International Conference on Advanced Material and Manufacturing Science, ICAMMS 2012; Beijing; China.
7. Kudaykulov A., Zhumadillayeva A. Numerical simulation of temperature distribution field in beam bulk in the simultaneous presens of heat insulation, heat flux and heat exchange // Actaphysicapolonica A, 2016. – P. 335-336.
8. Muheyat Niyazbek, Kuenssaule Talp, Kudaykulov A.K. Computer and Mathematical Modeling of Thermomechanical Processes in Element of Constructions //Acta Scientiarum Naturalium Universitatis Nankaiensis, Vol 50. № 3. 2017, Tianjin, China.
9. Kudaykulov A., Tashev, A.A., Askarova, A. A computational algorithm and the method of determining the temperature field along the length of the rod of variable cross section //Open Engineering, 8(1), 2018. – P. 170-175.
10. Krieth, F. Principles of Heat Transfer, 3-rd ed. Index Educational Publishers, N.Y. 1977.

11. Biot, M.A. Thermoelasticity and irreversible thermodynamics // Journal of Applied Physics, 27 (3), 1956.- P.240-253.
12. Presley, M.A., Christensen, P.R. Thermal conductivity measurements of particulate materials 1. // Journal of Geophysical Research E: Planets, 102 (E3). 1997. – P. 6535-6549.
13. Harr, M.E. Ground Water and Seepage / McGraw-Hill, N. Y.- 1962.- 315p.
14. Johnson, C. Numerical Solution of Partial Differential Equations by the Finite Element Method. /Cambridge University Press, Cambridge. – 1987. – 278p.
15. Gaspar, J.R., Moreira, M.L., Desampaio, P.A.B. Temperature Distribution Fuel Rods: A Study on the Effect of Eccentricity in the Position of UO₂ Pellets. // 0-th International Conference "Nuclear Energy for New Europe". – 2011.

РАЗРАБОТКА АЛГОРИТМА УПРАВЛЕНИЯ И ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ СОЛНЕЧНОГО КОТРОЛЛЕРА

**Кунелбаев М. М. Сундетов Т.Р., Айткулов Ж.,
Черикбаева Л. Ш., Орманов Т.А.**
murat7508@yandex.kz

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** В настоящей работе рассматривается алгоритм управления и программное обеспечение создания цифровой платформы для мониторинга за функционированием двухконтурного плоского солнечного коллектора с термосифонной циркуляцией.*

Введение

Использование возобновляемых источников энергии способствуют распространению постоянного увеличения спроса на энергию и на расширение участия экологически чистых технологий. Фототермическое преобразование в солнечных тепловых системах является одним из наиболее распространенных способов использования этой энергии [1,2]. Необходимость предоставления соответствующих рабочих параметров системы обеспечивает использование правильно выбранных и настроенных контроллеров. Солнечные контроллеры производятся многими компаниями и фирмами [3,4,5,6]. Несмотря на широкий спектр доступных решений, контроллеры по-прежнему в основном являются устройствами с замкнутой структурой и практически не имеют возможностей расширения. С другой стороны, гибкие решения, основанные на программируемых контроллерах (PLC Programmable Logic Controller), применимы только в больших системах из-за очень высокой стоимости ПЛК [7]. Однако в последнее время можно заметить растущий спрос на модульные контроллеры, полностью программируемые, которые позволяют интегрировать несколько систем в единую операционную систему. Не менее важным для пользователей представляется

реализация ряда дополнительных функций. Например, система управления и контроля через Интернет [9] или использование источника питания с фотоэлектрическими панелями [10, 11]. До настоящего времени реализации таких систем полагались только на ПЛК, однако из-за их высокой стоимости они редко используются в жилых установках. Хотя модульность и расширяемость уже предлагаются несколькими производителями солнечных контроллеров, например, [5] Однако программное обеспечение этих устройств все еще закрыто. С распространением недорогих платформ микроконтроллеров, таких как Arduino [8], стало возможным проектировать контроллеры в небольшой серии и адаптировать их программное обеспечение к потребностям конкретного клиента. Также стоит отметить, что такие контроллеры могут быть легко модернизированы при расширении поддерживаемой системы. Независимо от конструкции, солнечные контроллеры для нестандартных, сделанных на заказ установок должны соответствовать требованиям, определенным стандартом PN-EN 12977 [12, 13].

2. Методы исследование

Плоские солнечные коллекторы с термосифонной циркуляцией используются для преобразования падающего солнечного излучения в тепловую энергию. Эта энергия накапливается в виде ощутимого тепла в резервуаре для хранения жидкости и используется по мере необходимости для питания помещений и нагрева воды. На рисунке 1. показана двухконтурная гелиоустановка с термосифонной циркуляцией.

Работа предлагаемой установки осуществляется следующим образом. Солнечная энергия E с температурой t_0 поглощается гелиоколлектором 1, с температурой t_1 , нагревая поток солнечной энергии проходит через светопрозрачный изоляционный стеклопакет 2. Тепло, полученное от солнечного потока, нагревает жидкость в змеевиках 3, которая удаляется из коллектора, а на ее место поступает холодная вода из трубопровода с вентилем для холодной воды 8, а из сифона бака дозатора 7 происходит постоянная терма сифонная циркуляция с помощью циркуляционной трубы 10. Далее жидкость попадает в тепловой насос 11, который состоит из испарителя 12 конденсатора с температурой t_2 , в котором теплообменник выполнен в виде спирали, поглощая тепло теплоносителя, опускает его температуру ниже температуры атмосферного воздуха (Q_2) с помощью дросселирующего клапана 14, тем самым способствуя дополнительному поглощению тепла из атмосферного воздуха.

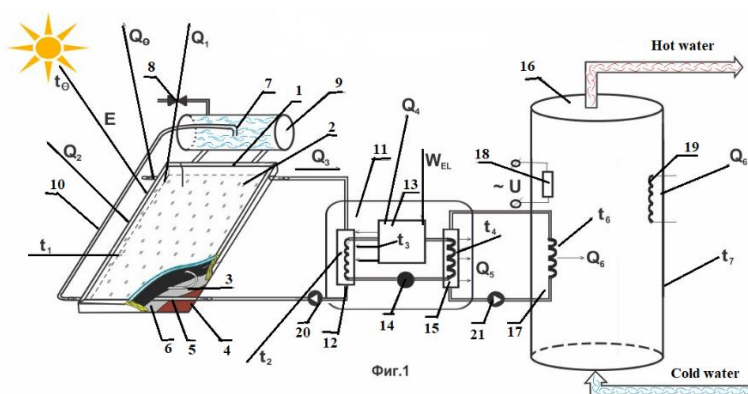


Рис. 1. Принципиальная схема двухконтурной гелиоустановки с термосифонной циркуляцией

В схеме также показано солнечное излучение, отраженное от полупрозрачного покрытия (Q_0) и поверхности поглощающей панели (Q_1). В тепловом насосе осуществляется передача энергии теплоносителя, с относительно низкой температурой, к теплоносителю теплообменника конденсатора 15 в виде спирали с более высокой температурой t_2 , которая увеличивает площадь, а также интенсивность теплообмена. Для осуществления такого цикла используется компрессор 13 с температурой t_3 , с электроприводом 17. Далее, посредством теплообменника конденсатора 15 с температурой t_4 , тепло от теплового насоса (Q_5) передается в бак аккумулятор теплообменника Q_6 с температурой t_6 системы отопления 18. Так как установка имеет два контура, она снабжена автоматическими циркуляционными насосами 19 и 20 для циркуляции жидкости между гелиоколлектором и испарителем, конденсатором и баком- аккумулятором.

Прототип модульного контроллера для солнечной тепловой системы

Контроллеры, предназначены для управления солнечными тепловыми установками, в частности, должны характеризоваться модульной структурой. Это позволяет обновить функциональность контроллера с дальнейшим развитием установки. В этой главе представлен пример прототипа модульного солнечного контроллера, основанного на свободно программируемой платформе Mojo v3 FPGA board. Блок-схема контроллера показана на (рис. 2). Прототип контроллера состоит из центрального блока, объединяющего все основные функции управления, и трех основных модулей расширения: системный мониторинг, резервное электропитание и метеостанция.

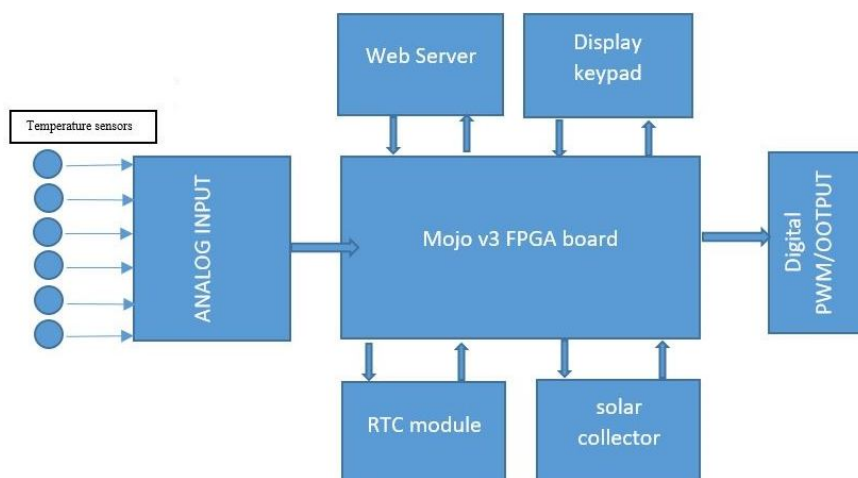


Рис. 2. Блок-схема контроллера управления двухконтурной гелиоустановки с термосифонной циркуляцией

Для управления двухконтурной гелиоустановки с термосифонной циркуляцией, была разработана функциональная программа Веб Сервера на рисунке 3.

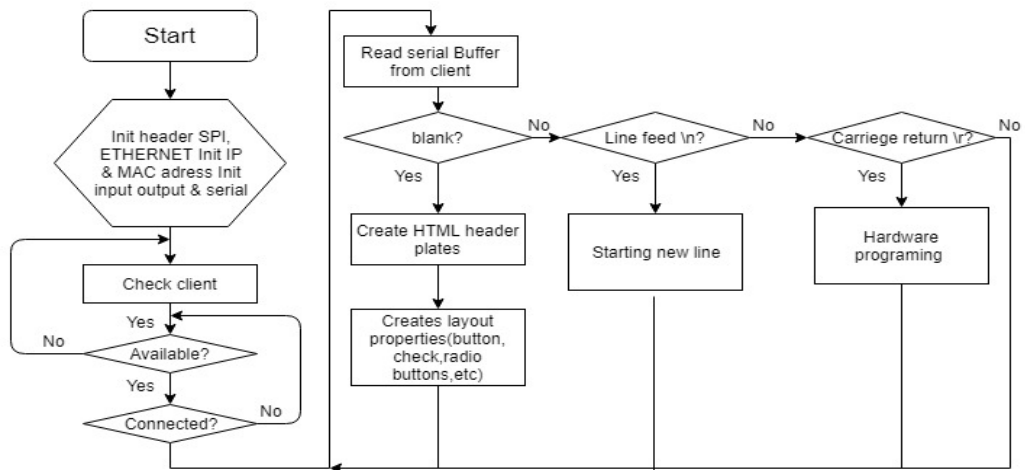


Рис. 3. Блок-схема Веб Сервера для контроллера управления двухконтурной гелиоустановки с термосиффонной циркуляцией

Как видно из рисунка 3 Веб Сервер, при начале работы инициализируют библиотеки SPI, Ethernet, а также активизируются IP и MAC адреса. Активизируются входные, выходные и последовательные данные. От инициализации библиотеки данных модуль Check Client соединит с клиентом. Если есть клиент тогда соединение устанавливается, если нет клиента тогда заново ищет клиента. После соединения открывается новый HTML файл и создается для управления контроллера слой, который имеет шесть температурных показателей и четыре кнопки управления клапанами и различные элементы для управления.

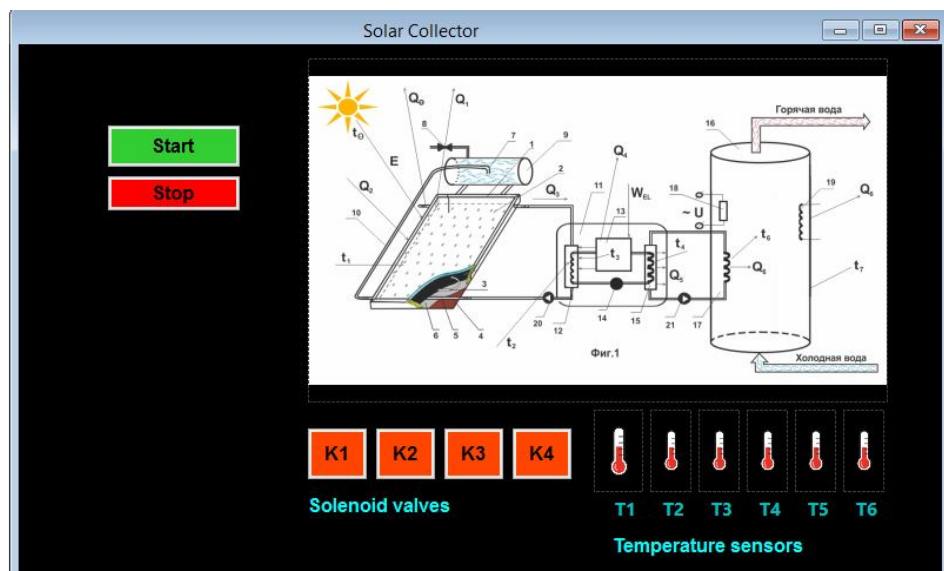


Рис. 4. Интерфейс управления программного обеспечения солнечного котроллера

На рисунке 4 приведена система управления двухконтурной гелиоустановки с термосиффонной циркуляцией. Интерфейс управления программного обеспечения солнечного котроллера состоит из самой системы рисунок (1), четырех индикаторов состояния клапанов, шести температурных датчиков для связи с контроллером, кнопкой старт и стоп.

Внизу приведен программный код на языке VHDL, для управления двухконтурной гелиоустановки с термосифонной циркуляцией, который содинен через Ethernet модуль с Веб Сервером.

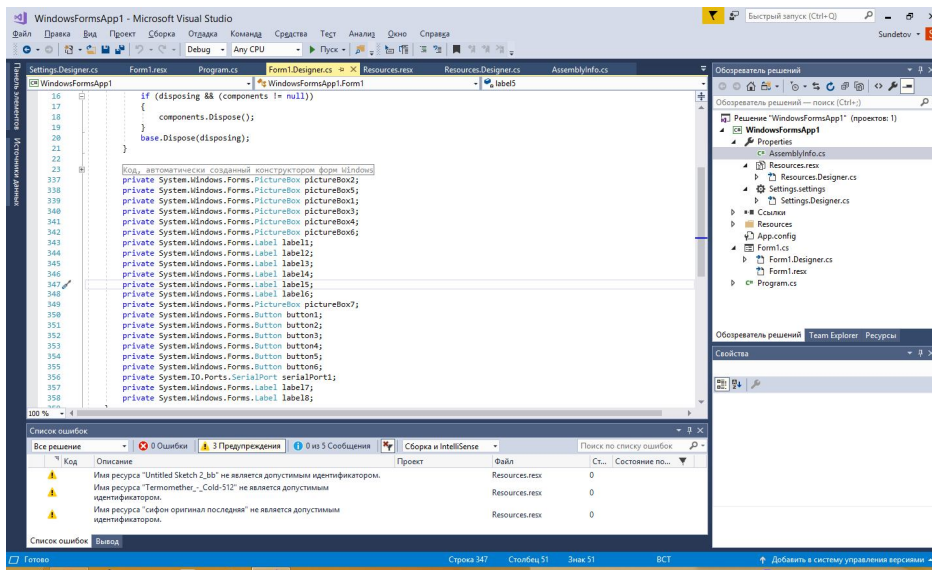


Рис. 5. Среда разработки Visual Studio для объявления элементов на программном ресурсе

На рисунке 5 показан код каждого элемента программного обеспечения кнопок, графических элементов и лэйбэлы.

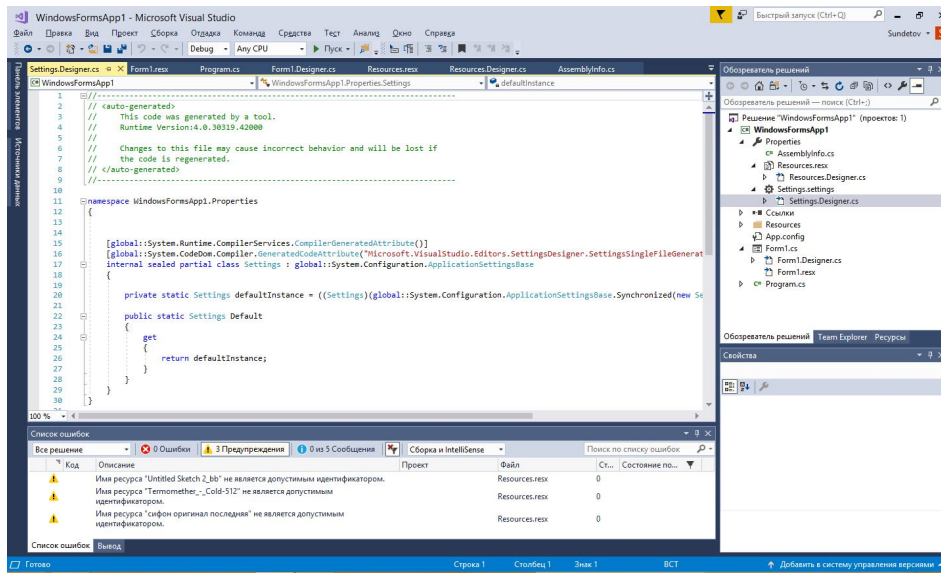


Рис. 6. Среда разработки Visual Studio для настройки программного обеспечения связки элементов программно ресурсе

Полное программное обеспечение, которое показывает, считывает данные гелиоустановки.

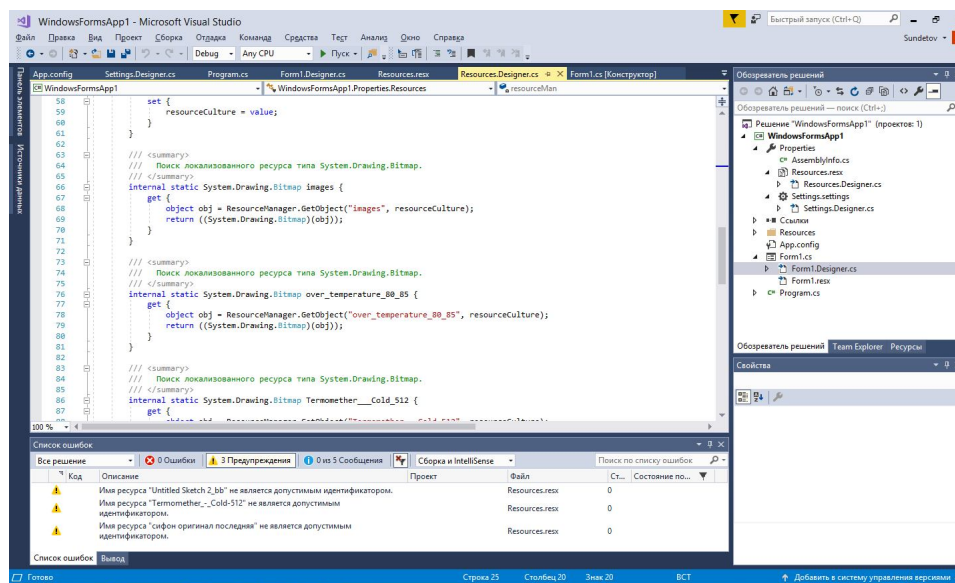


Рис. 7. Среда разработки Visual Studio для связки элементов программного ресурса

В программном обеспечении NET Ethernet_MAC_COL LOC являются локальными MAC адресами для соединения с Веб Сервером. NET LED [0] LOC-порты для управления клапанами и температурными датчиками. NET uart1_sin LOC-порты для обмена данными. NET "CLK_100MHZ" TNM_NET = sys_clk_pin-скорость соединения Ethernet модуля с Веб Сервером. TIMESPEC TS_sys_clk_pin = PERIOD sys_clk_pin 100000 kHz- период NET "RESET" LOC = "B6" | IOSTANDARD = LVCMOS33 | DRIVE = 8 | SLEW = FAST | PULLDOWN = TRUE-порты для перезагрузки связи с Веб Сервером.

Работа программного обеспечения основывается для считывания температурных данных через Ethernet и визуально показывает на ЖК- дисплее и Веб Сервере. А также можно управлять через Веб Сервер электромагнитными клапанами. В программном обеспечении имеется дополнительная интеллектуальная программа, которая зависит от температурных данных термосифонного гелиоустановки.

Заключение

В настоящей работе разработан алгоритм программы, которая активирует библиотеки и константы, в дальнейшем подключается модуль часы в реальном времени, активация температурных датчиков и активация связи с сервером. Связь устанавливается с сервером, сохраняет xml файл в Ethernet модуле, а так же считываются температурные данные датчиков и записываются данные о температуре в Ethernet модуле. Создана блок-схема Веб Сервера для контроллера управления двухконтурной гелиоустановки с термосифонной циркуляцией, которая инициализируют библиотеки SPI, Ethernet, а так же активизируются IP и MAC адреса. Работа программного обеспечения основывается для считывания температурных данных через Ethernet и визуально показывает на ЖК- дисплее и Веб Сервере

Литература

1. Kalogirou S., Solar thermal collectors and applications, Progress in Energy and Combustion Science, 30, 2004, 231–295.

2. Wang Z., Yang W., Qiu F., Zhang X., Zhao X., Solar water heating: From theory, application, marketing and research, *Renewable and Sustainable Energy Reviews*, 41, 2015, 68–84.
3. http://www.compit.pl/katalog/kategoria_12_regulatory-ukladow-solarnych.html [online: 02.06.2015]. 86
4. <http://www.geco.pl/prod.php?id=169&lang=pl> [online: 02.06.2015].
5. http://www.steca.com/index.php?Solar_Electronics_by_Steca [online: 02.06.2015].
6. <http://www.caleffi.com/usa/en-us/catalogue/solar-1> [online: 02.06.2015].
7. Nagaraju J., Garud S.S., Ashok Kumar K., Ramakrishna Rao M., 1MWth industrial solar hot water system and its performance, *Solar Energy*, 66, 1999, 491–497.
8. <http://www.arduino.cc> [online: 02.06.2015].
9. Stettler S. et al., Web-based functionality checks for solar heating systems, *Energy Procedia*, 48, 2014, 674–680.
10. Porzuczek J., Zabezpieczenie ciągłości zasilania w małych systemach HVAC, *Przegląd Naukowo-Metodyczny*, nr 24, Poznań 2014, 665–670.
11. Grassie T., MacGregor K., Muneer T., Kubie J., Design of a PV driven low flow solar domestic hot water system and modeling of the system collector outlet temperature, *Energy Conversion and Management*, 43, 2002, 1063–1078.
12. PN-EN 12977-5:2012, Słoneczne systemy grzewcze i ich elementy – Systemy wykonywane na zamówienie, Część 5: Metody badania wydajności wyposażenia regulacyjnego.
13. Peter M., Drück P., Testing of controllers for thermal solar systems, *Solar Energy*, 82, 2008, 676–685.

К ВОПРОСУ О ПОКОЛЕНИЯХ В БАЙЕСОВСКИХ СЕТЯХ

Литвиненко Н.Г., Литвиненко А.Г., Мамырбаев О.Ж.,
Шаяхметова А.С.

E-mail: n.litvinenko@inbox.ru, litvinenko@uq.rwth-aachen.de,
morkenj@mail.ru, a_sh_s83_83@mail.ru

*Институт информационных и вычислительных технологий МОН РК,
Казахстан*

Абстракт. *Статья посвящена вопросам исследования некоторых аспектов теории ориентированных графов при работе с байесовскими сетями. В некоторых статьях по теории байесовских сетей было введено понятие «Поколение вершин», обозначающее некоторое множество вершин, имеющих множество родителей, принадлежащих более ранним поколениям (определения будут даны ниже). Терминология для данного понятия, на наш взгляд, еще полностью не сложилась. Приносим извинения тем, кому наша терминология, наши понятия и определения покажутся недостаточно точными. Понятие «Поколение» в некоторых случаях позволяет более легко решать некоторые задачи в байесовских сетях, строить более простые алгоритмы.*

В статье будут использованы два учебных примера. Первый пример – известный, учебный пример Asia, описанный во многих статьях и книгах, а также

в технической документации к различным пакетам. Вторым примером – учебный пример *Wacan*, построенный нами для иллюстрации рассматриваемых понятий.

В статье для построения примеров использовались учебные версии пакетов *BayesiaLab* и *AgenaRisk*. Для первоначального понимания основ байесовских сетей читатель может воспользоваться [1, 2, 3, 4, 5, 6, 7, 8, 9]. Для понимания работы пакета *BayesiaLab* отсылаем к технической документации [10], а также к описанию основных возможностей работы с байесовскими сетями в данном пакете [11]. Для первоначального знакомства с возможностями пакета *AgenaRisk* можно посмотреть техническую документацию [12, 13].

Ключевые слова. Байесовские сети, *AgenaRisk*, ориентированные графы, поколение вершин.

Введение

Искусственный интеллект прочно вошел в жизнь исследователей проблем в различных областях науки. Наиболее популярным направлением использования искусственного интеллекта с начала 2000 годов стало использование в исследованиях аппарата байесовских сетей. Для корректного использования байесовских сетей был разработан соответствующий математический аппарат. Появились широкие возможности для исследования проблем в различных областях науки. Однако массивные, сложные вычисления при использовании аппарата байесовских сетей настоятельно требовали привлечения как хорошей вычислительной техники, так и наличия хороших программных продуктов, обеспечивающих удобную работу с байесовскими сетями.

На рынке программных продуктов появилось много интересных пакетов для работы с байесовскими сетями. К наиболее популярным пакетам можно отнести: *BayesiaLab*, *AgenaRisk*, *Bayes Server*, *Netica*, *Hugin Expert*, *BayesFusion*. Несмотря на то, что некоторым пакетам уже более 15 лет, до сих пор идет притирка между математиками, занимающимися теоретическими основами байесовских сетей, разработчиками алгоритмов и программистами, а также исследователями в прикладных областях науки.

В данной статье будет обозначено одно из теоретических направлений в теории байесовских сетей – разбиение вершин байесовской сети на поколения. На наш взгляд данное направление теории байесовских сетей достаточно слабо представлено в литературе, однако в некоторых случаях разбиение вершин на поколения может позволить по-новому взглянуть на общую теорию байесовских сетей, а также позволит облегчить разработку алгоритмов для некоторых задач.

Основные понятия

БС представляют собой удобный инструмент для описания достаточно сложных процессов с неопределенностями.

Кратко напомним основные понятия, связанные с байесовскими сетями.

Определение. Байесовской сетью называется ациклический ориентированный граф, в котором соблюдается Марковское условие. Вершины графа часто называют узлами. Узлы представляют некоторые переменные, отражающие основные сущности в разрабатываемой модели.

Определение. Байесовская сеть называется **дискретной**, если переменные, представляющие узлы, дискретные.

Определение. Байесовская сеть называется **непрерывной**, если переменные, представляющие узлы, непрерывные.

Определение. Байесовская сеть называется **гибридной**, если в байесовской сети существуют узлы, представляющие как **дискретные**, так и **непрерывные** переменные.

Определение. Байесовская сеть называется **динамической**, если переменные, представляющие узлы, изменяются со временем.

Некоторые узлы байесовской сети могут быть связаны между собой направленными дугами (ребрами). Дуги определяют некоторую вероятностную связь между соответствующими узлами. Иногда такая связь носит причинно-следственный характер. Причиной является узел откуда выходит направленная дуга, следствием называется узел, куда приходит направленная дуга.

Определение. Ребрами называют дуги без учета направления.

Определение. Если узлы не соединены дугой, то считается, что данные узлы **условно независимы**.

Определение. Скелетом байесовской сети называется граф, полученный из байесовской сети заменой дуг на **ребра**.

Определение. Если из вершины **A** выходит дуга в вершину **B**, то **A** называют родителем **B**, а **B** называют дочерней вершиной вершины **A**.

Определение. Пусть **Y** – некоторое подмножество вершин байесовской сети. Через **P(Y)** часто обозначают множество всех **родителей**, принадлежащих **Y**. Через **C(Y)** часто обозначают множество всех **детей**, принадлежащих **Y**.

Определение. Узел **x** и дуга **e** называются **инцидентными**, если **e** входит или исходит из **x**.

Определение. Узлы называются **смежными**, если они **инциденты** одной дуге, т.е. если один из них является родителем другого.

Определение. Маршрутом из вершины **a₀** в вершину **a_n** в ориентированном графе (в байесовской сети) называют чередующуюся последовательность вершин и дуг, вида **a₀, {a₀ a₁}, a₁, {a₁ a₂}, a₂, {a₂ a₃}, ..., a_n**.

Определение. Путем называют маршрут без повторяющихся дуг.

Определение. Если из вершины **A** существует ориентированный путь в вершину **B**, то **A** называется **предком B**, а **B** называется **потомком A**.

Определение. Два узла называются **связными**, если между ними есть маршрут.

Определение. Если у вершины нет **предков**, то её локальное распределение вероятностей называют **безусловным**, иначе **условным**.

Определение. Топологической нумерацией узлов байесовской сети называется такая нумерация узлов, при которой номер любого узла больше чем номер его родителя.

Определение. Иногда для разработки алгоритмов расчета байесовской сети удобно разбить множество узлов сети на поколения. Поколения могут быть двух типов: поколения потомков и поколения предков. Поколения потомков определяются следующим образом:

- Узлы, не имеющие родителей, относятся к **0** поколению потомков.
- Узлы, имеющие родителей только **0** поколения, относятся к **1** поколению потомков.
- Узлы, имеющие родителей только **0** и **1** поколения, относятся ко **2** поколению потомков.
-
- Узлы, имеющие родителей только **0, 1, 2, ... K** поколения, относятся к **K+1** поколению потомков.

-

Определение. Поколения предков определяются следующим образом:

- Узлы, не имеющие детей, относятся к **0** поколению предков.
- Узлы, имеющие детей только **0** поколения, относятся к **1** поколению предков.
 - Узлы, имеющие детей только **0** и **1** поколения, относятся ко **2** поколению предков.
-
- Узлы, имеющие детей только **0, 1, 2, ... К** поколения, относятся к **К+1** поколению предков.
-

Примеры байесовской сети

Рассмотрим два примера байесовской сети. Первый пример - распространенный пример байесовской сети, часто используемый во многих статьях.

На рисунке 1 изображена байесовская сеть, отражающая причинно-следственные связи между заболеванием раком и туберкулезом в зависимости от некоторых сопутствующих факторов.

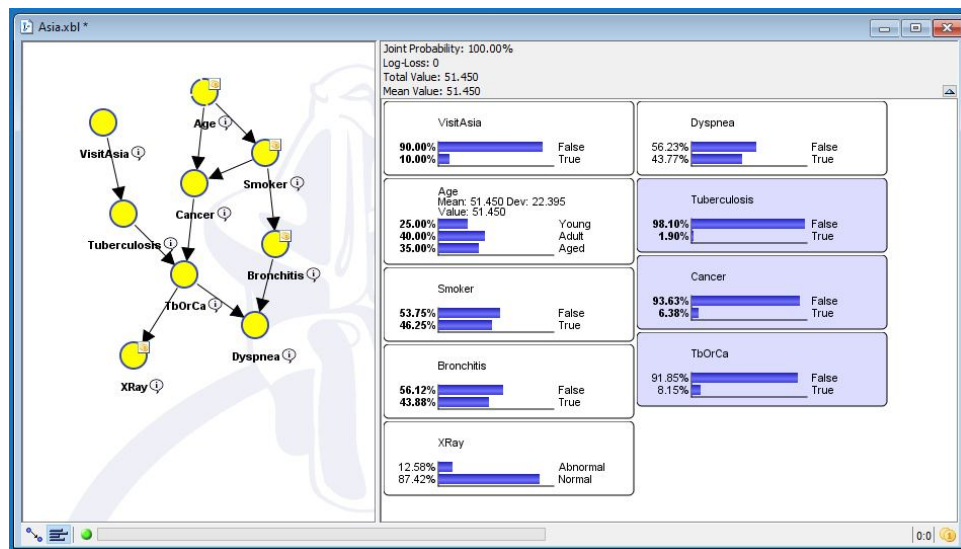


Рисунок 1

Рассмотрим поколения потомков и поколения предков для данной байесовской сети.

Поколения потомков для данной байесовской сети:

- К нулевому поколению будут отнесены вершины **Age** и **VisitAsia**.
- К первому поколению относятся вершины **Smoker** и **Tuberculosis**.
- Ко второму поколению относятся вершины **Cancer** и **Bronchitis**.
- К третьему поколению относится только одна вершина **TbOrCa**.
- К четвертому поколению относятся вершины **XRay** и **Dyspnea**.

Поколения предков для данной байесовской сети:

- К нулевому поколению будут отнесены вершины **XRay** и **Dyspnea**.
- К первому поколению относятся вершины **TbOrCa** и **Bronchitis**.

- Ко второму поколению относятся вершины **Cancer** и **Tuberculosis**.
- К третьему поколению относятся вершины **Smoker** и **VisitAsia**.
- К четвертому поколению относится вершина **Age**.

Второй пример – учебный пример, показывающий возможность оценить претендента на вакантную должность. Подробное описание данного примера приведено в [14].

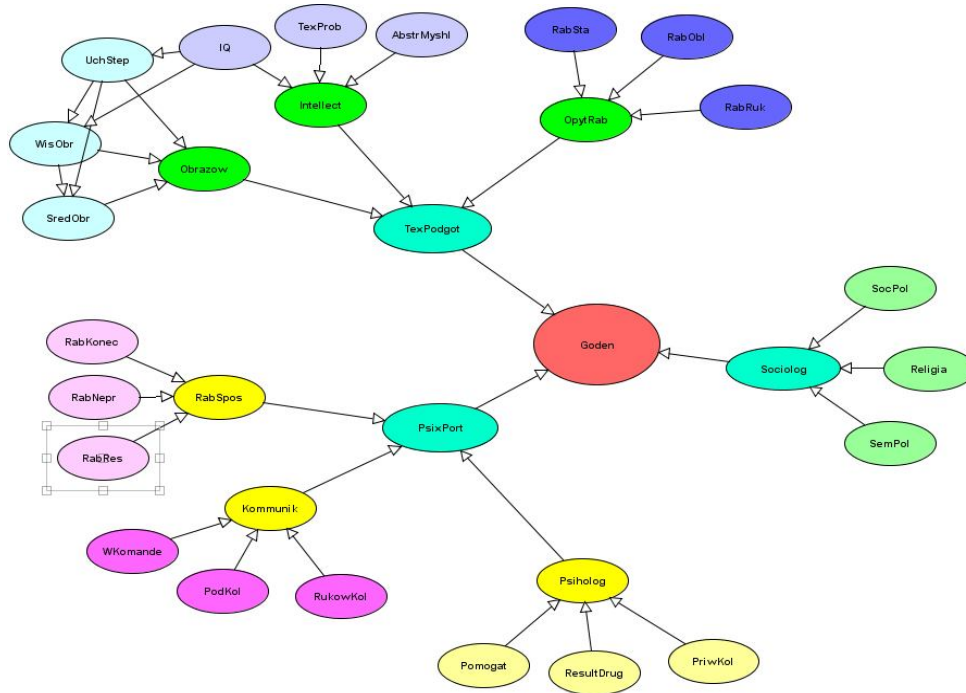


Рисунок 2.

Рассмотрим поколения потомков и поколения предков для данной байесовской сети.

Вначале рассмотрим поколения потомков. В данной байесовской сети находятся в основном узлы нулевого поколения:

- К нулевому поколению будут отнесены вершины **RabSta, RabObl, RabRuk, SocPol, Religia, SemPol, Pomogat, ResultDrug, PriwKol, WKomande, PodKol, RukowKol, RabKonec, RabNepr, RabRes, IQ, TexProb, AbstrMyshl**.

- К первому поколению относятся вершины **OpytRab, Sociolog, Psiholog, Kommunik, RabSpos, Intellect, UchStep**.

- Ко второму поколению относятся вершины **PsixPort, WisObr**.
- К третьему поколению относится только одна вершина **SredObr**.
- К четвертому поколению относится вершина **Obrazow**.
- К пятому поколению относятся вершины **TexPodgot**.

- К шестому поколению относятся вершины **Goden**.

Поколения предков для данной байесовской сети:

- К нулевому поколению будут отнесена вершина **Goden**.
- К первому поколению относятся вершины **PsixPort, TexPodgot, Sociolog**.
- Ко второму поколению относятся вершины **SocPol, Religia, SemPol, Psiholog, Kommunik, RabSpos, Obrazow, Intellect, OpytRab**.

- К третьему поколению относятся вершины **RabSta, RabObl, RabRuk, TexProb, AbstrMyshl, RabKonec, RabNep, RabRes, WKomande, PodKol, RukowKol, Pomogat, ResultDrug, PriwKol, SredObr.**

- К четвертому поколению относится вершина **WisObr.**

- К пятому поколению относится вершина **UchStep.**

- К шестому поколению относится вершина **IQ.**

Использование понятия «Поколение»

Разбиение вершин на поколения потомков достаточно просто. Вначале ищем вершины, не имеющие родителей. Относим их к нулевому поколению, отмечаем выбранные вершины. Далее просматриваем неотмеченные вершины и выбираем только те вершины, у которых родители принадлежат только к нулевому поколению. Получаем первое поколение потомков, также отмечаем вновь выбранные вершины. Опять просматриваем оставшиеся неотмеченные вершины и выбираем только те вершины, у которых родители принадлежат только или к нулевому, или к первому поколению. Получаем второе поколение потомков, также отмечаем выбранные вершины. Так продолжаем до тех пор, пока все вершины не окажутся отмеченными.

***Теорема 1.** Если на каком-то этапе построения поколений какое-то поколение оказывается пустым, но неотмеченные вершины остались, в графе присутствуют ориентированные циклы.*

Доказательство.

Если на некотором этапе из оставшихся неотмеченных вершин не оказалось вершин, имеющих родителей из предыдущих поколений, то у рассматриваемого узла родителями могут быть только узлы рассматриваемого поколения. Выбираем любого родителя, у которого также родители не принадлежат предыдущим поколениям. В результате имеем неограниченный путь с ограниченным числом узлов, то есть цикл.

Таким образом, мы имеем достаточно простой способ выяснить имеет ли ориентированный граф цикл и найти место, где находится данный цикл.

***Теорема 2.** Две изоморфные байесовские сети имеют одинаковое разбиение на поколения.*

Доказательство опускаем.

Задача выяснения изоморфности двух байесовских сетей значительно упрощается при использовании понятия «Поколение».

***Теорема 3.** Байесовская сеть допускает топологическую нумерацию.*

Доказательство.

Первый способ нумерации

Вначале произвольно нумеруем узлы нулевого поколения, затем продолжим нумерацию для узлов первого поколения, затем для узлов второго поколения, и т.д.

Так как родители каждого узла находятся на предыдущих поколениях, построенная нумерация будет топологической.

Второй способ нумерации

Рассмотрим все узлы нулевого поколения. Возьмём произвольный узел нулевого поколения и рассмотрим множество **M1** всех потомков данного узла на всех поколениях. Нумеруем узлы множества **M1** следующим образом. Узлу нулевого поколения присваиваем номер 1. Далее нумеруем узлы множества **M1**

первого поколения, затем узлы множества M_1 второго поколения и т. д. Отмечаем все пронумерованные узлы.

Далее возьмем следующий узел нулевого поколения и присвоим ему очередной номер. Рассмотрим множество M_2 всех потомков данного узла, не отмеченных ранее. Далее нумеруем узлы множества M_2 первого поколения, затем узлы множества M_2 второго поколения и т. д. Опять отмечаем все пронумерованные узлы.

То же самое проделываем со всеми оставшимися узлами нулевого поколения. Нумеруем аналогичным способом.

Так как родители каждого узла находятся на предыдущих поколениях, построенная нумерация будет топологической.

Теорема 4. *Байесовская сеть допускает такую топологическую нумерацию, при которой номер любого узла X будет больше номеров любого узла из произвольного множества узлов M , не содержащего узла X , а также потомков узла X .*

Доказательство.

Немного изменим второй способ нумерации из теоремы 3.

Пусть M_X – множество узлов рассматриваемой байесовской сети, состоящее из узла X и его потомков. Пусть L – количество узлов в рассматриваемой байесовской сети. Проведем топологическую нумерацию узлов сети вторым способом, описанным в теореме 3.

Ко всем номерам узлов множества M_X добавим число L . Непрерывная нумерация узлов байесовской сети будет нарушена. Однако номер любого узла останется по-прежнему больше номера его родителя. Теперь необходимо восстановить непрерывность нумерации байесовской сети.

Расположим в порядке возрастания полученные номера узлов байесовской сети. Пронумеруем данные номера узлов. Если за новые номера узлов взять нумерацию номеров узлов, то данная нумерация по-прежнему останется топологической нумерацией. Кроме того, номера узлов массива M_X будут больше номеров всех оставшихся узлов сети, а значит и больше номеров узлов массива M .

Теорема 5. *Байесовская сеть допускает такую топологическую нумерацию, при которой номер любого узла X будет меньше номеров любого узла из произвольного множества узлов M , не содержащего узла X , а также предков узла X .*

Доказательство.

Немного изменим второй способ нумерации из теоремы 3.

Пусть M_X – множество узлов рассматриваемой байесовской сети, состоящее из узла X и его предков. Пусть L – количество узлов в рассматриваемой байесовской сети. Проведем топологическую нумерацию узлов сети вторым способом, описанным в теореме 3.

Ко всем номерам узлов байесовской сети кроме множества M_X добавим число L . Непрерывная нумерация узлов байесовской сети будет нарушена. Однако номер любого узла останется по-прежнему больше номера его родителя. Теперь необходимо восстановить непрерывность нумерации байесовской сети.

Расположим в порядке возрастания полученные номера узлов байесовской сети. Пронумеруем данные номера узлов. Если за новые номера узлов взять нумерацию номеров узлов, то данная нумерация по-прежнему останется

топологической нумерацией. Кроме того, номера узлов массива **МХ** будут меньше номеров всех оставшихся узлов сети, а значит и меньше номеров узлов массива **М**.

Заключение

Данная статья показывает некоторые возможности использования понятия «Поколение» при доказательстве некоторых стандартных теорем в теории байесовских сетей. Использование данного понятия также упрощает некоторые алгоритмы для задач в байесовских сетях, например, поиск циклов или проверка изоморфизма двух сетей.

Приведена лишь небольшая часть возможностей использования понятия «Поколение».

Данная статья была опубликована в рамках грантового проекта «Разработка и программная реализация пакета прикладных программ для решения прикладных задач по байесовским сетям». ИРН проекта: AP05131293.

Список литературы

1. F.V. Jensen, T.D. Nielsen, Bayesian Networks and Decision Graphs. Springer. – 2007. – 447 p.
2. K.P. Murphy, An introduction to graphical models. – 2001. - 19 p. <http://www.stat.yale.edu/~jtc5/BioinformaticsCourse2001/MurphyBayesNetIntro.pdf>
3. D. Barber, Bayesian Reasoning and Machine Learning. – 2017. – 686 p. <http://web4.cs.ucl.ac.uk/staff/D.Barber/textbook/020217.pdf>
4. D. Barber, Machine Learning. A probabilistic Approach. – 343 p. <https://pdfs.semanticscholar.org/7bc7/54bc548f32b9ac53df67e3171e8e4df66d15.pdf>
5. K.P. Murphy, Machine Learning A Probabilistic Perspective. The MIT Press. – 1067 p.
6. R.E. Neapolitan, Learning Bayesian Networks. – 704 p. [http://www.cs.technion.ac.il/~dang/books/Learning%20Bayesian%20Networks\(Neapolitan,%20Richard\).pdf](http://www.cs.technion.ac.il/~dang/books/Learning%20Bayesian%20Networks(Neapolitan,%20Richard).pdf)
7. М. Гончаров, Байесовские сети. – 2011. – 46 с. <http://www.businessdataanalytics.ru/download/BayesianNetworks.pdf>
8. П.И. Бидюк, А.Н. Терентьев, Построение и методы обучения Байесовских сетей //Таврический вестник информатики и математики. – 2004. – №2. - с.139-154
9. Литвиненко Н.Г., Литвиненко А.Г., Шаяхметова А.С., Мамырбаев О.Ж. К вопросу о классификации типов свидетельств в байесовских сетях // Матер.научн. конф. ИИВТ МОН РК «Современные проблемы информатики и вычислительных технологий». – Алматы. – С. – 2018
10. S. Conrady, L. Jouffe, Bayesian Networks & BayesiaLab. A Practical Introduction for Researches. ISBN: 978-0-9965333-0-0
11. Литвиненко Н.Г., Литвиненко А.Г., Мамырбаев О.Ж., Шаяхметова А.С. Работа с Байесовскими сетями в BAYESIALAB. – Алматы: ИИВТ, 2018. – 311 с.
12. AgenaRisk 7.0 User Manual. – 2016. - 204 p. <http://www.agenarisk.com>
13. Getting Started with AgenaRisk. – 2013. - 31 p. <http://www.agenarisk.com>
14. Литвиненко Н.Г., Литвиненко А.Г., Шаяхметова А.С., Мамырбаев О.Ж., Сатымбеков М., Турдалыулы М. Использование байесовских сетей при исследовании соответствия претендента вакантной должности // Матер. III Международ. научн. конф. «Информатика и прикладная математика». – 2018. – Алматы. – С.324-332

ОБНАРУЖЕНИЕ ОНКОЛОГИЧЕСКИХ ПАТТЕРНОВ МЕТОДАМИ ДИСКРЕТНОЙ ДИФФЕРЕНЦИАЛЬНОЙ ГЕОМЕТРИИ

Макаренко Н.Г., Пак А.А., Панина А., Каримова Л.М.,
Пазовский Н.С.

e-mail: ng-makar@mail.ru, aa.pak83@gmail.com, nikita.pazovskii@gmail.com

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** В настоящей работе рассматривается компьютерный анализ медицинских изображений, основанный на кривизне Формана-Риччи, которая известна из дискретной дифференциальной геометрии. Медицинские снимки представляют собой гистологические образцы лимфатических узлов пациентов с онкологическими заболеваниями молочной железы. Типичный снимок достигает размеров 10 гигапикселей, а свойства его фотометрической меры затрудняют использование стандартных методов распознавания, основанных на стистиках второго порядка. Дискретная кривизна содержит информацию о локальных и глобальных особенностях геометрии изображений и удобна для эффективного машинного обнаружения онкопаттернов.*

Введение

Рак молочной железы является одним из самых смертоносных и широко распространенных онкологических заболеваний среди женщин [1]. Диагностические методы демонстрируют хорошие результаты средней выживаемости, то есть для пациентов с 5-летней выживаемостью доля равна 90%, а для пациентов с 10-летней выживаемостью доля составляет 83%. Однако, при метастазировании рака молочной железы доля выживаемости значительно сокращается [2]. Например, показатель 5-летней выживаемости при локализованном раке молочной железы составляет 99%, но снижается до 85% в случае наличия регионарных метастазов (в лимфатические узлы) и до 26% в случае отдаленного метастазирования. Следовательно, для выбора адекватного метода лечения и увеличения выживаемости, критически важным является установить наличие или отсутствие метастазов. Это отражено в критериях классификации TNM [3].

Традиционно для идентификации стадии метастазирования первым делом обследуют региональные лимфатические узлы, что является очень важным диагностическим фактором. Иными словами, наличие раковых клеток в региональных лимфатических узлах говорит о возможном наличии отдаленных метастазов [4]. При раке молочной железы наиболее распространенным способом оценки состояния регионарных лимфоузлов является процедура обнаружения «сторожевого» лимфатического узла. При этой процедуре рядом с опухолью в молочной железе вводят специальное контрастное вещество или радионуклид. Тот лимфоузел, в который контрастное вещество/радионуклид поступит быстрее всего, является «сторожевым» и, вероятней, метастатически поражен. Впоследствии «сторожевой» лимфоузел иссекается и направляется на патоморфологическое исследование.

Врач-патоморфолог исследует предметное стекло, содержащее препарат тканей лимфатического узла, окрашенный гематоксилином и эозином (H&E). Основываясь на размере единичных опухолевых клеток или диаметра скоплений опухолевых клеток метастазы можно разделить на три категории: макрометастазы, микрометастазы или изолированные опухолевые клетки. На основании наличия или отсутствия одной, или нескольких из этих групп метастазов пациенту назначается начальная N-стадия (pN-стадия). Далее, учитывая установленную начальную стадию в сочетании с характеристиками основного опухолевого узла, может быть проведена дальнейшая лимфодиссекция или лучевая терапия на аксиллярную область. В случае лимфодиссекции, извлеченные лимфоузлы подвергаются патоморфологическому анализу для установления заключительной стадии. Распределение по категориям зависит от количества и размеров метастазов, а также от анатомического расположения лимфатических узлов [7]. (см. Таблицы 1 и 2 в Приложении)

Обработка гистологических проб с целью выявления стадия заболевания является трудозатратной задачей, из-за большой площади ткани, которую нужно исследовать с целью выявления метастазов, которые могут быть столь же малы, как и отдельные клетки. Для дозорных лимфатических узлов необходимо исследовать не менее трех срезов на разных уровнях детализации, для недозорных узлов должен быть исследован один срез по крайней мере из 10 лимфатических узлов [8,9]. В результате этого ресурсоемкого процесса врач гистолог может пропустить небольшие метастазы [10].

Проблема автоматической обработки и анализа медицинских гистологических снимков является сложной с практической точки зрения. Оцифровка всего слайда - это метод, при котором высокоскоростные сканеры слайдов оцифровывают препараты с очень высоким разрешением (например, 240 нм на пиксель). Это приводит к изображениям размером порядка 10 гигапикселей, которые обычно называются изображениями с полным слайдом (WSI).

В рамках традиционных задач текстурного анализа такие снимки представляют собой дискретные множества высокой размерности, отягощенные эффектами "curse of dimensionality". Поэтому, применение стандартных статистических методов становится неэффективным: при машинном обучении паттернам, практически любой случайный вектор будет ортогонален вектору данных [11]. Мы применяем в этой работе методы дискретной дифференциальной геометрии [12]. Более точно, мы вычисляем для цифровых изображений дискретный аналог кривизны Риччи, так называемую кривизну Формана-Риччи [13,14]. В последнее время ее успешно используют в задачах связанных с обработкой изображений. Мы проверяем этот подход на базе данных SAMEION17 пяти медицинских центров Голландии [15].

Эти данные содержат WSI изображения образцов иссечения лимфатических узлов окрашенных контрастирующими (H&E) веществами. Основной целью работы была оценка эффективности кривизны Формана-Риччи для машинного распознавания новообразований в гистологических пробах.

В разделе 1 описана предобработка данных и необходимый формализм из дискретной дифференциальной геометрии. В разделе 2 обсуждение и анализ полученных результатов. Заключение содержит методические рекомендации по дальнейшему использованию рассмотренного подхода.

1. Предобработка данных и элементы теории кривизны Формана – Риччи

Размер анализируемого изображения равен 200 000 на 100 000 пикселей, большая часть снимка покрыта черным и не несет какой-либо полезной информации. В наборе данных вместе с гистологическим изображением представлена бинарная маска для локализации здоровых и раковых клеток, обычно маска представляет собой некоторую связную область произвольной формы. Одним из первых шагов предобработки изображения является генерация покрытия в форме разбиения на отдельные, не пересекающиеся между собой квадраты размером 304 на 304 пикселя. Следует отметить, что после такого измельчения каждый квадрат содержит либо здоровые, либо раковые клетки. Далее для тестирования из набора квадратов случайным образом были выбраны 200 примеров здоровых и раковых клеток, соответственно.

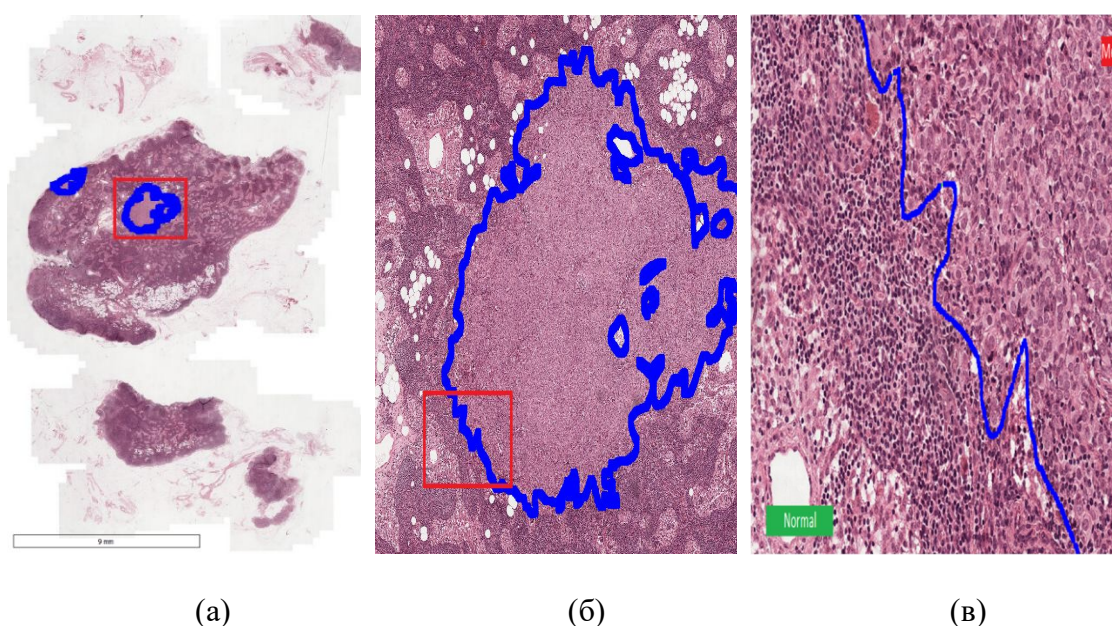


Рис. 1 Примеры гистологических данных и элементов их предобработки

- (а) Пример изображения без увеличения;
- (б) Пример бинарной маски, обозначающей раковые клетки;
- (в) Разделение здоровых и больных клеток

Приведем теперь краткие сведения и кривизне Формана-Риччи в приложении к цифровым изображениям [16]. Напомним, прежде всего, что кривизна Риччи определяется для дуг и ребер, ее усреднение для вешины по всем дугам называют скалярной кривизной. Поясним эристическую идею этой кривизны. Рассмотрим две точки (x,y) на дважды дифференцируемой поверхности размерности N . Предположим что они связаны геодезической длиной $d(x,y)$. Рассмотрим вектор w длиной ε , касательный к некоторому направлению по геодезической (См. Рис. 2). Изменение его направления образует ε окрестность точки $x - S_x$. Ради простоты, можно думать что (x,y) две точки на экваторе сферы $N=2$, а вектор w направлен вдоль меридиана. Перенесем этот вектор параллельно самому себе в точку y . Он будет замечать ε окрестность точки $y - S_y$. Измерим все расстояния между концами пар векторов w и его образа в окрестности S_y при разных положениях вектора w и усредним их. Для плоского многообразия, это усредненное расстояние совпадет с

$d(x,y)$. Эта ситуация соответствует нулевой кривизне Риччи. В нашем примере для сферы, сдвиг на ε по меридиану соответствует углу широты φ . Поэтому:

$$d_\varepsilon(x,y) = d(x,y) \cos \varphi \approx d(x,y) \left(1 - \frac{\varphi^2}{2}\right) = d(x,y) \left(1 - \frac{\varepsilon^2}{2R^2}\right)$$

Мы видим, что уменьшение расстояния пропорционально кривизне сферы: $K = 1/R^2$. В общем случае, расстояние уменьшается пропорционально кривизне Риччи $\sim (1 - \varepsilon^2 \frac{Ric}{2N})$ (См. Рис. 2). Для дальнейшего, важен основной принцип приведенной эвристики: кривизна участка дуги определяется сравнением своей геодезической длины, с аналогами, полученными параллельным сдвигом вдоль направлений, ортогональных участку.

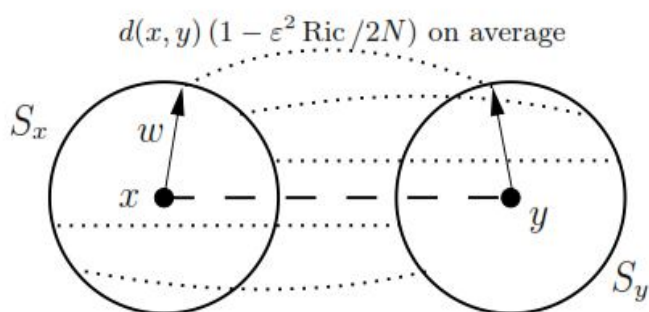


Рис. 2 Определение кривизны Риччи

На изображении, роль длины играют веса ребер, которые можно определить разными способами, на основе значений фотометрической меры (контраста или почернения) в гранях, которые разделяются выбранным ребром. Усреднение, можно учесть влиянием ребер, смежных с выбранным, по концам выбранного ребра. Это и составляет дискретный вариант кривизны по формуле Бохнера, полученной Форманом, в рамках дискретной теории Морса [17]. Точнее, рассмотрим фрагмент изображения, приведенный на Рис. 3.

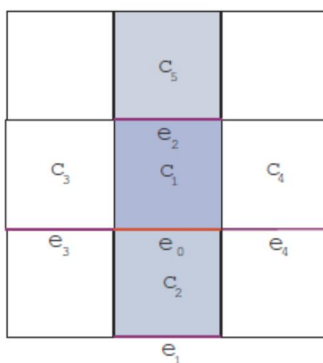


Рис. 3 К определению кривизны Риччи для изображения

Кривизна Риччи, для ребра e_0 , которое разделяет грани c_1, c_2 определяется формулой, содержащей веса ω граней и ребер [16]. Легко заметить что кривизна учитывает лишь ребра параллельные e_0 и смежные с ним по грани: e_2 смежно по грани c_1 и e_1 , смежное по грани c_2 . Ребра e_3 и e_4 не учитываются, потому что являются продолжением выбранного ребра и смежны с ним лишь по его вершинам.

$$Ric(e_0) = \omega(e_0) \left[\left(\frac{\omega(e_0)}{\omega(c_1)} + \frac{\omega(e_0)}{\omega(c_2)} \right) - \left(\frac{\sqrt{\omega(e_0)\omega(e_1)}}{\omega(c_1)} + \frac{\sqrt{\omega(e_0)\omega(e_2)}}{\omega(c_2)} \right) \right] \quad (1)$$

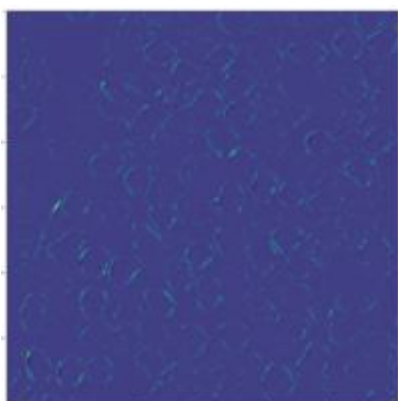
Веса ребер можно определить двумя способами [18]. Самый простой, который дает так называемые комбинаторные веса, заключается в нахождении среднего арифметического контрастов $\omega(c_i) \equiv I(c_i)$, содержащихся в гранях разделяющихся ребром. Так, для Рис. 3, комбинаторный вес ребра e_0 определяется как $\omega(e_0) = \omega(c_1)/\omega(c_2)$. Геометрические веса определяются несколько сложнее, как длины дуг, выраженные в метрике на изображении:

$$ds^2 = dx^2 + dy^2 + dI^2, \quad dI = \partial I/\partial x + \partial I/\partial y$$

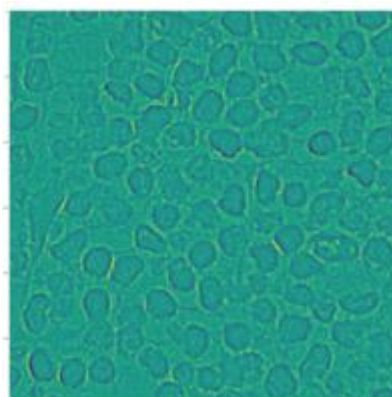
В этом случае, вес ребра для горизонтального и вертикального ребра определяются выражениями:

$$\omega(e_x) = \sqrt{\beta + (\partial I/\partial x)^2} dx; \quad \omega(e_y) = \sqrt{\beta + (\partial I/\partial y)^2} dy$$

Здесь β – масштабный фактор, выравнивающий числовые и фотометрические координаты. В качестве примера, на Рис. 4 приведены примеры вычисления кривизны Риччи для комбинаторных и геометрических весов.



(a)



(б)

Рис. 4 Кривизна Риччи, вычисленная с использованием геометрических (а) и комбинаторных (б) весов

2. Результаты и их анализ

Используя комбинаторные веса, для гистологических срезов вычислялись кривизны Формана-Риччи и строились их гистограммы. В качестве простейшей статистической характеристики использовались первые 4 центрированных момента гистограммы, используя стандартные алгоритмы [19]. На Рис. 5 приведены примеры таких вычислений для конкретных экземпляров гистологических изображений. Заметное отличие гистологических срезов демонстрируют 3 и 4 моменты (см. Рис. 6)

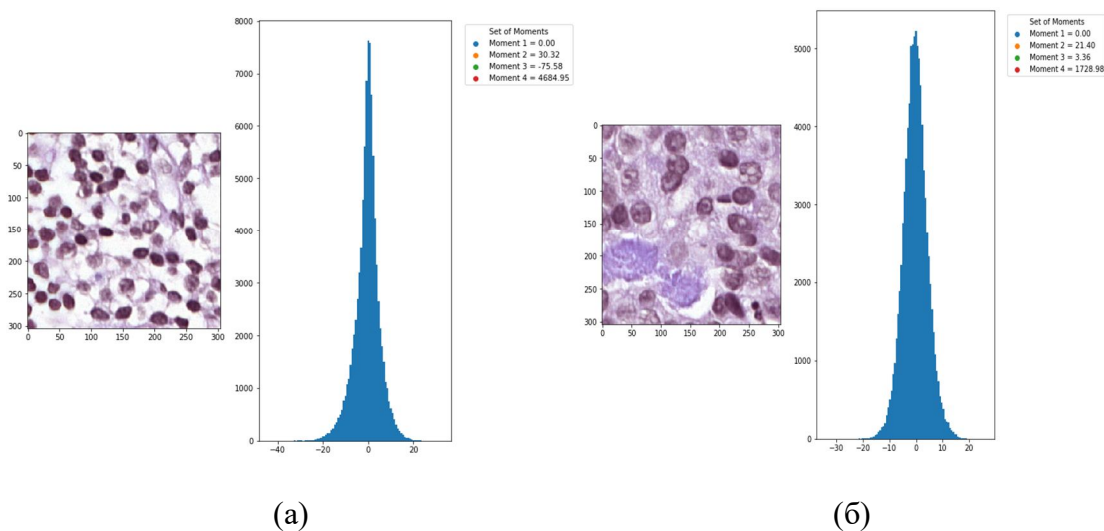


Рис. 5 Пример квадрата с нормальными(а) и раковыми(б) клетками и соответствующей ему гистограммы четвертых моментов кривизн Риччи

В качестве дополнительной дискриминатора для распознавания онкопаттернов мы использовали так называемое диффузное расстояние [20]. На эвристическом уровне вычисление диффузного расстояния между двумя гистограммами заключается в следующем. Каждая гистограмма рассматривается как начальное распределение температуры на вещественной оси. Под действием 1D оператора диффузии, первоначальное распределение «размывается» распределяя температуру равномерно вдоль оси. Рассмотрим побиноую разность 2-х гистограмм $d(x) = h_1(x) - h_2(x)$. Будем редуцировать к нулю эту разность с помощью решения уравнения теплопроводности

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2}, T(x, 0) = d(x)$$

Решение записывается в стандартном виде: $T(x, t) = d(x)\phi(x, t)$, где $\phi(x, t)$ - гауссовское ядро шириной t . Когда $t \rightarrow \infty$, решение $T(x, t) \rightarrow 0$, нивелируя разность двух гистограмм. Следовательно, мерой их диссимилиации может служить некоторая норма решения, например, L_1 :

$$K(h_1, h_2) = \int_0^\tau |T(x, t)| dt$$

которая измеряет диффузное расстояние числом шагов решения уравнения диффузии для получение нулевой разности гистограмм. На Рис. 7 представлена тепловая карта попарных диффузных расстояний для здоровых и раковых клеток на изображениях.

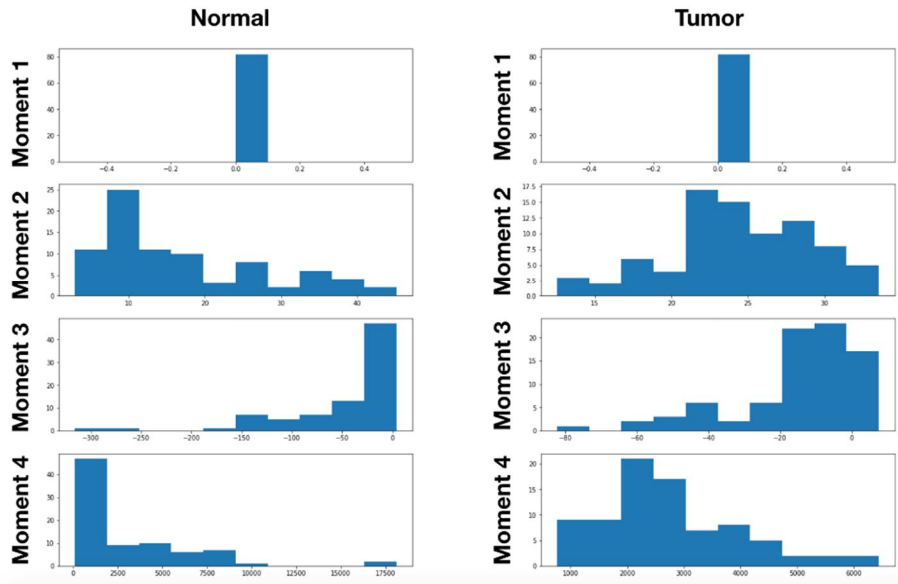


Рис. 6 Гистограммы статистических моментов кривизн Риччи для 200 изображений

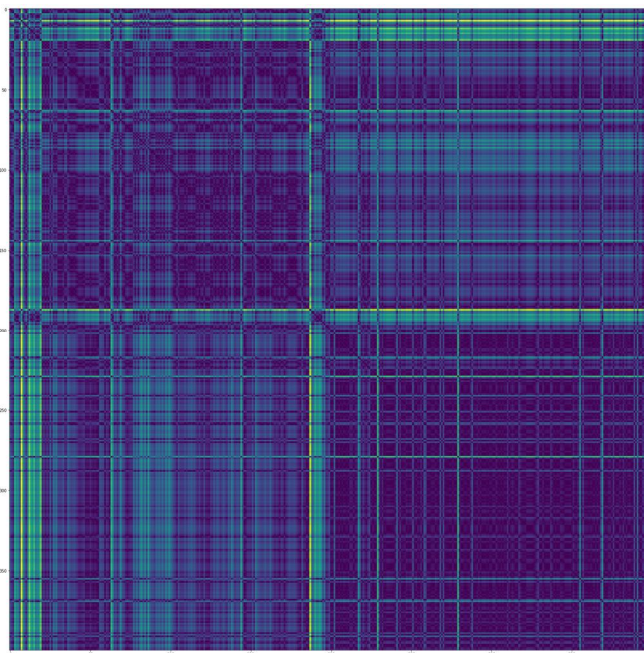


Рис. 7 Тепловая карта парных диффузных расстояний между гистограммами кривизн Риччи. Верхний левый угол - здоровые клетки против здоровых. Нижний правый угол - раковые клетки против раковых. Верхний правый и нижний левый - раковые клетки против здоровых. Темный цвет соответствует меньшему расстоянию, светлый большому расстоянию

Заключение

В статье обсуждаются подходы к машинному обучению онкопаттернов на гистологических образцах лимфатических узлов пациентов с онкологическими заболеваниями молочной железы.

Использовалась голландская база данных [1] содержащая полнослайдовые (WSI) изображения, образцов окрашенные контрастирующими веществами, обычно (H & E). Диагностика заключается в обнаружении макрометастазов, микрометастазов или изолированных опухолевых клеток. Последние две категории особенно трудны для визуального обнаружения. Так, для дозорных лимфатических узлов необходимо исследовать не менее трех срезов WSI изображений размером до 10 мегапикселей каждый, на разных уровнях детализации. Для недозорных узлов должен быть исследован один срез по крайней мере из 10 лимфатических узлов.

В качестве дескриптора WSI изображений мы использовали так называемую кривизну Формана–Риччи, известную из дискретной дифференциальной геометрии. Она содержит глобальную и локальную информацию о паттернах, содержащихся в фотометрической мере. Выходным продуктом оценок являлась гистограмма распределения кривизна для ребер пиксельного изображения. Для сравнения гистограмм использовались два метода. В первом из них вычислялись 4 первых центральных момента гистограмм. Во втором, проводилась оценка так называемого диффузного расстояния между гистограммами. Распознавание считалось корректным, если разность расстояний между двумя группами – нормальные и метастазированные образцы, отличалось от парных расстояний внутри каждой группы. По выборке, содержащей 200 образцов гистологических срезов, мы обнаружили, что предложенный нами метод может быть использован для эффективного машинного обнаружения онкопаттернов.

Мы благодарны за финансовую поддержку Института Информационных и Вычислительных Технологий (Гранты AP05132760, AP05134227 Казахстан).

Литература

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016;66(1):7–30.
2. Howlader N, Noone AM, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2014. National Cancer Institute, Bethesda, MD. http://seer.cancer.gov/csr/1975_2014/ based on November 2016 SEER data submission, posted to the SEER web site, April 2017; http://seer.cancer.gov/csr/1975_2014/.
3. Amin MB, Edge SB, Greene FL et al. AJCC Cancer Staging Manual. Springer-Verlag GmbH; 2016. http://www.ebook.de/de/product/26196032/ajcc_cancer_staging_manual.html.
4. Voogd AC, Nielsen M, Peterse JL et al. Differences in risk factors for local and distant recurrence after breast-conserving therapy or mastectomy for stage I and II breast cancer: pooled results of two large European randomized trials. J Clin Oncol. 2001; 19:1688–97.
5. Giuliano AE, Hunt KK, Ballman KV et al. Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial. JAMA. 2011; 305:569–75.
6. Giuliano AE, Ballman KV, McCall L, et al. Effect of axillary dissection vs no axillary dissection on 10-year overall survival among women with invasive breast cancer and sentinel node metastasis: the ACOSOG Z0011 (Alliance) randomized clinical trial. JAMA. 2017; 318:918–26.

7. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of. *Ann Surg Oncol.* 2010; 17:1471–4.
8. Weaver DL. Pathology evaluation of sentinel lymph nodes in breast cancer: protocol recommendations and rationale. *Mod Pathol.* 2010;23(Suppl 2): S26–32.
9. Somner JEA, Dixon JMJ, Thomas JSJ. Node retrieval in axillary lymph node dissections: recommendations for minimum numbers to be confident about node negative status. *J Clin Pathol.* 2004; 57:845–8.
10. van Diest PJ, van Deurzen CHM, Cserni G. Pathology issues related to SN procedures and increased detection of micrometastases and isolated tumor cells. *Breast Disease.* 2010; 31:65–81.
11. Gorban A. N., Tyukin I. Y. Blessing of dimensionality: mathematical foundations of the statistical physics of data // *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Science.* 2018. V. 376. №. 2118. P. 20170237.
12. Naitsat A, Saucan E, Zeevi Y.Y. A Differential Geometry Approach for Change Detection in Medical Images. // *CBMS 2017.* P.85-88
13. Saucan E, Weber M. Forman's Ricci Curvature - From Networks to Hypernetworks. // *Complex Networks 2018,* P. 706-717
14. Weber M., Saucan E, Jost J. Characterizing complex networks with Forman-Ricci curvature and associated geometric flows. // *J. Complex Networks 2017,* V.5(4) P.527-550
15. Geert Litjens Jeroen van der Laak Oscar Geessink Peter Bandi. CAMELION17 Dataset. 2017. url: <https://camelyon17.grand-challenge.org> (visited on 11/10/2018).
16. E. Saucan, E. Appleboim, G. Wolanski, and Y. Zeevi. Combinatorial ricci curvature for image processing. // *MICCAI 2008 Workshop Manifolds in Medical Imaging: Metrics, Learning and Beyond,* 2008.
17. Forman, R., *Discrete and Computational Geometry,* 2003, V.29, 323.
18. Saucan, E., et al. // *Proceedings of CISP'09,* Vol. 2, 992-997, (2009).
19. Weber E., Ehrenfeld, S., Littauer S. B. *Introduction to Statistical Method.* McGraw-Hill Book Comp., New York, 1964. 533 S.
20. Ling H., Okada K. Diffusion Distance for Histogram Comparison // *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06).* V.1. 2006, pp. 246.

АВТОМАТИЧЕСКОЕ ФОРМИРОВАНИЕ ТЕМАТИЧЕСКИХ СЛОВАРЕЙ СОЦИАЛЬНО-ЗНАЧИМЫХ ПОНЯТИЙ

Мейрамбеккызы Ж., Аманбай А., Мусабаев Р.

e-mail: jazirahanim@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Аннотация. В работе описывается проблема компьютерного анализа социально значимых тем в электронных новостях Казахстана. Предлагается выделение типов новостных тем для дальнейшей более точной автоматической

классификации текстов по темам. Для кластеризации текстов по темам существует два способа: словарный и метод машинного обучения. В данной работе выбран первый способ и приведен метод разработки тематических словарей, построенный на анализе семантической схожести с помощью матрицы *co-occurrence* и инструмента *Word2Vec*.

Введение

Еще в 60—70-х годах XX века составлялись словари разных типов с помощью ЭВМ, которые применялись для сортировки разных материалов. В 80-е годы XX века проводились работы по созданию словарных баз данных и разрабатывались другие виды автоматических словарей (тезаурусы). Как такого единого термина нет, используются разные обозначения как «машинный словарь», «компьютерный словарь», «электронный», «автоматический» в равных значениях [1]. Для разработки словаря любого типа необходимо создать общую концепцию, его структуру, принципы, по которым осуществится организация материала. Для общей задачи кластеризации новостных данных и распределении по темам, необходимы так называемые тематические словари. Каждый словарь имеет набор слов и понятий, относящийся к одной теме или подтеме. Многие сферы жизни сложно определить однозначно в одну тему, потому есть необходимость построить иерархию тем. Например, *Казахстан*, *премьер-министр* являются более универсальными, потому не относятся к конкретной теме. А *футбол*, *Головкин*, *Олимпийские игры* и т.д. включены в общую тему «Спорт». Для описания содержания предметных областей и тем в новостных текстах необходим процесс кластеризации текстов по темам. Тезаурусы, словари позволят более полно представить поисковое содержание кластера. Необходима технология формирования концептуальных словарей по корпусу текстов, по которой будут разработаны словари по темам.

Структура тематических словарей

Перед нами стоит задача формировать тематические словари с социально значимыми понятиями. Известно, что общественное мнение в основном формируется благодаря активной деятельности средств массовой информации. Не все социально значимые проблемы могут освещаться в новостях, потому читатели могут не подозревать об их существовании. В итоге, у населения складывается искаженная и неполная картина о действительно существующих социально значимых проблемах. Необходимо отметить, что термин "социально" в данном контексте употребляется не в широком смысле, означающее все, что относится к обществу. В наших задачах в узком смысле, как понятие, охватывающее только определенную часть общественных явлений. Например, разновидность общественных сфер – экономическая, социологическая, политическая и др. Как отмечают социологи понятие «социальной значимости» используется с разными значениями. К примеру, в статье [2] *образование* и *здравоохранение* отнесены к социально значимым, только опираясь на научные определения этих сфер. Однако автор подчеркивает, что среди исследователей-социологов нет единого научного определения понятия. Исследователь выделяет главные критерии социально значимых тем: (1) данное понятие выражает важность(значимость) ряда социальных сфер, проблем, (2) важна целевая направленность, адресность социально значимых проблем/тем. Автор приводит в пример социально значимые *образование*, *здравоохранение* являющиеся значимыми как для общества, так и для индивидуальных граждан. Однако для нашей задачи выделения социально

значимых тем, важен показатель того, что данная тема направлена на реализацию интересов общества и его ценностных установок. Таким образом, в нашем исследовании мы применим следующее определение. *Социальная значимость новостей* – новости, нацеленные на обозрение актуальных социальных проблем общества, в которых сообщается о решении какой-либо проблемы или события, важные для общества (*новая госпрограмма «Әлеуметтік қамқорлық», новость о гибели 5-детей в столице, митинги в день собрания «Нур Отан», выборы*). Предполагается, что многие социально значимые темы, новости должны быть резонансными, с большим количеством просмотров, комментариев, т.к. в статье затрагиваются проблемные темы для общества. Однако надо учесть, что в резонансные новости также входят скандальные события, новости шоу-бизнеса, которые, по сути, не являются социально значимыми. Поэтому среди множества резонансных новостей необходимо отличать социально значимые (*выборы, менингит, пенсионная реформа*) и скандально-событийные (*истории с Баян Есентаевой, шоу-бизнес и др*). Резонансные статьи - новости, вызвавшие общественный интерес, зачастую скандальные события. Пример **резонансного и одновременно социально-значимого** события: все новости о *смерти Д.Тэна*. Социальная значимость появилась после события. До происшествия темы «фигурное катание», «Денис Тэн» не были социально значимыми, поэтому данная тематика относится к динамичным. При обработке новостных данных и применении алгоритмов для тематического моделирования необходимо учесть эти 2 группы новостных тем.

Предлагается следующий алгоритм действий выявления социально-значимых новостей:



Прежде чем разработать способ, распределяющий в тексте социально значимые новости и темы в отдельные кластеры, необходимо сформировать определенный их перечень. В этой задаче, как один из вариантов решения, мы опираемся на социологические исследования по Казахстану. По данным ОФ «Центр социальных и политических исследований «Стратегия»⁵, можно выделить следующие 15 основных социально значимых направлений. Как видно на графике, их значимость меняется по годам. Для нашей технической задачи необходимо составить набор понятий, лемм, которые войдут в данные направления, поскольку в автоматической обработке текста процесс поиска происходит по формальным признакам, чем больше понятий можно включить в одну тему, тем глубже

⁵ <http://eurasiamonitor.org/>

результаты поиска. Например, к социально значимой теме «рост цен, инфляция» в тематические словари мы также включаем близкие понятия: *обесценивание, высокие цены, повышение цен* и тд.

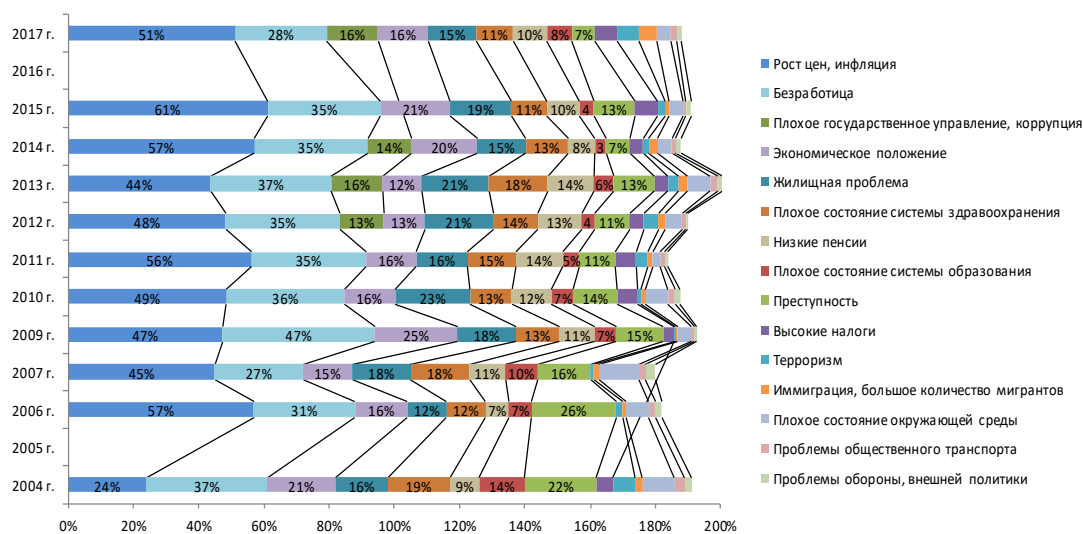


Рис. 1. Статистика по данным ЦСПИ «Стратегия»

Все указанные социально значимые проблемы мы включаем в группу *статичных тем*. Опираясь на указанное исследование видно, как данные социально значимые темы сохраняются на протяжении долгого периода с небольшими изменениями в значимости для населения в разные годы.

Помимо набора социально значимых тем, проблем, упомянутых в новостях, необходимо выяснить какую роль имеют электронные новостные порталы в получении информации для казахстанских граждан. По исследованию группы социологов, предпочтения казахстанцев каналов информации с 2004 по 2017 гг. привели к следующим результатам. Ключевой вопрос заключался в Таблице 1.

Таблица 1. Распределение ответов на вопрос "Какими каналами информации Вы пользуетесь чаще всего для получения сведений о событиях, происходящих в нашей стране?" (%)

	2004 г.	2008 г.	2012 г.	2017 г.
Передачи республиканского телевидения	78%	57%	37%	63%
Передачи местного телевидения	30%	48%	30%	17%
Центральные газеты	36%	13%	10%	6%
Местные газеты	32%	29%	18%	8%
Радио	33%	16%	4%	9%
Разговоры с родственниками, знакомыми	18%	26%	23%	31%
Интернет	4%	9%	31%	38%
Другое	1%	0%	0%	1%
Затрудняюсь ответить	1%	1%	1%	1%

В данных виден рост по годам интересов казахстанцев к интернету как к источнику информации происходящих событий, что включает также электронные новостные порталы. В дальнейших наших исследованиях отдельно будут рассмотрены исследования по 2018-2019 годам. В данный момент в проекте

имеется корпус новостей с глубиной в 7 лет в объеме 2 млн.300 тысяч документов разных новостных порталов. Все последующие работы и эксперименты будут проводиться на данном большом корпусе.

Наша общая задача заключается в автоматическом выделении новостей по темам без участия экспертных оценок, подключения социологов, аннотаторов. Потому перед нами стоит задача заранее выявить важные свойства новостных данных для учета в работе кластеризации. Таким образом, выдвигаем следующие гипотезы: (1) частотность темы/подтем в одном кластере может определять степень резонансности данной темы; (2) Из множества резонансных тем/тематик предполагается выделить социально-значимые, на основе тематических словарей. (3) Социально значимые темы/подтемы внутри кластера должны делиться на статические и динамические. В данном исследовании эксперименты мы начали с разработки тематических словарей.

Эксперименты

I Формирование тематических словарей на основе co-occurrence матрицы

В основу работы взят корпус из 200.000 новостных статей tengrnews. Существуют два подхода к решению задачи кластеризации по темам: словари и алгоритмы машинного обучения. В данной статье мы приводим несколько способов автоматического формирования тематических словарей. В работе [5] приведены результаты гибридного способа, являющийся более точным. В дальнейшем планируется сравнить оба метода на базе нашего корпуса казахстанских новостей. Были сформированы тематические словари на основе вычисленной матрицы по методу *co-occurrence*. Словари были обработаны с помощью языка Python (библиотека NLTK), почищены от лишних символов, приведены в нормальную форму. Для лемматизации мы применили Mystem. В отличие от его популярного аналога – Rymorhy, первый инструмент лучше распознает именованные сущности, свойственные казахстанскому содержанию. Однако данный метод автоматической разработки словаря требует дальнейшего развития, поскольку он требует ручной чистки. Из словаря в размере 1000 слов только 20% процентов относятся к конкретной теме. Метод встречаемости слов *co-occurrence* вычисляет как слова встречаются вместе, т.е. фиксирует связь между словами. Статистика встречаемости основывается на простой идее – как два и более слов связаны в одном контексте. Например: $count(W_{next} | W_{current})$ представляет, сколько раз слово W_{next} следует за словом $W_{current}$, мы можем суммировать статистику совместного появления для слов "Головкин" и "бокс" как:

Таблица 2. Матрица смежности слов

	спорт	бокс	аэропорт	театр	Казахстан	Головкин	Геннадий
Головкин	8	9	1	0	6	0	10
бокс	3	0	0	0	2	8	7

В таблице 2 показана частота биграмм – количество встречаемости 2 парных слов. Такая таблица очень разрежена, так как большинство частот равно нулю. На практике подсчеты встречаемости преобразуются в вероятности. Это приводит к тому, что записи строк для каждой строки складываются по одной в матрице совместного вхождения. В дальнейшем будут проведены исследования со словосочетаниями, например, мы можем подсчитать, количество встречаемости

пары слов в одном предложении независимо от их положения в предложениях. На данный момент имеются предварительные результаты формирования словарей социально значимых тем по (1) способу: «льготы» (1000), «пенсия» (500), политика: «президент» (1000), по «спорту» (700), «культуре» (500). Также для задачи распределения новостей по темам применяются готовые тематические словари в открытом доступе.

II Тематические словари с на основе word2vec

Для извлечения тематических слов также важен эксперимент по кластеризации текста с помощью word2vec. В отличие от co-occurrence, данный метод рассматривает не близость слов, а связь между контекстами. Как отмечено в работе [6] важно увеличить косинусную близость между векторами слов похожих контекстов и уменьшить косинусную близость слов из разных контекстов. Однако точность кластеризации слабая, поэтому слова одной темы не всегда собирались в одном наборе. В одном кластере могут попасть статьи, не относящиеся к данной теме. Это проблема не только несовершенства алгоритма, но и сложности структуры текстов СМИ. Одна новостная статья может заключать в себе темы ДТП, правонарушения, убийства, происшествий, политику.

III Формирование словаря именованных сущностей с помощью Regular Expressions

Данный лингвистический способ формирования словаря с помощью Regular Expressions позволяет из корпуса выделить имена собственные – так называемые именованные сущности. Данный словарь необходим для задачи отделения фамилий, имен, названий географических объектов. К примеру, именованная сущность «министр» может выражаться в самых разных формах: *министр образования, Ерлан Сагадиев, Сагадиев, министр образования и науки, министр Сагадиев, министр МОН РК.*

Заключение

В данном исследовании мы пытались показать разные способы автоматического формирования тематических словарей. Для начала мы формировали тематические словари распространенных тем: спорт, политика. Следующим этапом перед нами стояла задача выделить социально значимые понятия и сформировать их словари. Для точности понятия социальной значимости мы опирались на определение и результаты социологических опросов специалистов. Для формирования словаря применили 3 способа. В перспективе данная работа разработки тематических словарей будет продолжен. Словари будут применены для кластеризации текста с помощью алгоритмов машинного обучения.

Работа выполнена в рамках проекта ПЦФ №BR05236839 «Разработка информационных технологий и систем для стимулирования устойчивого развития личности как одна из основ развития цифрового Казахстана»

Литература

1. Ю.Е. ВОЛКОВ. СОЦИАЛЬНЫЕ ОТНОШЕНИЯ И СОЦИАЛЬНАЯ СФЕРА Журнал «Социальная политика. Социальная структура» №5, 2003
2. Т.Д. Бурменко, Е.Г. Тарханова Состояние сферы социально значимых услуг в мире (на примере здравоохранения и образования) / Журнал «Социально-экономическое развитие и образование». 2013

3. Сидорова Е.А. Российский НИИ Искусственного Интеллекта, г. Новосибирск. ТЕХНОЛОГИЯ РАЗРАБОТКИ ТЕМАТИЧЕСКИХ СЛОВАРЕЙ НА ОСНОВЕ СОЧЕТАНИЯ ЛИНГВИСТИЧЕСКИХ И СТАТИСТИЧЕСКИХ МЕТОДОВ. Диалог-2005. МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ по компьютерной лингвистике и интеллектуальным технологиям/Москва/ 1-7 июня 2015

4. Е.И. Большакова. О ПРИНЦИПАХ ПОСТРОЕНИЯ КОМПЬЮТЕРНОГО СЛОВАРЯ ОБЩЕНАУЧНОЙ ЛЕКСИКИ/ Диалог-2002/ МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ по компьютерной лингвистике и интеллектуальным технологиям/ Москва/

5. А.А. Тузовский. РАЗРАБОТКА АЛГОРИТМА АВТОМАТИЧЕСКОЙ КАТЕГОРИЗАЦИИ ТЕКСТА НА ОСНОВЕ КОМБИНАЦИИ СЛОВАРЕЙ, ПРАВИЛ И МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ/Сборник материалов XIV Международной школы-конференции студентов, аспирантов и молодых ученых, С.412-415., 2018 г.

6. Левченко С.В. Разработка метода кластеризации слов по смысловым характеристикам с использованием алгоритмов Word2Vec/Журнал «НОВЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ», 2017 Г

МЕТОД МУЛЬТИМОДАЛЬНОЙ ОЦЕНКИ СРЕДСТВ МАССОВОЙ ИНФОРМАЦИИ НА ОСНОВЕ ТЕМАТИЧЕСКОЙ МОДЕЛИ КОРПУСА ТЕКСТОВ

Мусабаев Р., Мухамедиев Р., Кучин Я., Сымагулов А., Якунин К.,
Мурзахметов С.

e-mail: ravil.muhamedyev@gmail.com

*Институт информационных и вычислительных технологий МОН РК,
Алматы, Республика Казахстан*

***Аннотация.** Предложен метод расчета вероятностей распределения средств массовой информации (СМИ) по темам, признакам и классам на основе тематической модели корпуса текстов СМИ. Суть метода заключается в использовании тематической модели, представленной матрицей условных вероятностей распределения документов СМИ по темам, для получения распределений условных вероятностей СМИ по темам, а также, классам и признакам, определяемым экспертами. Описан алгоритм расчета, основанный на байесовских правилах и результат его работы на небольшом модельном примере.*

***Ключевые слова:** обработка естественного языка, тематическое моделирование, правило Байеса*

Введение

В настоящее время достигнут значительный прогресс в моделях и методах обработки текстов на естественном языке. Обработка естественного языка (NLP) как область исследований включает в себя широкий спектр прикладных разделов:

–автоматический перевод [1],

- автоматическое реферирование,
- генерация ответов на пользовательские запросы (Question answering) [2].
- извлечение информации (IE) [3],
- поиск информации [4] [5].
- анализ тональности [6],
- другие области, которые так или иначе связаны с обработкой устной и письменной естественной речи.

Потребности в решении практических задач NLP повлияли на развитие многих методов обработки данных, среди которых можно упомянуть машинное обучение, нейронные сети, глубокое обучение [7], нечеткую логику, логику первого порядка, семантические сети и другие.

В последнее время научные исследования, которые поддерживаются ростом вычислительной мощности, привели к ряду прорывных результатов в области NLP, среди которых успехи в области машинного перевода, автоматического суммирования, поиска информации, ответа на вопросы, анализа тональности, извлечение информации [8], определение авторства [9].

Ключевыми аспектами, позволившими получить впечатляющие результаты в области автоматической обработки текстов на естественном языке, являются, согласно [10], достижения в развитии методов машинного обучения, многократное увеличение вычислительной мощности, наличие большого объема лингвистических данных и развитие понимания структуры естественного языка в приложении к социальному контексту.

NLP как область исследований меняется крайне динамично. Со времени работы [11] качественно новые результаты были получены при разработке статистических моделей языка. Большие объемы доступных текстов в социальных сетях и использование глубоких нейронных сетей [12] приводят к формулированию задач извлечения образов из огромного количества неструктурированной информации на основе современных методов распределенной лингвистики и так называемого контролируемого обучения [10].

Одним из методов, продуктивно применяемых в области NLP, является тематический анализ или тематическое моделирование. Тематическое моделирование – метод, основанный на статистических характеристиках коллекций документов, который применяется в задачах автоматического реферирования, извлечения информации, информационного поиска и классификации [13]. Смысл данного подхода основан на интуитивном понимании того, что документы в коллекции образуют группы, в которых частота встречаемости слов или сочетаний слов различается.

Расцвет данного направления исследований пришелся на 2012-2013 годы после чего в 2018 году количество публикаций с термином «тематическое моделирование» уменьшилось более чем вдвое (166000 в 2012 и 61800 в 2018) по данным google scholar [14].

Использование кластеров документов, относящихся к совокупности тем, позволяет решать задачи синонимии и полисемии терминов [15]. Вероятностные тематические модели описывают документы (M) дискретным распределением на множестве тем (T), а темы дискретным распределением на множестве терминов [16]. В результате построения тематической модели формируется матрица условных вероятностей, далее обозначаемая как $p_2(k|m)$, где $k \in T$, $m \in M$.

В настоящей работе тематическая модель использована для расчета условных распределений вероятности средств массовой информации по темам, классам и

признакам. Другими словами, предлагается мультимодальная оценка средств массовой информации (СМИ), основанная на первоначальном распределении документов по темам. Предложенный подход примечателен тем, что тематическая модель, созданная путем кластерного анализа (обучение без учителя) затем применена в сочетании с экспертно определяемыми классами и признаками. Другим словами, семантику требуемого распределения задает пользователь, хотя изначальный тематический анализ зависит только от корпуса документов.

Алгоритм мультимодального оценивания СМИ

Цель: Используя распределение вероятностей документов корпуса агрегировать показатели соответствия статьи тематикам, тематик признакам (словарям) и классам для получения оценок соответствия средств массовой информации (СМИ) в трех модальностях: тематикам, признакам и классам.

Метод: Для оценки указанных соответствий используется вероятностный и байесовский подходы в предположении, что вероятности статей, тематик, классов и признаков взаимно не зависимы.

Ожидаемый результат: Оценка "принадлежности" СМИ к тематикам, признакам и классам в виде распределений вероятностей.

Краткое описание алгоритма: Используя множество тематик корпуса, во первых, получаем дискретное распределение вероятности статей по тематикам (p2). Во-вторых, получаем распределение словарей по тематикам (p1), то есть, определяем, в какой степени словарь описывает конкретную тематику. В третьих, с помощью аналитического иерархического процесса (АИР) рассчитываем важность словарей для классов (отдельно для каждого класса) (p3). Затем, используя p1 и p3, рассчитываем условное распределение тематик по классам (p4). Зная распределение вероятностей тематики по классам (p4) и распределение вероятностей статьи по тематикам (p2) можно вычислить распределение статьи по классам (p5). В свою очередь, распределение статьи по признакам или словарям (p6) зависит от p1 и p2. Исходные данные и получаемые матрицы условных вероятностей показаны на рисунке. Корпус документов описывается словарем (Corpus Dictionary). СМИ (MMS) являются источником m документов. (Рис. 1)

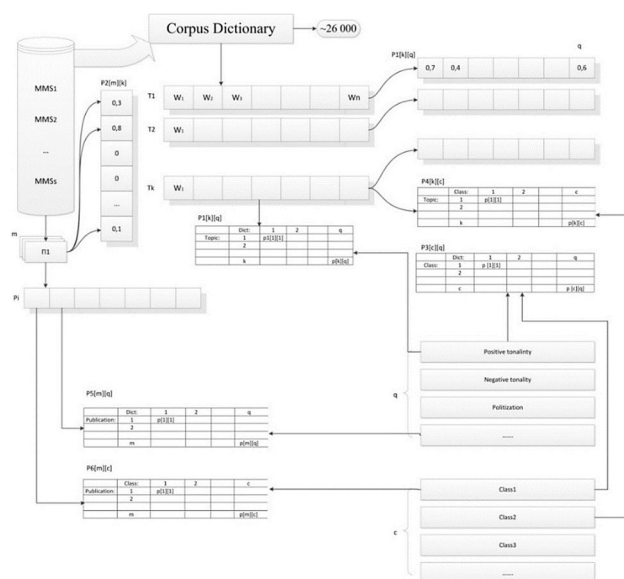


Рисунок 1. Процессы определения условных вероятностей

После получения указанных оценок для каждой статьи, производится оценка СМИ с использованием цепочки байесовских правил.

В соответствии с формулой Байеса мы можем выразить условную вероятность справедливости гипотезы h при наступлении события e в виде

$$p(h|e) = \frac{p(e|h) \times p(h)}{p(e|h) \times p(h) + p(e|\sim h) \times p(\sim h)} \quad (B1)$$

где $p(e|h)$ – условная вероятность наступления события e при справедливости h , $p(h)$ – априорная вероятность гипотезы h , $p(e|\sim h)$ – условная вероятность e при несправедливости h , $p(\sim h)$ – вероятность того, что событие h не верно, которое, в соответствии с формулой полной вероятности, может быть вычислено как

$$p(\sim h) = 1 - p(h) \quad (B2)$$

Таким образом, для вычисления условной вероятности $p(h|e)$ достаточно знать вероятности $p(e|h)$, $p(e|\sim h)$ и априорную вероятность $p(h)$.

В терминах задачи оценки СМИ, условные вероятности $p(e|h)$ могут быть интерпретированы как вероятности появления некоторой статьи $e \in E$ (где E – множество статей конкретного СМИ) при условии справедливости гипотез трех видов (каждая гипотеза для своей модальности):

1. h_1 : СМИ работает в одной из тематик из множества тематик T : $H_1 = \{h_1[1], \dots, h_1[k]\}$

2. h_2 : СМИ публикует статьи, относящиеся к классу из множества C : $H_2 = \{h_2[1], \dots, h_2[c]\}$

3. h_3 : СМИ отличается признаками из Q : $H_3 = \{h_3[1], \dots, h_3[q]\}$

Рассмотрим первую задачу, заключающуюся в определении предпочтительных тематик СМИ.

Тематическая ориентированность СМИ выражается в публикуемых статьях. Статья e_i появляется в зависимости от того работает ли СМИ в тематике $t_j \in T$, где $t_j = (1, \dots, k)$

Положим, что $p(e_i|h_1(t_j)) = p_2[t_j][e_i]$, где p_2 – дискретное распределение условных вероятностей тематик по документам корпуса.

В свою очередь, вероятность появления статьи данной тематики при справедливости гипотезы о том, что СМИ не работает в тематике k_j , примем равной $p(e_i|\sim h_1) = 0.5$, что по существу означает, что у нас нет сведений. Априорную вероятность $p(h_1(t_j))$ примем равновероятной для всех тематик.

$$p(h_1(t_j)) = 1/k,$$

и, в соответствии с выражением (B2)

$$p(\sim h_1(t_j)) = 1 - p(h_1(t_j))$$

Аналогично допустим

$p(e_i|h_2(c_j)) = p_5[e_i][c_j]$, где p_5 – условные вероятности распределения документов по классам

$$p(h_2(c_j)) = 1/c_j,$$

$$p(\sim h_2(c_j)) = 1 - p(h_2(c_j))$$

$$p(e_i|\sim h_2(c_j)) = 0.5$$

Также

$$p(e_i|h_3(q_j)) = p_6[e_i][q_j],$$

где p_6 - условные вероятности распределения документов по признакам

$$p(h_3(q_j))=1/q,$$

$$p(\sim h_3(q_j))=1- p(h_3(q_j))$$

$$p(e_i|\sim h_3(q_j))=0.5.$$

Опуская номер конкретной тематики (t_j), класса (c_j) и признака (q_j) для указанных гипотез $h_1 \in H_1$, $h_2 \in H_2$, $h_3 \in H_3$, получим следующий

Алгоритм мультимодальной оценки СМИ

Для всех $h_1 \in H_1$

Для всех $e_i \in E$

$$p(h_1|e_i) = \frac{p(e_i|h_1) \times p(h_1)}{p(e_i|h_1) \times p(h_1) + p(e_i|\sim h_1) \times p(\sim h_1)}$$

$$p(h_1) := p(h_1|e_i)$$

$$p(\sim h_1) := 1 - p(h_1|e_i)$$

Для всех $h_2 \in H_2$

Для всех $e_i \in E$

$$p(h_2|e_i) = \frac{p(e_i|h_2) \times p(h_2)}{p(e_i|h_2) \times p(h_2) + p(e_i|\sim h_2) \times p(\sim h_2)}$$

$$p(h_2) := p(h_2|e_i)$$

$$p(\sim h_2) := 1 - p(h_2|e_i)$$

Для всех $h_3 \in H_3$

Для всех $e_i \in E$

$$p(h_3|e_i) = \frac{p(e_i|h_3) \times p(h_3)}{p(e_i|h_3) \times p(h_3) + p(e_i|\sim h_3) \times p(\sim h_3)}$$

$$p(h_3) := p(h_3|e_i)$$

$$p(\sim h_3) := 1 - p(h_3|e_i)$$

В результате выполнения алгоритма получим множество $\Psi_i(H) = \{\Psi_i(H_1), \Psi_i(H_2), \Psi_i(H_3)\}$, для конкретного СМИ s_i где

$$\Psi_i(H_1) = \{ p(h_1[1]), \dots, p(h_1[k]) \}$$

$$\Psi_i(H_2) = \{ p(h_2[1]), \dots, p(h_2[c]) \}$$

$$\Psi_i(H_3) = \{ p(h_3[1]), \dots, p(h_3[q]) \}$$

Экспериментальный корпус текстов

С целью оценки применимости описанного алгоритма был сформирован очень небольшой корпус документов, условно распределенный между двумя СМИ. Для описанного корпуса были рассчитаны матрицы p_1, \dots, p_6 с применением меры Жаккара, довольно часто применяемой как мера близости документов в алгоритмах кластеризации [17,18]. Кратко исходные данные для работы алгоритма следующие:

Имеем $m=5$ статей следующей направленности:

1. экономика,
2. спорт(бокс),
3. политика,
4. шоу бизнес,
5. образование и наука.

Корпус текстов разбит на следующие тематические кластеры ($k=4$)

1. политика,
2. спорт,
3. шоу бизнес,
4. образование и наука.

Рассматриваются два класса статей и СМИ ($c=2$)

1. социально значимый
2. объективный

Определены следующие признаки($q=4$), для каждого из которых определен свой словарь терминов.

1. манипулятивность,
2. политизированность,
3. негативная тональность,
4. позитивная тональность.

Определена матрица условных распределения тематик по статьям в виде матрицы $p2[1..k][1..m]$, где k - количество тематик (построчно), m - статей (по столбцам)

$$p2 = \begin{bmatrix} 0.03, & 0, & 0.03, & 0.004, & 0.005 \\ 0, & 0.01, & 0, & 0.008, & 0.005 \\ 0, & 0.007, & 0.004, & 0.03, & 0.005 \\ 0.04, & 0, & 0.007, & 0, & 0.005 \end{bmatrix}$$

С использованием аналитического иерархического процесса (АНР) получены веса каждого признака для классов в виде матрицы $p3[1..c][1..q]$, где c – классы, q - признаки

$$p3 = \begin{bmatrix} 0.55, & 0.27, & 0.18, & 0.18 \\ 0.23, & 0.43, & 0.34, & 0.34 \end{bmatrix}$$

Получены следующие результаты расчета матриц $p4[1..k][1..c]$ (условные вероятности распределения тематик по классам), $p5[1..m][1..c]$ (условные вероятности распределения документов по классам), $p6[1..m][1..q]$ (условные вероятности распределения документов по признакам):

$$p4 = \begin{bmatrix} 0.1967 & 0.2239 \\ 0.0371 & 0.0531 \\ 0.0316 & 0.0508 \\ 0.0427 & 0.0443 \end{bmatrix}$$

$$p5 = \begin{bmatrix} 0.007609 & 0.008489 \\ 0.0005922 & 0.0008866 \\ 0.0063263 & 0.0072303 \\ 0.0020316 & 0.0028444 \\ 0.0015405 & 0.0018605 \end{bmatrix}$$

$$p_6 = \begin{bmatrix} 0.0058 & 0.0149 & 0.0014 & 0.0008 \\ 0.00027 & 0.00017 & 0.00055 & 0.00166 \\ 0.00452 & 0.01329 & 0.00094 & 0.00046 \\ 0.00102 & 0.0021 & 0.00174 & 0.00328 \\ 0.00105 & 0.0025 & 0.00055 & 0.00105 \end{bmatrix}$$

Заметим, что процессы формирования словарей признаков, выбор классов для признаков, формирование таблиц АНР и т.п. в данной работе не приводятся.

Результаты

На основании приведенных значений с применением алгоритма мультимодальной оценки СМИ рассчитаны распределения СМИ по тематикам, классам и признакам в предположении, что имеющиеся статьи предоставлены двумя СМИ, первое из которых опубликовало статьи #1 и #3 (экономической и политической направленности), второе является источником статей 2,4,5 (“спорт(бокс)”, “шоу бизнес”, “образование и наука”).

Распределение СМИ по тематикам

СМИ\Тематики:	Политика	Спорт	Шоу бизнес	Экономика и финансы	Статьи
СМИ0	0.047109	0.000000	0.003012	0.036757	[экономическая, политическая]
СМИ1	0.001514	0.006877	0.014395	0.000000	[спорт (бокс), шоу бизнес, образование и наука]

Распределение СМИ по классам

СМИ\Классы:	Социально значимый	Объективный	Статьи
СМИ0	0.0	0.000456	[экономическая, политическая]
СМИ1	0.0	0.000180	[спорт (бокс), шоу бизнес, образование и наука]

Распределение СМИ по признакам

СМИ\Признаки:	Манипулятивность	Политизированность	Негативная тональность	Позитивная тональность	Статьи
СМИ0	0.006855	0.020653	0.000812	0.000000	[экономическая, политическая]
СМИ1	0.000000	0.000916	0.000188	0.001377	[спорт (бокс), шоу бизнес, образование и наука]

Заключение

Описанный алгоритм мультимодального оценивания СМИ на основе публикуемых статей показал логичные результаты на небольшом модельном примере, кратко описанном в статье. Полученные результаты позволяют рассчитывать, что на больших корпусах алгоритм позволит получить требуемые

оценки СМИ путем агрегирования оценок принадлежащего им массива документов. Очевидно, что точность алгоритма критически зависит от алгоритма получения условных распределений статей по темам, тем по словарям и классам. Методы получения указанных распределений предполагается рассмотреть в следующей работе.

Работа финансировалась за счет гранта BR05236839 Министерства образования и науки Республики Казахстан «Разработка информационных технологий и систем для стимулирования устойчивого развития личности как одна из основ развития цифрового Казахстана».

Литература

1. Sreelekha, S., Pushpak Bhattacharyya, Shishir K. Jha, and D. Malathi. "A survey report on evolution of machine translation." *Int. J. Control Theory Appl* 9, no. 33 (2016): 233-240. - Вработеописывается история развития автоматического перевода
2. Höffner, Konrad, Sebastian Walter, Edgard Marx, Ricardo Usbeck, Jens Lehmann, and Axel-Cyrille Ngonga Ngomo. "Survey on challenges of question answering in the semantic web." *Semantic Web* 8, no. 6 (2017): 895-920.
3. Daniel Jurafsky and James H. Martin. 2009. *Speech and Language Processing* (2Nd Edition). Prentice-Hall, Inc., Upper Saddle River, NJ, USA
4. Deo, Arpit, Jayesh Gangrade, and Shweta Gangrade. "A SURVEY PAPER ON INFORMATION RETRIEVAL SYSTEM." *International Journal of Advanced Research in Computer Science* 9, no. 1 (2018).
5. Шокин, Юрий Иванович, Анатолий Михайлович Федотов, and Владимир Борисович Барахнин. "Проблемы поиска информации." (2010).
6. Sun, Shiliang, Chen Luo, and Junyu Chen. "A review of natural language processing techniques for opinion mining systems." *Information Fusion* 36 (2017): 10-25.
7. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521, no. 7553 (2015): 436.
8. Hogenboom F. et al. A survey of event extraction methods from text for decision support systems // *Decision Support Systems*. – 2016. – Т. 85. – С. 12-22.
9. Potthast M., Hagen M., Stein B. Author Obfuscation: Attacking the State of the Art in Authorship Verification // *CLEF (Working Notes)*. – 2016. – С. 716-749.
10. Hirschberg J., Manning C. D. *Advances in natural language processing* // *Science*. – 2015. – Т. 349. – №. 6245. – С. 261-266.
11. Manning, Christopher D., Christopher D. Manning, and Hinrich Schütze. *Foundations of statistical natural language processing*. MIT press, 1999.
12. Goldberg Y. A primer on neural network models for natural language processing // *Journal of Artificial Intelligence Research*. – 2016. – Т. 57. – С. 345-420.
13. И. В. Машечкин, М. И. Петровский, Д. В. Царёв, "Методы вычисления релевантности фрагментов текста на основе тематических моделей в задаче автоматического аннотирования", *Выч. мет. программирование*, 14:1 (2013), 91–1021
14. <https://scholar.google.com>
15. https://ru.wikipedia.org/wiki/Тематическое_моделирование
16. Воронцов К. В., Потапенко А. А. Регуляризация, робастность и разреженность вероятностных тематических моделей // *Компьютерные исследования и моделирование*. – 2012. – Т. 4. – №. 4. – С. 693-706.

17. Пархоменко П.А., Григорьев А.А., Астраханцев Н.А. Обзор и экспериментальное сравнение методов кластеризации текстов. Труды ИСП РАН, том 29, вып. 2, 2017 г., стр. 161-200. DOI: 10.15514/ISPRAS-2017-29(2)-6

18. O’Callaghan, Derek, Derek Greene, Joe Carthy, and Pádraig Cunningham. "An analysis of the coherence of descriptors in topic modeling." *Expert Systems with Applications* 42, no. 13 (2015): 5645-5657

РЕФЕРИРОВАНИЕ ТЕКСТОВОГО ДОКУМЕНТА С ПОМОЩЬЮ WORD MOVER’S DISTANCE И ИЗВЛЕЧЕННЫХ КЛЮЧЕВЫХ СЛОВ ДОКУМЕНТА

Мусабаев Р.Р., Сейткали Д.

e-mail: rustam@ipic.kz

*Институт информационных и вычислительных технологий КН МОН РК,
Алматы*

Аннотация. Целью данной работы является разработка метода автоматической суммаризации текстового документа. Основная идея метода заключается в использовании ключевых слов, определенных на основе плотного векторного представления слов (*word embeddings*) и нахождения центроида документа, и ранжирования предложений с помощью *Word Mover’s Distance*. Результаты *ROUGE* на наборе данных *DUC 2002* показали, что качество результата суммаризаций разработанного подхода находится на одном уровне с другими более сложными современными системами автоматического реферирования, а в некоторых случаях превосходит их. В этой работе также представлен анализ качества оценок суммаризации рассматриваемого набора данных, составленные экспертами.

Ключевые слова: обработка текстов естественных языков, реферирование текстового документа, *word embeddings*, *word mover’s distance*, центроид документа, ключевые слова документа, *DUC 2002*

1. Введение

Целью систем суммаризации текстовых документов является создание более сжатой (плотной) версии документа, которая должна содержать ключевую выжимку исходного текста без потери смысла документа, его темы. Данная задача предполагает воспроизведение способности людей в понимании документов, поэтому является сложной в решении. Существует два типа систем суммаризации документов: экстрактивные и абстрагирующие. В данной статье представлено экстрактивное обобщение документа, где сгенерированное резюме представляет собой подборку релевантных предложений исходного документа. Хороший метод экстрактивного суммирования должен обладать по возможности наиболее оптимальным образом свойством “охвата”, так и свойством “разнообразия”, при чем выбранные предложения должны избегать избыточности информации в резюме. Система экстрактивного реферирования обычно состоит из двух основных компонент. Во-первых, в ней должна быть заложена модель по представлению

предложений. Во-вторых, она должна иметь способность ранжировать предложения, основываясь на некоторой функции сходства.

Первоначально для этих целей в качестве представления использовалось множество слов (bag of words, BOW). Однако представление на основе BOW не позволяет в достаточной мере уловить смысловое сходство при сравнении предложений [1]. В последнее десятилетие изучение возможных представлений текстовой информации стало активной исследовательской темой [2], [3] в области обработки текстов естественного языка (natural language processing, NLP). В результате это привело к прогрессу при решении различных задач, требующих понимание информации в текстовом виде, например, в системах автоматического реферирования документов. В одном из направлений данных исследований широко стали использоваться нейронные сети для получения представления слов в вещественном векторном пространстве. Модели нейронных сетей, например, word2vec, путем кодирования семантического и синтаксического значения слов позволяют представлять слова, употребляемые в одном и том же контексте, схожим векторным представлением. Например, два предложения, в которых нет общих слов: «Трамп говорил с репортерами» и «Президент был опрошен журналистами», вероятно, будут рассматриваться как сходные по значению, если использовать модель word2vec. Представление же на основе BOW будет рассматривать их векторные (разреженные) представления как ортогональные, так как предложения не имеют общих слов.

Для представления предложения или документа широко распространенной практикой является усреднение векторных представлений всех слов предложения или документа. Применение метрики косинусного сходства к любой паре векторных представлений позволит количественно определить степень сходства между предложениями или документами. Хотя представление предложения с помощью усреднения представлений его слов дают сравнительно хорошие результаты при вычислениях сходства между предложениями, такой метод представления текста может не уловить основную тему семантического содержания, не определить ключевые слова в тексте. Разработанный метод Word Mover's Distance (WMD) наиболее качественно решает эту проблему по сравнению с другими имеющимися методами [4]. На данный момент вопрос о применении WMD к задаче автоматической суммаризации не изучен тщательно. Основной причиной этого может быть вычислительная сложность метода WMD.

В данной работе представляется система экстрактивной суммаризации документов, которая использует преимущества векторных представлений на основе нейронных сетей [2], так и метрики WMD. На первом этапе используются плотные векторные представления слов для извлечения определения центроида документа и ключевых слов документов. Далее WMD применяется для расчета сходства между предложениями и ключевыми словами документов. Для оценки качества работы разработанного метода приведены сравнительные результаты его работы с другими современными методами, основываясь на результатах ROUGE на наборе данных DUC-02.

2. Обзор схожих работ

2.1. Экстрактивная суммаризация

На сегодняшний день разработан ряд unsupervised алгоритмов суммаризации документов. Наиболее заметным является CP3 [5], который дает лучшие оценки ROUGE1 (R-1), ROUGE-2 (R-2) и ROUGE-SU4 (RSU4) по DUC-02 среди всех ранее разработанных алгоритмов. Среди них хорошо известные алгоритмы, такие

как TextRank [6], Mead [7], пара жадных алгоритмов, алгоритм динамического программирования DocSumm [8]. В статье [9], который используют извлечение ключевых слов для дальнейшей суммаризации. Хотя алгоритм CP3, по-видимому, является лучшей системой экстрактивной суммаризации, основанной на наборе данных DUC-02, он проигрывает другим алгоритмам с точки зрения эффективности. Основная причина этого заключается в использовании ILP для максимизации важности, отсутствия избыточности и согласованности предложений на основе шаблонов [10]. Вычисление ILP занимает много времени даже на документах среднего размера.

2.2. Word Mover's Distance

Суть подхода. В 2015 году Куснер (Kusner) и соавторы представили новый метод под названием Word Mover's Distance (WMD), который вычисляет расстояния между двумя предложениями [4] на основе транспортной задачи, где рассматривает взаимные расстояния между векторными представлениями слов и соответствующие им веса. WMD сначала определяет индивидуальное расстояние между парами слов w_i и w_j через расчет расстояния между их векторными представлениями $e(w_i)$ и $e(w_j)$:

$$d(w_i, w_j) = |e(w_i) - e(w_j)| \quad (1)$$

WMD формулирует степень сходства между парой предложений S и S' , рассчитывая минимальные совокупные суммы транспортных расходов для “переноса” слов в предложении S в слова, встречающиеся в предложении S' :

$$D_{WMD}(S, S') = \sum_{w_i \in S} \sum_{w_j \in S'} T_{w_i, w_j} \times d(w_i, w_j) \quad (2)$$

$$\sum_{w_j \in S'} T_{w_i, w_j} = p_i \quad (3)$$

$$\sum_{w_i \in S} T_{w_i, w_j} = p_j \quad (4)$$

где $T \in R_{n \times n}$ - матрица значения, которой указывают, сколько вероятностной массы должно “перенестись” от слова w_i в предложении S к слову w_j в другом предложении S' и наоборот, при этом каждое слово может быть перенесено целиком или по частям. Полное преобразование вероятностной массы S в S' обеспечивает условие (2). Кроме того, количество входящего потока для слова w_j должно соответствовать условию (3).

WMD в суммаризации. В двух источниках указывается об использовании WMD в отношении суммаризации. В них WMD используется для расчета попарных расстояний между предложениями при ранжировании предложений [11]. Их системы оцениваются на новостном корпусе и наборе научных данных, соответственно. В статье [10] предлагается использовать WMD в качестве инструмента измерения для оценки суммаризации.

3. Предложенный метод

В этой статье предлагается разработанный метод по ранжированию предложений для суммаризации документов. Основная идея состоит в выделении тех предложений, которые являются близкими к ключевым словам документов. Ключевые слова определяются как слова, которые наиболее близкие к центру масс представления документа (центроиду документа). В качестве метрики сходства

используется Word Mover's Distance (WMD). WMD является одной из лучших метрик для определения смыслового сходства текстов, но процесс вычисления ее является вычислительно затратным. Основной причиной этого является то, что WMD основана на решении транспортной задачи. Далее будет представлена система автоматической суммаризации, которая использует преимущества метрики WMD, при этом избегая попарных вычислений, применяет ее к измерению расстояния между предложениями и ключевыми словами документа.

3.1. Предобработка текста

Для апробации данного метода для представления слов в векторном пространстве использовалась предобученная модель word2vec на GoogleNews2. Предварительная обработка документов произведена была таким образом, чтобы максимизировать количество слов, которые могут быть найдены в словаре данной модели. Первым шагом было проведено разбиение документов на предложения, затем проведена токенизация и удаление stopwords. Понижение регистра всех слов осуществлялось с учетом того, что есть слова, которые хранятся в модели только в верхнем регистре.

3.2. Центроид документа

Метод экстрактивной суммаризации на основе центроидов был предложен в [12]. Центроид документа является своего рода абстрактным документом, который представляет единственным вектором наиболее значимую информацию исходного документа.

Рассмотрим документ с словарем V всех слов размером N , построим матрицу $E \in R_{n \times k}$, где i -й строка матрицы векторное представление i -го слова из V . Центроидом документа будет являться сумма векторных представлений всех уникальных слов в документе.

3.3. Выделение ключевых слов

Ключевые слова — это набор слов ограниченного количества, которые описывают документ. Предполагается, что рассчитанный центроид документа (на предыдущем этапе) представляет центральную тему документа, а слова (то есть их векторное представление) расположенных близко к данному центроиду могут образовывать набор ключевых слов документа. С целью их определения отранжируем все уникальные слова в документе, на основе косинусного сходства между каждым векторным представлением слова $e(w_i)$ и центроидом документа. Выберем 25% наиболее близких к центроиду документа уникальных его слов; они образуют ключевые слова документа.

3.4. Ранжирование предложений с помощью WMD

Для определения предложений документа, которые и будут являться суммаризацией документа, используем метрику сходства WMD, описанную в разделе 2.2:

$$U(S_i) = D_{WMD}(S_i, K)$$

где S_i – bag of words i -го предложения, K - набор ключевых слов.

Выбрав ближайшие предложения так, чтобы они по количеству были равны 100 слов (если предложения ровно не укладываются в сто слов, то последнее предложение обрезается), получим резюме документа.

4. Результаты

Для оценки качества разработанного метода суммаризации рассмотрим набор данных DUC 2002, который является де-факто набором данных по которому проводится оценки систем экстрактивной суммаризации с помощью метрики ROUGE. DUC2002 состоит из 567 документов; каждый документ имеет резюме, состоящее из 100 слов, подготовленных экспертами. Большинство документов содержат два резюме от двух разных экспертов.

Во многих работах по методам суммаризации, используется набор данных DUC и оценки качества их систем по ROUGE. Тем не менее, только в некоторых указаны недостатки генерируемых человеком суммаризаций в оценке ROUGE [13]. Оценим максимальное достижимое значение ROUGE с помощью систем автоматических суммаризаций, то есть сравним схожесть резюме одного эксперта от другого. Для этого используем набор данных DUC-02, где есть резюме от двух экспертов для 545 документов. В таблице 1 показаны результаты:

System	Recall	Precision	F-measure
gold to gold	0.50576	0.50566	0.50549

Table 1: ROUGE scores between two gold summaries on DUC-02 dataset.

Сравним результаты разработанной системы суммирования с результатами известных современных системы, по которым есть результаты ROUGE для набора данных DUC 2002 [14]. Из таблицы 1 видно, что наша система набрала 0,45582 балла по ROUGE F-1, обогнав TextRank, MEAD и четыре из семи лучших результатов по DUC-02.

System	Recall	Precision	F-measure
S28	.47813	.45779	.46729
S19	.45563	.47748	.46309
S21	.47543	.44635	.46029
Baseline	.47788	.44680	.46172
Our system	0.45747	0.45453	0.45582
TextRank	.46165	.43234	.44640
S29	.46100	.44557	.45269
S23	.43188	.47585	.45018
S27	.45485	.44808	.45014
MEAD	.44506	.45290	.44729
s15	.44805	.43323	.44014

Table 2: ROUGE-1 evaluation scores for our system, top 7 DUC02 systems, MEAD, TextRank, and the baseline.

Однако, по показателю Recall разработанный метод уступает в Baseline, TextRank и другим, которые показывают более лучшие результаты DUC-02.

5. Заключение

Основываясь на результатах эксперимента разработанный метод генерирует достаточно успешно краткое изложение текста с приемлемой оценкой ROUGE. Более того, он может успешно конкурировать с уже существующими системами. Как продемонстрировано в статье сравнение суммаризаций двух экспертов может достигать в среднем не более 50 процентов по ROUGE-1 F-1. Из этого наблюдения следует, что оценки ROUGE, превышающие этот порог, фактически “переобучаются” на одного конкретного эксперта. Предложенный подход экстрактивной суммаризации успешно справляется с охватом тем документа, но не гарантирует разнообразие предложений, потому что наш подход не учитывает попарные отношения между предложениями. В дальнейшем предлагается изучить возможность модернизации предлагаемого подхода путем применения дополнительных методов, которые обеспечат свойство “разнообразия” генерируемых резюме.

Работа выполнена в рамках проекта ПЦФ №BR05236839 «Разработка информационных технологий и систем для стимулирования устойчивого развития личности как одна из основ развития цифрового Казахстана»

Литература

1. Gaetano Rossiello, Pierpaolo Basile, and Giovanni Semeraro. Centroid-based text summarization through compositionality of word embeddings. - In Proceedings of the MultiLing 2017 Workshop on Summarization and Summary Evaluation Across Source Types and Genres. - 2017. – p. 12–21.
2. Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. - arXiv preprint arXiv:1301.3781. - 2013.
3. Yoshua Bengio, Rjean Ducharme, Pascal Vincent, and Christian Jauvin. A neural probabilistic language model. - In Journal of Machine Learning Research. - 2003. – p. 1137-1155.
4. Matt Kusner, Yu Sun, Nicholas Kolkin, and Kilian Weinberger. From word embeddings to document distances. - In International Conference on Machine Learning. - 2015. – p. 957–966.
5. D. Parveen, M. Mesgar, and M. Strube. Generating coherent summaries of scientific articles using coherence patterns. – In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. - 2016. – p. 772-783.
6. R. Mihalcea and P. Tarau. TextRank: Bringing order into text. - In Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing. - 2004. – p. 404-411.
7. D. R. Radev, T. Allison, S. Blair-Goldensohn, J. Blitzer, A. Celebi, S. Dimitrov, E. Drabek, A. Hakim, W. Lam, D. Liu, and et al. Mead a platform for multi document multilingual text summarization. - In LREC. - 2004a.
8. H. Takamura and M Okumura. Text summarization model based on maximum coverage problem and its variant. In Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics. – 2009. – p. 781-789.
9. Veera Venkata Murali Krishna Ravinuthala and Satyananda Reddy Chinnam. A keyword extraction approach for single document extractive summarization based on topic centrality. - In International Journal of Intelligent Engineering Systems. - 2017.

10. Liqun Shao, Hao Zhang, Ming Jia, Jie Wang. Efficient and effective single-document summarizations and a word-embedding measurement of quality. - In arXiv preprint arXiv:1710.00284. - 2017.
11. Nikola I. Nikolov, Michael Pfeiffer, and Richard H.R. Hahnloser. Data driven summarization of scientific articles. - In arXiv preprint arXiv:1804.08875. - 2018.
12. Dragomir R. Radev, Hongyan Jing, Malgorzata Stys, and Daniel Tam. Centroid-based summarization of multiple documents. In Inf. Process. Manage. - 2004b. – p. 919-938.
13. Rakesh Verma and Daniel Lee. Extractive summarization: Limits, compression, generalized model and heuristics. - Computaci' on y Sistemas 21(4): - 2017. – p. 787– 798.
14. Shih-Hung Liu, Kuan-Yu Chen, Yu-Lun Hsieh, Berlin Chen, Hsin-Min Wang, Hsu-Chun Yen, and WenLian Hsu. Exploring word mover's distance and semantic-aware embedding techniques for extractive broadcast news summarization. - In INTERSPEECH. - 2016. – p. 670–674.

ОПРЕДЕЛЕНИЕ КЛАССА ПРОЧНОСТИ ЗАКЛАДКИ, ОСНОВАННОЕ НА МЕТОДАХ РАСПОЗНАВАНИЯ ОБРАЗОВ

¹Мустафин С.А., ¹Арсланов М.З., ²Шангитбаев Ж.К.
sam@ipic.kz, mars@ipic.kz, zh.shagibayev@almu.edu.kz

¹*Институт информационных и вычислительных технологий МОН РК*
²*Алматинский университет менеджмента*

***Аннотация.** Предложена методика определения класса эксплуатационной готовности складочного массива по данным наблюдений за показателями его состояний, выполненных в заданные моменты времени.*

Введение

Применение математического моделирования и вытекающих из них методов интеллектуальной обработки информации, представляет собой новый шаг в повышении эффективности производства на основе внедрения современных технологий, в том числе инструментов цифровой трансформации. Как правило, построение операций по решению конкретных задач без полной или с частичной формализацией осуществляется опытными специалистами в значительной степени интуитивно, и решение каждой новой задачи требует всей работы заново. В результате часто довольно большая часть данных остается необработанной и невостребованной. Это является существенным недостатком текущей ситуации, которая определяет основное требование для перспективных подходов к решению прикладных задач обработки данных: должна быть обеспечена простота и надежность сложных процедур, не требующих интуитивной работы специалистов [1].

В статье на примере развития прочности складочного материала, применяемого при ряде разработок месторождений полезных ископаемых, рассматриваются методические проблемы прогнозных рисков операций в горном деле, связанных с безопасностью людей и экологией. Показаны преимущества

использования алгоритмов распознавания. Это помогает снизить потери, разубоживание добываемой руды при безопасной добыче, а также обосновать решения при добыче полезных ископаемых. Подчеркивая важность традиционных подходов, мы отмечаем, что использование предлагаемого подхода позволяет получить более точный и объективный прогноз, работать с большими наборами данных, оптимизировать затраты на добычу, рассчитать эффективность добычных работ с учетом минимального воздействия на окружающую среду и повышения безопасности людей.

В горном деле закладку определяют, как заполнение закладочным материалом выработанного пространства, которое образуется в недрах земли в результате выемки полезного ископаемого. Закладочные материалами могут быть измельченные горные породы, и отходы производства. Закладка бывает полной, если заполняется всё выработанное пространство, и частичной при заполнении определённой его части (в виде полос или слоев). В зависимости от способа транспортирования и укладки различают гидравлическую, пневматическую, гидропневматическую, механическую, самотёчную и ручную.

Применение закладки на горнодобывающих предприятиях вызвано обеспечением безопасного процесса ведения добычных работ, сохранением строений на поверхности земли, обеспечением безопасности и сохранения окружающей среды и т.д. С этой целью выработанное пространство заполняют закладочным материалом, который после достижения определенного состояния материала, должен выполнять функцию поддерживающих целиков.

Цели, преследуемые при использовании закладочного материала, зависят от предназначения. Закладка применяется для управления горным давлением, для снижения потерь и разубоживания добываемого полезного ископаемого при добыче, для предотвращения подземных пожаров, для уменьшения деформаций поверхности земли и для охраны объектов на земной поверхности от разрушения, повышения безопасности горных работ, улучшения проветривания подземных выработок, для уменьшения транспортных затрат.

Требования к свойствам закладочного массива могут быть разными и зависят от его назначения. Так, требования к закладочному массиву, предназначенному для предотвращения просадки земной поверхности и охраны тем самым зданий и сооружений намного выше и здесь особенно важно проводить прогноз его состояний закладки, чем в случаях, когда например, закладка выполняет функции заполнителя пустот и предотвращения разубоживания и потерь руды.

В зависимости от назначения и систем разработки месторождений применяются сухая, гидравлическая, твердеющая и др. закладки. Естественно, отличаются свойства и способы их создания.

При твердеющей закладке добавляют вяжущий компонент, что существенно повышает стоимость закладки, вследствие дороговизны вяжущего материала. Данный вид закладки намного превышает стоимость остальных и применяется в строго определенных случаях и только при условии обеспечения полной окупаемости материалов и работ по закладке.

При использовании закладки встает проблема определения момента эксплуатационной готовности состояния твердеющего закладочного массива к выполнению предназначенных ему функций [2].

С конца 50-х годов для решения проблем прогноза состояний технологических процессов стали активно применять различные методы моделирования, основанные на различных показателях процесса. Для этого данные

сложного технологического процесса расчлениют на разные по физической природе составляющие, проводится раздельное их изучение, после чего их взаимное влияние определяется математическими методами с использованием ЭВМ.

Это вызвано тем, что прямой натурный эксперимент часто слишком дорог, опасен, и невозможен, в большинстве случаев, для воспроизведения в лабораториях во всех особенностях реальный процесс, сопровождающийся процессами переносом тепла и превращениями вещества из одного состояния в другое.

Как известно, исследование любого технологического процесса состоит из последовательных этапов: формулировка цели и постановка задачи; изучение всей информации о процессе; построение физико-химической структуры модели процесса; предварительное построение математической модели; решение задач по оптимальным режимным параметрам процесса; проведение вычислительных экспериментов в различных условиях; планирование и осуществление натурального лабораторного эксперимента; сопоставление результатов натурального и вычислительного экспериментов.

Естественно, ряд этапов может отсутствовать или объединен в один. В результате проведения перечисленных этапов возникает модель технологического процесса, готовая к промышленной реализации. Поэтому большую актуальность приобретает разработка моделей поведения процессов. Особенно большое значение модели имеют при оптимальном проектировании, когда требуется оценить оптимальные параметры поведения процессов в различных условиях.

При этом к математической модели предъявляется ряд важных требований адекватности, достаточной простоты, устойчивости относительно погрешностей в исходных данных, простоты, продуктивности, наглядности, стоимости получения исходных данных и других условий [3].

В процессе создания твердеющего закладочного массива, состоящего из заполнения и формирования при этом искусственного массива вследствие усадки материалов закладки, затруднительно проведение натуральных исследований в производственных условиях по широкому спектру показателей, характеризующих его состояние.

Для оценки состояния закладочного массива, возможно исследование его отдельных элементов, выбуривая, например, из закладки керны в определенные моменты времени (неделя, месяц и т.д.) и исследуя их характеристики проводить прогнозирование состояния всего закладочного массива.

Другим способом прогноза состояния закладочного массива может быть физическое моделирование эквивалентными материалами. При этом сам объект исследования может быть исследован механическими, ультразвуковыми, электрическими и другими методами. Оценка изменения в течение времени прочностных, ультразвуковых, электрических и тепловых параметров закладки в целом составляют прогнозную картину состояния закладочного массива.

Попытки решать задачи макрокинетики на основе теории подобия и физического моделирования оказались неполными из-за несовместимости условий подобия химических и физических составляющих процесса. Для решения проблем макрокинетики должны быть известны закономерности собственно химического превращения, не искажённые влиянием процессов переноса, и законы массо- и теплопередачи. Закономерности химического превращения выражаются в виде кинетических уравнений, отражающих зависимость скорости химической реакции от состава реакционной смеси, температуры, давления, свойств катализатора и др.

Таким образом, к основным методам, которые позволяют контролировать состояние закладки, относятся: радиометрический контроль, методы местных разрушений, методы ударных воздействий, акустический метод, метод измерения температуры материала и т.д.

Основным элементом перечисленных методов является то, что они требуют построения индивидуальных градировочных зависимостей по результатам исследований стандартных образцов-кубов, изготовленных из бетона такого же состава и возраста, что и исследуемый объект, конструкция, в нашем случае закладка. Непосредственно при этом измеряется некоторая косвенная физическая характеристика, связанная с прочностью корреляционной зависимостью. Для установления этой зависимости, а соответственно, и для установления прочности конструкции предварительно необходимо установить градировочную характеристику между прочностью и некоторой косвенной характеристикой – температурой, влажностью, проводимостью и т.д.

Для этого выводят зависимость прочности R массива от его косвенной характеристики, изменяющейся по времени, - на основе метода наименьших квадратов строится зависимость выбранной характеристики (температуры, влажности, акустического параметра, электрической проводимости и т.д.) от времени t с соответствующим вычислением остальных показателей.

Значения (характеристики) в момент времени t , принимают за параметры динамического состояния закладки.

Следует отметить, что каждая характеристика в отдельности определяет состояние искусственного массива, но не полностью, а только односторонне, и не может быть глобальной оценкой динамического состояния процесса.

Таким образом, объект исследования имеет набор характеристик состояния объекта, изменяющихся во времени.

Исходя из приведенных выше утверждений, можно констатировать, что определение состояния закладки является важной и актуальной проблемой, поэтому существует потребность в разработке новых способов и методов определения класса состояний технологического процесса готовности.

Во всех случаях для прогнозной оценки состояния закладочного массива представляется необходимым измерять в различные моменты времени $T = \{t_1, \dots, t_T\}$ параметры закладки, характеризующие статическое состояние и их динамику [1].

Под состоянием закладочного материала будем понимать набор значений физических свойств материала в некоторый момент времени t . Такими признаками являются механические, тепловые, ультразвуковые, электрические и другие параметры закладочного материала, измеряемые в момент времени $t - X$.

Пусть на входе предполагаемой системы оценки состояния закладочного массива зафиксированы значения входных параметров X , а на выходе принимаются значения выходных параметров Y .

Под моделью объекта будем понимать некоторый оператор $F(x) (\hat{y} = F(x))$, устанавливающий соответствие между входными и выходными переменными, позволяющий с определенной точностью восстанавливать выходные переменные объекта по ее входным переменным.

Необходимо построить такое отображение $X \rightarrow Y$, чтобы на каждый возможный входной сигнал X формировался правильный выходной сигнал Y .

Практика статистического моделирования показала, что наиболее употребительны непрерывные регрессионные модели (линейные и нелинейные),

построенные для всей выборочной совокупности объектов. Их использование равносильно предположению о качественной однородности рассматриваемой совокупности объектов, об отсутствии разрывов и скачков в моделируемой зависимости. Работа с линейными моделями накладывает еще более жесткие требования к исходной совокупности - многомерное распределение признаков, описывающих объект, должно быть близким к нормальному. При этом предполагается стабильность характера зависимости и параметров модели во всех подобластях признакового пространства.

Между тем на практике эти условия, как правило, не соблюдаются, что делает неправомочным использование в качестве оператора $F(x)$ непрерывных регрессионных моделей. Например, пространственные совокупности траекторий развития динамических объектов одной задачи, как правило, неоднородны и обладают определенной структурой. Последнее означает, что общая совокупность траекторий развития объектов распадается на ряд подсовокупностей со специфическим характером зависимости показателей от входных переменных.

Довольно эффективны для аппроксимации функции методы кусочной аппроксимации. Основная идея этих методов состоит в разбиении выборочного пространства на области, в каждой из которых строится своя функция. Однако можно отметить ряд малоисследованных вопросов, связанных с разделением выборочного пространства на области, в которых строятся отдельные функции. Построение частных моделей выделенных областях также, на наш взгляд, не полностью решает задачу идентификации.

Таким образом, применение традиционных методов прогноза готовности показывает низкую достоверность из-за игнорирования сложности объекта исследования, что значительно сужает область их использования на практике.

Нами предлагается процедура, в результате которой происходит объединение полученных частных моделей в единую статистическую модель объекта (процесса).

Идея предлагаемой методики прогнозирования состояния закладки состоит в учете структуры данных, а именно, в разбиение пространства признаков, характеризующих состояние закладочного материала в определенные моменты времени, на классы близких объектов в пространстве признаков, в каждом классе в эти моменты времени строится своя функция прогноза. Построение частных моделей на выделенных областях позволяет объединить частные модели прогноза в единую модель прогноза состояния закладки.

Другими словами, предпринята попытка учесть структурную неоднородность исходных данных по состояниям объекта исследования во времени, т.е. аппроксимировать реальную зависимость состояния закладки от физических параметров массива функцией с учетом времени t и принадлежности классу [4].

Приведем формальную постановку задачи.

Рассматривается множество траекторий технологического процесса твердения $\tilde{S} = \{S_1, \dots, S_m\}$, представляющих собой сложные развивающиеся во времени однотипные системы, описание состояний которых дано в признаковом пространстве.

Наблюдается процесс развития каждого объекта S_v , $v=1,2,\dots,M$, на отрезке $[t_0, t_T]$. Наблюдение ведется в дискретные моменты времени.

В каждый момент времени на отрезке $[t_0, t_T]$, любой объект из \tilde{S} может находиться в одном из своих состояний, описываемых с помощью набора из $h(t_k)$ признаков x_1, x_2, \dots, x_h .

Состояние объекта S_v в момент времени t_k описывается следующим вектором

$$\bar{a}_v(t_k) = (\bar{a}_{v1}(t_k), \dots, \bar{a}_{vn}(t_k)) \quad (1) \\ v = 1, \dots, M; \quad k = 0, 1, \dots, T$$

Каждый из признаков x_i принимает значения из множества $M_i, i = 1, \dots, h$.

Считаем $M_i, i = 1, \dots, n$ метрическим или полуметрическим пространством, в котором не выполняется аксиома треугольника.

Состояния динамических объектов \tilde{S} в произвольный объект момент времени t_k являются представителями некоторых зависящих от времени множеств K_1, \dots, K_l , называемых в дальнейшем классами. Предполагаем, что $K_i \cap K_j = \emptyset, i \neq j, i, j = 1, \dots, l$.

Подмножество $\tilde{S}_0 = \{S_1, \dots, S_m\}$ объектов \tilde{S} выбираем в качестве обучающей выборки.

Основная задача. Пусть задана траектория S , которая принадлежит одному из заданных классов траекторий $K_j, j = 1, 2, \dots, l$ и наблюдаемая на интервале $[t_0, t_T]$. Требуется определить процедурами распознавания принадлежность траектории S к одному из заданных классов траекторий.

Заключение. Методика, использующая методы распознавания образов, позволяет повысить достоверность определения и получить модель, которая более точно выражает определяемую траекторию изменения закладочного материала по наборам признаков его состояния.

Рассматриваемый подход позволяет придать статистическим моделям свойства оптимизируемости, управляемости, а также повысить их адекватность изучаемым химико-технологическим процессам.

Предложена формальная постановка задач прогноза состояния закладочного массива, используемый при разработке месторождений.

Разработана методика определения состояния закладочного массива, использующая методы распознавания на основе вычисления оценок, позволяют выявлять эксплуатационную готовность закладки.

Применение разработанной методики к решению конкретных задач позволяет детально исследовать состояние закладочного массива по мере развития его твердения, устанавливать класс эксплуатационной готовности на основе использования алгоритмов распознавания.

Другими словами, методика, использующая методы распознавания образов, позволяет повысить достоверность определения класса траектории и получить модель, которая более точно выражает прогнозируемую траекторию изменения закладочного материала по наборам признаков его состояния.

Данная работа выполнена при финансовой поддержке МОН РК в рамках научных проектов лаборатории "Распознавание образов и принятия решений" Института информационных и вычислительных технологий.

Литература

1. Калимолдаев М.Н. и др. (2018) Методологическая основа стратегии развития систем искусственного интеллекта в свете послания президента РК от

05.10.2018 г. // Известия НАН РК. Серия Геологические и Технические науки. 2018. Том. 6, С.47-54

2. Байконуров, О.А. Подземная разработка месторождений с закладкой / О.А. Байконуров, Л. Крупник, В. Мельников. - Алма-Ата: Наука, 1972. - 384 стр.

3. Блехман И.И. Механика и прикладная математика: Логика и особенности приложений математики / И.И. Блехман, А.Д. Мышкис, Я.Г. Пановко - Москва: Наука, 1990. - 356 стр.

4. Журавлев Ю.И.: Избранные труды. - Москва: Магистр, 1998. – 568 стр.

РЕАЛИЗАЦИЯ И ОПУСКАНИЕ ТИПОВ В ФОРСИНГ-СТРУКТУРАХ

Нуртазин А.Т.

e-mail: abyznurtazin@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** Понятие «форсинг-структуры» было введено Абрахамом Робинсоном для построения экзистенциально замкнутых структур, но в результате изучения форсинг-структур оказалось намного сложнее. Для продвижения на этом пути ранее автором было введено понятие «форсинг-типа». В предлагаемой работе изучены некоторые возможности реализации и опускания этих типов.*

1. Структуры общего вида и переход к реляционной сигнатуре. Как известно, объектами исследования в математике являются математические структуры. Любая такая структура состоит из некоторого фиксированного множества, называемого базисным, и двух фиксированных последовательностей функций и отношений на нём. При этом в математической логике обычно рассматриваются структуры, в которых такие последовательности не более чем счётны. Более того, в наиболее интересных случаях они конечны. Для простоты в дальнейшем мы рассматриваем структуры счётной чисто предикатной сигнатуры S . Обоснование этого ограничения основывается на замене в структурах смешанной сигнатуры всех функций на их графики.

Замена функций на их графики. Переход к реляционной сигнатуре. Ограничение чисто предикатными сигнатурами значительно упрощает доказательства многих утверждений. Поэтому важно показать, что такой переход при рассмотрении индуктивных классов и теорий не приводит к принципиальным потерям. В частности, следующие утверждения показывают, что переход к реляционной сигнатуре оказывается вполне адекватным методом при рассмотрении формульных отношений, индуктивных классов и теорий и их экзистенциально замкнутых моделей.

¹⁰. Если в структуре A' общего вида все функции заменить на их графики, а в её сигнатуре S' все соответствующие функциональные символы -- на предикатные символы на единицу больших местностей, то семейства формульных

отношений в структуре A' и полученной в результате такой замены структуре A совпадают.

2⁰. Если класс C' моделей смешанной сигнатуры S' аксиоматизируем, то класс C моделей чисто предикатной сигнатуры S , полученных из моделей класса C' заменой функций на их графики, также аксиоматизируем.

3⁰. Если элементарная теория аксиоматизируемого класса моделей C' аксиоматизируется универсальными или $\forall\exists$ -предложениями, то теория T класса C моделей, полученных из моделей класса C' заменой всех функций на их графики, также является $\forall\exists$ -теорией.

4⁰. Пусть C' некоторый аксиоматизируемый класс моделей общего вида и класс реляционных моделей C получен из него заменой всех функций на их графики. Тогда в результате прямого перехода из любой модели, экзистенциально замкнутой в первом классе, получается модель, экзистенциально замкнутая во втором, и наоборот.

Сделанные замечания позволяют при изучении индуктивных аксиоматизируемых классов и теорий ограничиться рассмотрением реляционных структур и индуктивных теорий некоторой счётной чисто предикатной сигнатуры S .

2. Классы Фрэсе. В середине прошлого столетия выдающийся французский логик и один из классиков теории моделей Роланд Фрэсе показал, что при определении значений формул в данной реляционной структуре M конечной сигнатуры S важную роль играет семейство её конечных подструктур F . Естественно предположить, что для изучения структур произвольной реляционной сигнатуры S аналогичную роль должны играть классы Фрэсе.

Определение 1.

Класс Фрэсе $F(M) = \text{Age}(M)$ произвольной S -структуры M это класс всех конечных структур, изоморфных всевозможным ограничениям на конечные подсигнатуры сигнатуры S конечных подструктур структуры M .

Можно ожидать, что такой переход от конечной реляционной сигнатуры к бесконечной неизбежно приведёт к усложнению ряда определений и обозначений. В частности, для простоты дальнейшего изложения в отличие от общепринятого, конечную структуру $A_1 = \langle A_1; s_1 \rangle$ мы будем считать подструктурой конечной структуры $A_2 = \langle A_2; s_2 \rangle$ ($A_1 \leq A_2$), если выполняются включения $A_1 \subseteq A_2$ и $s_1 \subseteq s_2$ и равенство $A_1 = A_2 \upharpoonright A_1 \upharpoonright s_1$. Аналогично, вложение f основного множества структуры Фрэсе A_1 в основное множество структуры Фрэсе A_2 мы называем изоморфным вложением, если выполняется включение $s_1 \subseteq s_2$ и отображение f является изоморфным вложением A_1 в $A_2 \upharpoonright A_1 \upharpoonright s_1$.

Легко видеть, что для произвольной структуры M класс $\text{Age}(M)$ является абстрактным и замкнут относительно взятия подструктур и обединений. Кроме того, в $\text{Age}(M)$ любые две конечные структуры A_1 и A_2 могут быть одновременно вложены в некоторую третью. Это свойство класса $\text{Age}(M)$ обычно называется свойством совместного вложения. Эти свойства положены в основу другого определения "класса Фрэсе", которое в семидесятых и восьмидесятых годах двадцатого столетия было использовано для построения наиболее тонких примеров, опровергающих целый ряд трудных гипотез теории моделей того времени.

Определение 2.

Просто классом Фресе F счётной предикатной сигнатуры S называется произвольный наследственно замкнутый, обладающий свойством совместного вложения абстрактный класс конечных структур конечных подсигнатур сигнатуры S .

Компаньоны. Следующее определение уточняет сразу несколько имевшихся до него в теории экзистенциальной замкнутости понятий "модельного, индуктивного, экзистенциально замкнутого и форсинг-компаньонов".

Определение 3.

Структуры M и N счётной чисто предикатной сигнатуры называются компаньонами, если совпадают их классы Фресе: $F(M) = F(N)$.

Оказалось, что класс всех компаньонов данной структуры (или, просто, компаньон-класс) аксиоматизируется двумя множествами экзистенциальных и универсальных предложений. Внутри данного компаньон-класса C в порядке убывания естественно определяются индуктивный I , экзистенциально замкнутый E и форсинг F компаньон-подклассы. При этом первый из этих классов аксиоматизируем, а теория третьего - полна.

Диаграммы и бескванторные формулы. Пусть M некоторая структура счётной реляционной сигнатуры S . В отличие от общепринятого, конечной подструктурой A структуры M (конечной s -подструктурой) мы называем любое ограничение M её обычной конечной подструктуры $A = \langle A; S \rangle$ на произвольную конечную подсигнатуру s сигнатуры S . Простым упражнением в написании формул является поиск предложения, выделяющего все структуры, изоморфные структуре M . Более точно, если $a = \langle a_0, \dots, a_{n-1} \rangle$ перечисление всех элементов A , то формула $D_a^s(x) \equiv \{R^e(x): R \in s, e \in \{0,1\}, M \models R^e(a)\}$ называется s -диаграммой кортежа a или, просто, s - n -диаграммой. Легко видеть, что в этом случае, если в любой другой S -структуре M на некотором n -элементном кортеже b выполняется $M \models D_a^s(b)$, то покоординатное соответствие из a на b является изоморфизмом структуры A на ограничение $N \parallel b \parallel s$. Таким образом, в качестве искомого предложения, выделяющего класс структур, изоморфных структуре A , можно взять $\exists x D_a^s(x)$.

Так как для любой конечной сигнатуры s и натурального числа n число попарно неизоморфных n -элементных s -структур конечно, то также конечным оказывается и число различных s - n -диаграмм. Если в записи некоторой формулы $\varphi(x)$ участвуют лишь символы из некоторой сигнатуры s , то эту формулу мы называем s -формулой, а в случае $x = \langle x_0, \dots, x_{n-1} \rangle$ -- s - n -формулой. Все s -диаграммы являются бескванторными s -формулами. Из определения следует, что любая s - n -диаграмма выполнима. Также достаточно просто видеть, что любая s - n -диаграмма является минимальной в классе выполнимых бескванторных s -формул от переменных x относительно логического следования. Действительно, если произвольная бескванторная s -формула $\varphi(x)$ совместна с s -диаграммой $D_a^s(x)$, то можно считать, что она выполняется на самом кортеже a . Но тогда для любого другого кортежа b из $M \models D_a^s(b)$ следует, что на нём выполняются все атомные и отрицания атомных s -формул, истинных на a . Индукцией по сложности легко проверяется, что на этом кортеже b выполняется и формула $\varphi(x)$. Таким

образом, нами в точности доказано, что произвольная бескванторная s -формула $\varphi(x)$, выполняющаяся на некотором кортеже a , имеющем s -диаграмму $D_a^s(x)$, также выполняется и на любом другом кортеже b , с той же самой s -диаграммой. Теперь непосредственно из сделанного замечания также вытекает, что любая бескванторная s - n -формула логически эквивалентна дизъюнкции совместных с нею s - n -диаграмм.

Кванторная сложность формул. В математической логике хорошо известна естественная сложностная классификация формул, заданных в пренексной форме, в которой все кванторы расположены в начале (см., например, [Simmons]).

А именно, *кванторной сложностью* произвольной формулы в пренексной форме называется число блоков расположенных рядом переменных, связанных в этой форме одноимёнными кванторами. При этом наиболее простыми являются экзистенциальные формулы, имеющие пренексную форму, в которой на все переменные навешивается квантор существования, и универсальные формулы, в получении пренексной формы которых используются только кванторы всеобщности. Любая универсальная формула, не имеющая свободных переменных, называется универсальным предложением.

Как уже отмечалось, произвольная бескванторная s - n -формула $\varphi(x)$ логически эквивалентна дизъюнкции совместных с нею s - n -диаграмм: $\varphi(x) \sim D_1^s(x) \vee \dots \vee D_r^s(x)$. Непосредственно из определения истинности следует, что моделями универсального предложения $\forall x\varphi(x)$ являются все s -структуры, в которых любые n -элементные кортежи имеют диаграммы из множества $\{D_1^s(x), \dots, D_r^s(x)\}$. С другой стороны, если при этом $D_{r+1}^s(x), \dots, D_s^s(x)$ --- все s - n -диаграммы, несовместные с формулой $\varphi(x)$, то любая модель предложения $\forall x\varphi(x)$ не содержит n -элементных кортежей, имеющих одну из диаграмм $D_{r+1}^s(x), \dots, D_s^s(x)$ ([Мальцев]).

Индуктивные классы и теории. Класс формул кванторной сложности два, в пренексной форме которых вначале идут переменные, связанные кванторами всеобщности, а затем -- переменные, связанные кванторами существования, называются $\forall\exists$ -формулами. Этот класс является первым нетривиальным классом в упомянутой кванторной сложностной классификации и после подходящего обогащения сигнатуры без особых потерь с помощью этих формул можно формализовать большинство свойств и отношений, выразимых в логике первого порядка. В то же время класс структур, замкнутый относительно объединений возрастающих цепей, называется индуктивным классом. Одним из важных достижений периода бурного развития теории моделей, относящегося к шестидесятым годам прошлого столетия, явилось описание индуктивных аксиоматизируемых классов. Элементарная теория, заданная $\forall\exists$ -предложениями, называется $\forall\exists$ -теорией. Отметим, что стартовой и основополагающей здесь является теорема Лося -- Сушко, которая говорит, что понятия *индуктивный аксиоматизируемый класс* и *$\forall\exists$ -аксиоматизируемый класс* совпадают.

Форсинг. Несмотря на то, что исторически понятие форсинг-структуры явилось вторичным относительно понятия экзистенциальной замкнутости и даже, более точно, было заимствовано Абрахамом Робинсоном из теории множеств именно для построения экзистенциально замкнутых структур, как показал Роланд Фрэсе, его можно напрямую использовать для определения вынуждения формул в

классах Фрэсе, а затем получить все форсинг-структуры, как структуры, в которых на произвольном кортеже вынуждается любая формула либо её отрицание. Дадим принадлежащее Роланду Фрэсе и используемое здесь семантическое определение *форсируемости*. При этом мы считаем, что все рассматриваемые здесь логические формулы имеют пренексную форму, в которой все кванторы расположены в начале формулы. Известно, что тогда число следующих один за другим блоков одноимённых кванторов в начале такой формулы $\varphi(x)$ называется её кванторной сложностью $q(\varphi)$. Теперь *форсируемость* формул естественно определяется индукцией по их кванторной сложности.

Определение.

По Фрэсе для структуры $A = \langle A; s \rangle$ из класса F , лежащего в ней кортежа a и s -формулы $\varphi(x)$ отношение " A в F вынуждает $\varphi(a)$ " определяется индукцией по кванторной сложности формулы φ :

1. $q(\varphi) = 0$.
 A форсирует $\varphi(a)$, если и только если $A \models \varphi(a)$.
2. $q(\varphi) = k+1$ и $\varphi(x) \equiv \exists x^1 \varphi^1(x, x^1)$. A форсирует $\varphi(a)$, если и только если в A найдётся кортеж a^1 такой, что A форсирует $\varphi^1(a, a^1)$.
3. $q(\varphi) = k+1$ и $\varphi(x) \equiv \forall x^1 \varphi^1(x, x^1)$. A форсирует $\varphi(a)$, если и только если не существует расширения B структуры A , форсирующего формулу $\varphi^1(a, a^1)$.

Несложное рассуждение показывает, что правильным является следующее внутреннее определение форсинг-структуры, которое позволяет задать класс F всех форсинг-структур безотносительно какой-либо теории.

Определение.

Структура M является форсинг-структурой относительно вынуждения в классе $F(M)$, если для любых кортежа a и формулы $\varphi(x)$ в M найдётся конечная подструктура A , форсирующая в классе Фрэсе $F(M)$ либо формулу $\varphi(a)$, либо её отрицание $\neg \varphi(a)$.

Из этого определения легко получаются следующие известные свойства форсинг структур из произвольного форсинг-компаньона F .

1⁰. Данная M является форсинг-структурой, если и только если в ней для каждого кортежа a вынуждается любая истинная на нём формула.

2⁰. Данная M является форсинг-структурой, если и только если в ней для любой формулы $\varphi(x)$, истинной на произвольном кортеже a , нашлась бы также истинная на нём экзистенциальная формула $\varphi^1(x)$, такая, что $M \models \forall x [\varphi^1(x) \rightarrow \varphi(x)]$

3⁰. Объединение возрастающей цепочки форсинг-структур само является форсинг-структурой.

4⁰. Структура M , экзистенциально замкнутая в форсинг-структуре N , сама является форсинг-структурой.

Форсинг-типы. Реализация и опускание. В [5] автор показал, что экзистенциальные типы, реализующиеся в экзистенциально замкнутых структурах максимальны. Этот факт обусловил их важность при изучении этого класса структур. Из [1] следует, что элементарная теория любого форсинг-компаньона полна, а экзистенциальные типы, реализуемые в форсинг-структурах не только

максимальны, но и имеют единственные расширения до полных. В [7] автор назвал такие экзистенциальные типы форсинг-типами. Там же было доказано, что в случае существования простая модель над кортежом, реализующим форсинг-тип, является форсинг-моделью, и любой форсинг-тип можно реализовать в некоторой структуре, являющейся моделью форсинг-теории. В 2018 году автором доказано, что любой форсинг-тип может быть реализован в некоторой форсинг-модели. Приведённый результат немедленно делает актуальными следующие вопросы:

1. *Можно ли два форсинг-типа $p(x)$ и $q(y)$ реализовать в одной и той же форсинг-структуре?*

2. *При каких условиях данный форсинг-тип $p(x)$ можно опустить в некоторой форсинг структуре?*

3. *При каких условиях существует форсинг-структура, в которой реализуется форсинг-тип $p(x)$ и опускается форсинг-тип $q(y)$?*

На все три сформулированных вопроса были найдены приемлемые ответы, которые предполагается использовать для определения возможного числа счётных форсинг-структур.

Библиография

1. Barwise J., Robinson A., Completing theories by forcing, Ann. Math. Logic, 1970, 2, 119-142.
2. Cohen P.J., Set Theory and the Continuum Hypothesis, N.-Y., Benjamin, 1966.
3. Ершов Ю.Л., Палютин Е.А., Тайманов А.Д., Теория моделей, Справочная книга по математической логике, Часть 1, главы 2 и 4, М "Наука", 1982.
4. Fraisse R., Sur quelques classifications des systemes de relations. --- Publ. scient. de l'univ. d'Algers, 1955, A1, p. 35-182.
5. Нуртазин А.Т. Счетные экзистенциально замкнутые модели универсально аксиоматизируемых теорий // Математические труды. - Новосибирск, Институт математики СО РАН им. Соболева С.Л., 2015. - т.18, №1. - С.48-97.
6. Нуртазин А.Т., Некоторые свойства экзистенциально замкнутых компаньонов, Новосибирск, Алгебра и логика, 2018, № 2.
7. Нуртазин А.Т., Вынуждение в структурах и классах Фрэсе, Новосибирск, Алгебра и логика, 2018, № 4.
8. Robinson A., On the Metamathematics of Algebra, -- Amsterdam: North -- Holland, 1951.

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА СИММЕТРИЧНОГО БЛОЧНОГО ШИФРОВАНИЯ «QAMAL» ДЛЯ ИССЛЕДОВАНИЯ ЕГО КРИПТОГРАФИЧЕСКИХ СВОЙСТВ

Сақан Қ., Алғазы К.

e-mail: kairat_sks@mail.ru, kunbolat@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Аннотация. В данной работе описана программная реализация разработанной системы симметричного блочного шифрования «Qamal», которая предназначена для проведения расчетов промежуточных параметров, получения статистических данных и дальнейшего исследования полученных данных. На основе полученных результатов будет корректироваться этот алгоритм шифрования.

В настоящее время для Республики Казахстан обеспечение информационной безопасности является одной из основных задач в информационной сфере, которая достигается путем разработки и реализации комплекса мероприятий, направленных на защиту национальных интересов в различных сферах жизни и деятельности общества и государства. Важную роль здесь выполняют организационные меры, способы и методы обеспечения информационной безопасности. В связи с этим, создание отечественных средств защиты информации, в том числе и криптографических, является актуальной и необходимой [1].

В соответствии с календарным планом на 2018 год программы BR053236757 «Разработка программных и программно-аппаратных средств для криптографической защиты информации при ее передаче и хранении в инфокоммуникационных системах и сетях общего назначения» был предложен новый симметричный блочный алгоритм шифрования «Qamal» [2].

В алгоритме шифрования Qamal применены обратимые преобразования открытого текста для трех длин блоков файлов и ключей в 128, 192 и 256 бит. Ключам длиной в 128, 192 и 256 бит соответствуют число раундов шифрования 6, 8, 10. Процесс зашифрования состоит из разработанных процедур наложения ключа с помощью операции побитового сложения, блока замены и процедур перемешивания. В процессе расшифрования те же самые криптографические преобразования инвертируются и применяются в обратном порядке. Согласно плану НИР на 2019 год осуществлена программная реализация этой системы симметричного блочного шифрования Qamal и проведено исследование его надежности с использованием методов криптоанализа. На рисунке 1 приведено основное окно программы «Алгоритм «QAMAL»».

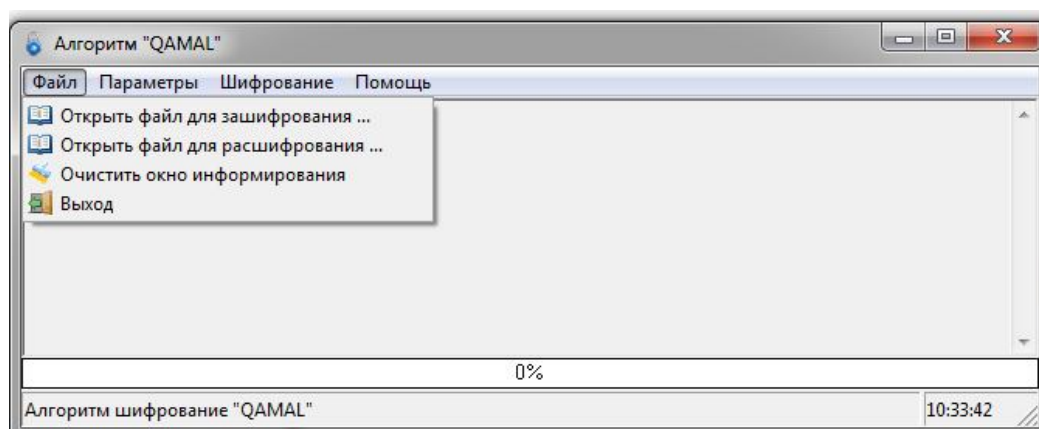


Рисунок 1 - Основное окно программы «Алгоритм «QAMAL»»

Запуск программы производится через «Qamal_RW.exe». Появляется диалоговое окно (рисунок 1). Окно меню состоит из четырех разделов: «Файл», «Параметры», «Шифрование» и «Помощь».

Раздел «Файл» – это диалоговое окно ввода файла для его зашифрования или расшифрования.

Раздел «Параметры» – окно ввода и установки длины блока, количества раундов, длины базового ключа и начального вектора инициализации шифрования. В зависимости от установленной длины блока шифрования программа сама определяет количество раундов зашифрования, длину базового ключа и вектора инициализации (рисунок 2).

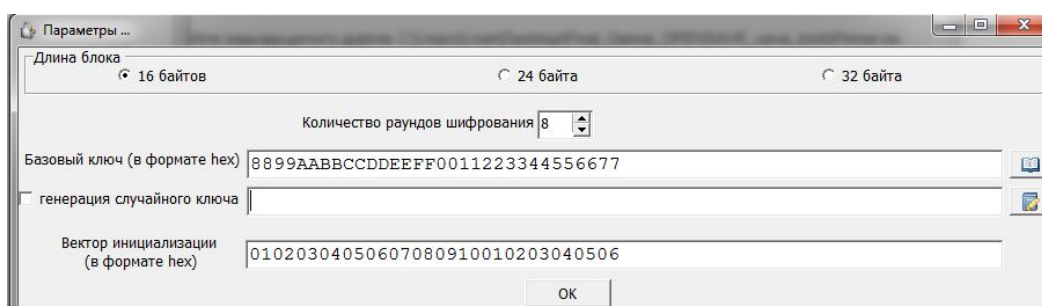


Рисунок 2 - Раздел «Параметры»

Раздел «Шифрование» – окно выбор направления шифрования и его запуска (рисунок 3). Чтение блока из файла, процесс зашифрования и сохранение полученного зашифрованного блока в файле производится поблочно.

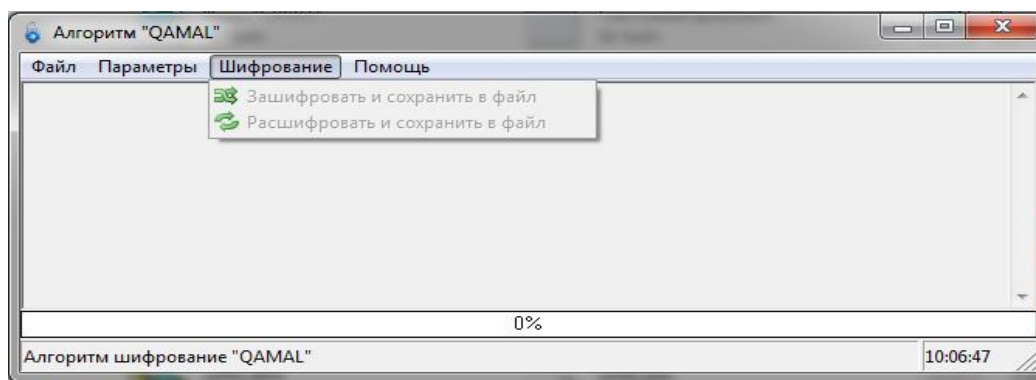


Рисунок 3 - Раздел «Шифрование»

Процесс зашифрования:

1. Ввод констант и начальных данных. Входными данными программного обеспечения являются:

- ввод файла для шифрования;
- ввод начальных параметров: K_0 - базовый ключ длиной 16, 24 или 32 байта, V_0 - начальный вектор инициализации.

2. В зависимости от длины блока шифрования определяется число раундовых ключей. Как определено в описании «Qamal», длина блока может иметь длину 128, 194 и 256 бит, количество раундовых ключей равно $R = 8, 10, 12$ соответственно.

Процесс генерации следующего раундового ключа K_{i+1} выполняется путем выполнения последовательности преобразований: нелинейного преобразования – процедуры $QamS_key$, процедуры перемешивания $Mixer1$ и $Mixer2$ и процедура $Module\ p_i(x)$. Данная последовательность действий повторяется 10 раз и в конце выполняется побитовое сложение (xor) полученного блока с ключом K_i . Блок-схема формирования раундовых ключей (подпрограмма $QamKeyGen$) приведена ниже на рисунке 4 [3].

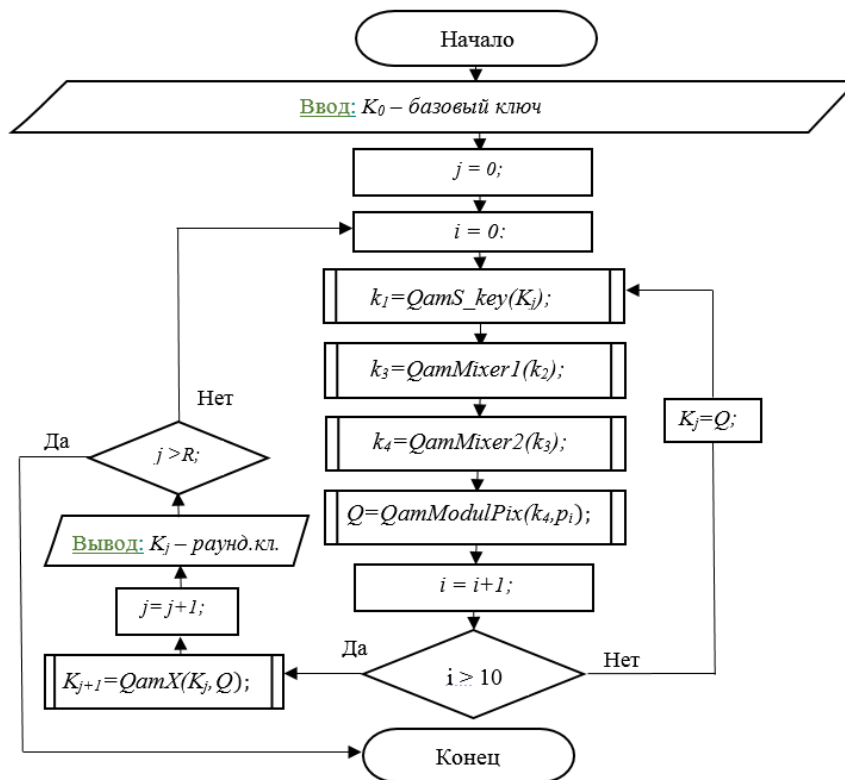


Рисунок 4 - Блок-схема формирования раундовых ключей (подпрограмма $QamKeyGen$)

3. Блок зашифрования. Блок-схема зашифрования данных, показана на рисунке 5. Процесс зашифрования данных состоит из следующих блоков преобразования:

3.1. Режим сцепления шифрования реализуется путем битового сложения по модулю 2 вектора инициализации и блока открытого текста (процедура $QamX$).

3.2. Выполняется операция xor ключа K_0 и полученного блока в подпункте 3.1. (процедура $QamX$).

3.3. Производится нелинейное биективное преобразование байтов (процедура $QamS$).

3.4. Преобразование по столбцам матрицы A (процедура $Mixer1$). При формировании матрицы A байты блока, полученного в подпункте 3.3, представляются в виде двумерного массива размером $m = 4, 6$ и 8 в зависимости от размера блока.

3.5. Преобразование по строкам матрицы B (процедура $Mixer2$). Структура и элементы матрицы B определяются аналогично, как для матрицы A (данные в результате формирования $Mixer1$).

3.6. Выполняется операция xor ключа K_{i+1} с полученным блоком из подпункта 1.5. (процедура $QamX$).

3.7. Подпункты 3.3 – 3.6 повторяются 8, 10 или 12 раз в зависимости длины блока с применением ключа K_{i+1} .

3.8. Подпункты 3.1 - 3.8 повторяются до зашифрования последнего блока файла. В случае неполного последнего блока шифрования недостающие байты дополняются нулевыми байтами до целого, при этом вместо последнего нулевого байта – код количества дополненных байтов.

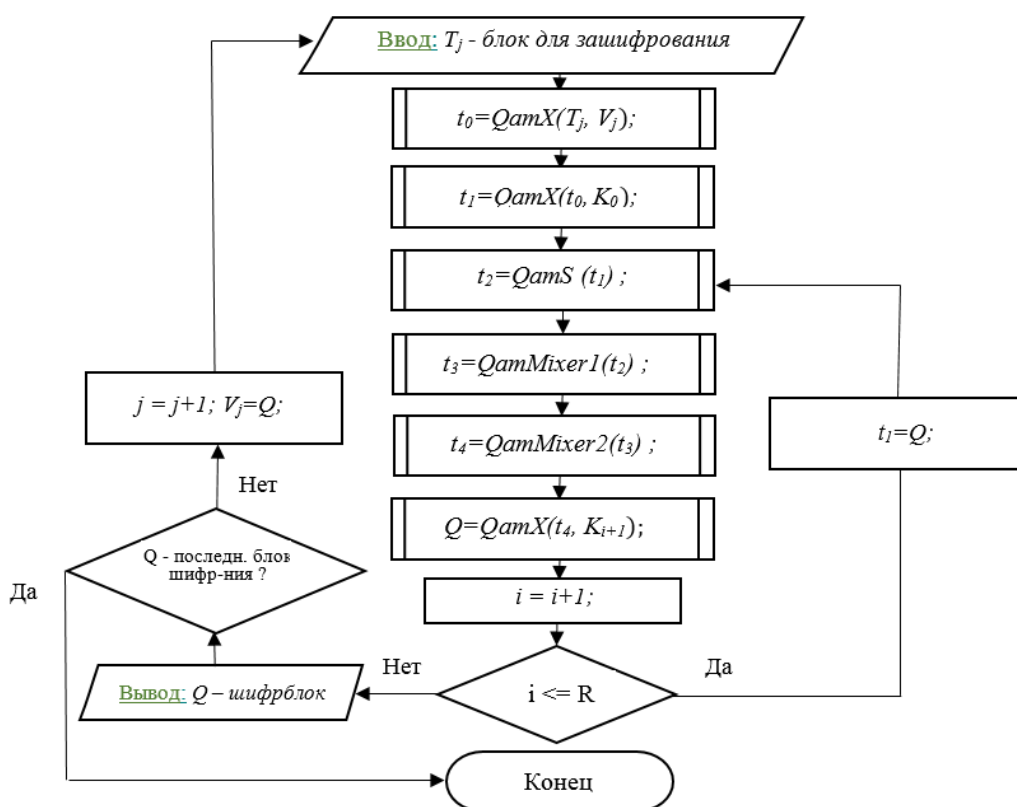


Рисунок 5. Блок-схема блока зашифрования

4. Блок расшифрования

Процесс расшифрования производится следующим образом. Все криптографические преобразования $QamS$, $Mixer1$ и $Mixer2$ инвертируются и $QamSR$, $MixerR1$ и $MixerR2$ и применяются в обратном порядке. Раундовые ключи K_i также используется в обратном порядке. Нелинейному преобразованию $S-box$ соответствует обратное преобразование $SR-box$. Блок-схема блока показана на рисунке 6.

4.1. Выполняется операция xor ключа K_i с блоком для расшифрования или с полученным блоком из подпункта 4.4 (процедура $QamXR$).

4.2. Выполнение преобразования по строкам матрицы C (процедура $MixerR2$).

4.3. Выполнение преобразования по столбцам матрицы D (процедура $MixerR1$).

4.4. После этого выполняется нелинейное биективное преобразование байтов (процедура $QamSR$).

4.5. Подпункты 4.1 – 4.4 повторяются 8, 10 или 12 раз в зависимости длины блока с применением раундовых ключей K_i , где $i = R, \dots, 1$.

4.6. Выполнение операции xor ключа K_0 с результатом пункта 4.4. (процедура $QamX$).

4.7. Режим сцепления шифрования реализуется через процедуры $QamX$.

4.8. Все пункты раздела циклически выполняются до расшифрования последнего блока зашифрованного текста.

Предложенная система шифрования реализована на языке программирования Delphi v.7. Объем рабочего файла - 1,13 МБ. Системное требование: тактовая частота – не ниже 2 ГГц, ОЗУ – не менее 1 МБ, объем свободного места на жестком диске – не меньше 10 Гб, операционные системы - Windows Me/2000/XP/7/8. При разработки программного обеспечения рассмотрены различные варианты вычисления и вывода промежуточных и конечных данных.

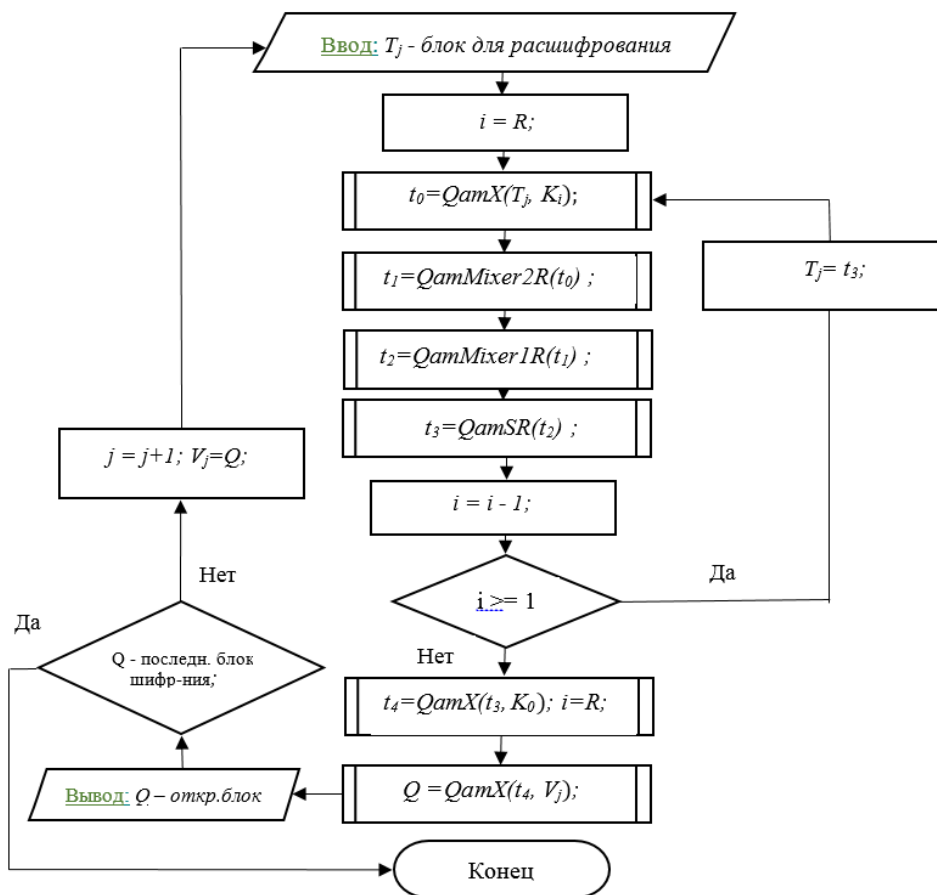


Рисунок 6 - Блок-схема блока расшифрования

Заключение

На разработанных компьютерных программах будут исследованы свойства предложенной системы шифрования. Будут проведены работы по повышению быстродействия программ.

Представленные работы проведены за счет средств финансирования проекта программно-целевого финансирования BR053236757 научных исследований на 2018-2020 годы «Разработка программных и программно-аппаратных средств для криптографической защиты информации при ее передаче и хранении в инфокоммуникационных системах и сетях общего назначения».

Список литературы

1. Концепция кибербезопасности ("Кибершит Казахстана") / Утверждена постановлением Правительства Республики Казахстан от 30 июня 2017 года № 407.
2. Отчет о научно-исследовательской работе «Разработка программных и программно-аппаратных средств для криптографической защиты информации при ее передаче и хранении в инфокоммуникационных системах и сетях общего назначения», 2018 г., № гос. регистрации 0118РК01064.
3. Брюс Шнайер. Прикладная криптография. Протоколы, алгоритмы и исходные тексты на языке С. 2-е издание. 2002 г. «Триумф». Москва.

ТЕХНОЛОГИЯ СОЗДАНИЯ ТЕМАТИЧЕСКИХ ИНТЕЛЛЕКТУАЛЬНЫХ РЕСУРСОВ ДЛЯ ПОДДЕРЖКИ НАУЧНО-ОБРАЗОВАТЕЛЬНОЙ ДЕЯТЕЛЬНОСТИ НА ОСНОВЕ ОНТОЛОГИЧЕСКОГО ПРОЕКТИРОВАНИЯ

Самбетбаева М.А., Батырханов А.Г.

e-mail: Madina_jgtu@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Евразийский национальный университет им. Л. Н. Гумилева, Казахстан

Аннотация. В статье представлена выработка концептуального видения технологии использования системы управления электронными библиотеками, как типовой системы хранения и доступа к информационным ресурсам, и описание основных требований к ее реализации и функционированию.

Информационные системы (ИС) для поддержки научно-образовательной деятельности оперируют с различного рода информацией, как например, публикации, электронные документы, электронные коллекции, онтологические описания, массивы данных, логические описания и др. Как правило, эти ресурсы, востребованные разными группами исследователей, оказываются недоступными из-за невозможности их поиска и идентификации. Семантические связи между информационными ресурсами повышают их ценность и предоставляет дополнительные возможности для информационного поиска и идентификации. Данные, интегрированные в открытое семантическое пространство, представляют

собой совокупность знаний о некоторой предметной области в виде семантической структуры, на основе которой становятся возможными качественно новые наукометрические измерения и исследования структурных свойств корпуса научных знаний.

Технология информационных хранилищ - это разновидность интегрированной технологии, предназначенной для реализации процедур, методов и средств хранения и применения комплекса БД в решении задач пользователя. Обширные массивы данных можно хранить на одном или нескольких серверах. Эти массивы обычно называют информационными хранилищами [1].

Представление данных. В качестве предметной области были рассмотрены материалы научного наследия в области ИТ-технологий. В информационном пространстве события, факты и любые другие сущности реального мира существуют только в форме документов. Документ является основным объектом, участвующим в любой информационной системе (ИС). Основной функцией документа является информационная, т.е. способность удовлетворять информационную потребность человека. Множество документов, содержащих фактографическую информацию, имеющих одинаковую физическую структуру и логическое, информативное назначение, образуют коллекции. Коллекции, согласно [2], характеризуются своими описаниями и описаниями структуры документов, из которых она состоит, и представляют собой систематизированную совокупность документов, объединенных по какому-либо критерию принадлежности, например, по содержанию, назначения, способу доступа и т.д., снабженные метаописанием (метаданными) в соответствии стандартами и схемами данных. Коллекции могут вкладываться друг в друга, но у одной коллекции может быть только одна родительская коллекция. Любой документ может помещаться в несколько коллекций.

Метаданные создаются в процессе явной или неявной каталогизации и соответствуют той или иной общепринятой сфере. Для извлечения метаданных используются алгоритмы разбора титульной страницы, извлечения ключевых слов при их отсутствии, алгоритм построения реферата, на основе графа связного с текстом. Основной каталог информационных ресурсов сервера метаданных ИС строится в соответствии со схемой метаданных. Разработанная схема учитывает основные требования схемы метаданных Dublin Core. В дальнейшем нашу схему метаданных будем называть внутренней.

Хранение данных. Для долговременного хранения документов использовался институциональный репозиторий Dspace [3], в силу его способности расширять список поддерживаемых метаданных, что позволяет настраивать его на различные предметные области. Для поддержки процесса наполнения полнотекстовых баз были настроены рабочие процессы и пользовательский интерфейс: созданные профили метаданных были зарегистрированы в системе DSpace.

Обмен данными. Для реализации обмена метаданными между DSpace, в соответствии с расширенным профилем, был создан сервис на языке XSLT, выполняющий преобразование схем метаданных из внутренней схемы DSpace в схему сервера метаданных и в схему Dublin Core с использованием квалификаторов. Также реализован OAI сервис, который в пакетном режиме периодически, в соответствии с расписанием, проводит синхронизацию метаданных репозитория и сервера метаданных. Для заполнения основного каталога метаданных в соответствии с созданными схемами метаданных используются контролируемые словари из справочного блока системы [4, 5].

Функционирование среды. Основывается на использовании протоколов Z39.50 и LDAP. При этом предусмотрены механизмы преобразования данных из предметных схем в абстрактную схему протокола Z39.50. Виртуальная среда состоит реестра объектов и ресурсов, основного сервера Z39.50, нескольких функциональных модулей, а также web-интерфейса с публичным и административным разделами для доступа к различным функциям среды. Для каждого источника устанавливается отдельный сервер Z39.50, который осуществляет преобразование данных из схемы источника в абстрактную схему данных [6]

Сбор данных из внешних источников. Выбранная технология интеграции разрабатываемой информационной системы для поддержки исследований по научному наследию с ЦД (цифровой депозитарий) позволяет использовать любые другие реализации ЦД, поддерживающие протокол OAI. Для работы с внешними ЦД, поддерживающими протокол OAI, реализован сервис, преобразующий схему метаданных Dublin Core во внутреннюю схему метаданных системы.

Извлекаемые из внешних ЦД метаданные также помещаются в базу данных электронной библиотеки (сервера метаданных). Если для конкретного ЦД это возможно, то метаданные извлекаются в XML-схеме с дальнейшим преобразованием в схему ГОСТ. Если схема данных внешнего ЦД неизвестна, то метаданные преобразуются из схемы, основанной на Dublin Core, в схему метаданных ГОСТ.

Для интеграции электронной библиотеки с внешними системами и приложениями по протоколу OAI используются сервисы преобразований внутренней схемы метаданных, например, в схему MARC XML 20 или RUSMARC и др. При необходимости этот же подход может быть использован для преобразования метаданных в другие схемы семейства MARC.

Поиск данных. Функциональность поиска документов (информационных ресурсов) доступна конечным пользователям тремя путями: через пользовательский интерфейс информационной системы (сервера метаданных), через специализированный поисковый сервис (для внешних приложений) и через пользовательский интерфейс ЦД DSpace [7]. Однако в любом случае сами документы всегда хранятся в ЦД, поэтому, через какой бы интерфейс пользователь ни нашел необходимый ему документ, документ будет напрямую извлечен по протоколу HTTP из ЦД.

Контроль доступа пользователей. На всех уровнях пользовательских интерфейсов (сервер метаданных, сервисы и ЦД) контроль доступа пользователей к ресурсам информационной системы осуществляется на основании идентификационной информации в соответствии с протоколом LDAP. DSpace имеет встроенную поддержку LDAP, а в программируемых сервисах и в сервере метаданных поддержка обеспечивается сервисами информационной системы.

Предлагаемая технология поддержки больших информационных хранилищ основана на клиент-серверной архитектуре ИС и отвечает требованиям, описанным в [12]:

- унификация процесс обмена результатами научных исследований;
- оперирование с данными и документами, интегрированными в открытое семантическое пространство;
- предоставление сервисов преобразования разнородных ресурсов, реализующих средства описания, представления, автоматического связывания ресурсов, а также взаимодействие с поисковыми и классификационными механизмами в соответствии с потребностями пользователей.

Определение требований к системе идентификации информационных ресурсов, системе управления метаданными, системе управления информационными ресурсами

Исходя из целей создания ИС для поддержки научно-образовательной деятельности и анализа существующих систем, направленных на поддержку научных исследований определены требования к системе идентификации информационных ресурсов [6].

Известно, что неструктурированные информационные ресурсы целиком представляются в виде онтологий. Для этого необходимо выделить термины предметной области. Термины можно выделить либо путем индексирования всех информационных ресурсов, либо с использованием тезауруса.

В процессе научно-образовательной, деятельности много времени и сил отнимает работа с литературными источниками [7], разного рода материалами и документами: поиск необходимых документов, систематизация и классификация документов в соответствии с поставленной задачей. Для удовлетворения информационных потребностей современных пользователей в сфере научно-образовательного деятельности необходима поддержка тонких функций поиска и классификации информации, а также просмотр ресурсов по категориям (рубрикам) и словарям-классификаторам. Наиболее важной задачей является задача систематизации ресурсов (тематическая классификация), для решения которой необходимо четко определить состав логико-семантических категорий (фасетов) и ключевых терминов (понятий), покрывающих избранную достаточно узкую предметную область, интересующую пользователя [8].

В исследовании были сформулированы основные требования к информационной системе поддержки в научно-образовательной деятельности [9]. В частности, такая система должна поддерживать надежное, долговременное и защищенное от исчезновения, хранение информации; включать большое количество словарей-классификаторов [10, 11] для обеспечения идентификации и классификации ресурсов; поддерживать слабо структурированные информационные ресурсы, взаимосвязи между информационными ресурсами; включать интеллектуальные службы обслуживания запросов пользователя, а также программные интерфейсы для поддержки аналитической работы пользователя; удовлетворять требованиям интероперабельности и т.д. (Рисунки 1, 2).

Был выделен набор наиболее общих функциональных требований к ИС поддержки научно-образовательной деятельности.

1) *Сбор информационных ресурсов.* Для сбора информации необходимо использовать различные варианты ввода данных:

- ввод данных пользователями;
- сбор данных в Internet посредством специальных программных агентов (пауков);
- обмен данными с другими ИС.

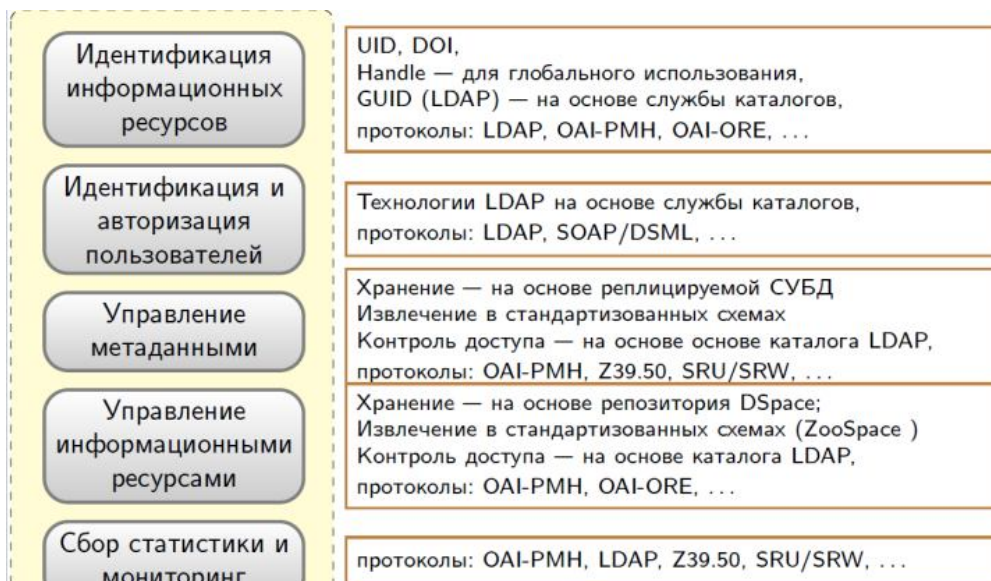


Рис. 1 Система управления информационными ресурсами

Страница администратора Самбетбаева Мадина

Редактирование записи N 265 в коллекции *publications* (22)

Проверьте правильность приведенной ниже формы:

url= АБРАМОВВГТРИФОНОВНПТРИФОНОВАГНВВЕДЕНИЕВЯЗЫКПАСКАЛЬ1988 (:15:28::15:28::ru)
url1= 262af97ab778903230e3b1862f0c341f

ID	Название	Значение
36	Идентификатор основной (UID)*	FB7AC580
42	Дополнительные идентификаторы (:)	
3	Идентификаторы авторов публикации (:author: по порядку)	:Абрамов:Трифонов:Трифорова:
50	Имена авторов публикации	Абрамов Владимир Геннадьевич, Трифонов Николай Павлович, Трифорова Галина Николаевна
4	Организации авторов	МГУ имени М.В. Ломоносова, Факультет вычислительной математики и кибернетики

Рис. 2 Идентификация информационных ресурсов

2) *Релевантность документов*. При автоматическом сборе информации в Internet может накапливаться и нерелевантная или малорелевантная информация для данной ИС поддержки научно-образовательной деятельности. Решение проблемы возможно следующими способами:

- Создание подробных форматов представления метаданных о ресурсах и структурированных справочников для тематической классификации ресурсов. ИС поддержки научно-образовательной деятельности должна вкладывать описания в метаданных на веб-страницы и предоставлять интерактивные средства пользователям для создания метаданных определенного формата при размещении ресурсов.

- Разделение информационных ресурсов в зависимости от варианта поступления в систему (размещенные экспертами/пользователями и пауком, а также указание степени достоверности информации с учетом ее источника.

- Указание поисковым средствам пространства поиска и классификации информации, а также задание критериев оценки качества введенной информации

– Использование схем классификации ресурсов согласно потребностям пользователей и классификация ресурсов в соответствии с этими схемами

3) *Актуальность, полнота, достоверность происхождения документов.* Способы решения проблем актуальности и полноты аналогичны способам решения проблемы охвата ресурсов. Способы определения достоверности происхождения информации следующие:

– для интерактивного ввода информация вводится только аутентифицированными пользователями;

– для автоматизированных систем сбора накладывание ограничений на область действия агента, выполняющего сбор информации;

– для обмена информацией с другими ИС задание фильтров на импортируемые информационные ресурсы;

– для всех способов ввода должна выполняться проверка и классификация всей введенной информации.

4) *Использование интеллектуальных служб обработки запросов пользователя.* Службы обработки запросов пользователей должны обеспечивать поиск по атрибутам, полнотекстовый поиск, просмотр ресурсов по категориям, семантический поиск (необязателен).

5) *Извлечение знаний.* Использование частичной автоматизации извлечения знаний. В основе подхода лежит представление смысла текста в виде семантической сети, принцип построения которой основывается на использовании частоты совместной встречаемости понятий в тексте. Пользователю сеть представляется в виде тематического дерева (дерево ключевых терминов и связанных с ними понятий), что позволяет выполнять навигацию и существенно облегчает процесс исследования текста и поиска требуемой информации. Данный подход также используется для решения таких задач, как автоматическое реферирование, тематическая классификация и кластеризация текстов, семантический поиск и т.д. Кроме того к ИС поддержки научно-образовательной деятельности, работающие с разными типами информационных ресурсов, предъявляются следующие требования

6) *Поддержка не централизованных архитектур информационных систем.* Это требование является необходимым условием для полноты, аутентичности и актуальности информации. Опыт эксплуатации ИС поддержки научно-образовательной деятельности показал сложность создания централизованных научных систем, охватывающих научную информацию в определенной области науки, или в какой-либо стране.

7) *Структурированность информационного пространства.* Для поддержки сложных функций поиска и классификации информации помимо хранения полнотекстового описания, необходимо реализовывать поиск по атрибутам, полнотекстовый поиск, просмотр ресурсов по категориям и словарям-классификаторам. При этом выбор классификаторов определяется степенью специализации системы.

8) *Адаптивное представления информации.* Для повышения скорости поиска и точности подбора информации пользователем без потери качества поиска ИС поддержки научно-образовательной деятельности должна учитывать запросы пользователей, их компетентность при работе с ИС поддержки научно-образовательной деятельности, ограничения по времени. ИС поддержки научно-образовательной деятельности должна обеспечивать возможность пользователю получать различные уровни абстракции при представлении информации от

кратких описаний для максимального быстрого поиска, до очень подробных описаний информационных объектов.

9) *Историчность информации.* Спецификой научной информации является ее быстрое устаревание и потеря актуальности. Для многих типов информационных ресурсов важно хранить всю информацию о всех изменениях и иметь возможность восстановить состояние ресурса на любой момент времени. Например, информация об авторах может меняться со временем при смене персоной фамилии, места работы. Необходимо учитывать переформирования и переименования организаций, наименования географических объектов, которые также могут меняться. Поэтому необходимо учитывать временной фактор и использовать актуальную информацию для сущностей, связанных с промежутками времени. При распознавании сущностей необходимо обеспечивать выполнение запросов на определенный момент времени в прошлом, то есть создание среза истинности информации о сущностях на произвольную дату.

10) *Архив.* Как было отмечено выше, большая часть научной информации быстро устаревает. Но существуют информационные ресурсы, к которым необходимо длительное время обеспечивать доступ. К таковым, например, относятся документы, имеющие длительную юридическую силу, патенты или мультимедийная информация об исторических событиях, которые могут быть востребованы в любой период времени. Кроме того, научные отчеты институтов, речи ученых могут также иметь огромную историческую ценность, приобретая со временем свою значимость. Поэтому системы должны поддерживать возможность длительного хранения информационных ресурсов с возможностью их восстановления.

В условиях работы в распределенной среде к ИС поддержки научно-образовательной деятельности предъявляются требования:

- поддержка принятых стандартов метаданных для экспорта и импорта данных;
- поддержка протоколов обмена информации с другими информационными системами;
- поддержка возможности ссылки на внутренние ресурсы как в интерфейсах пользователей, так и на системном уровне.

Разработка концептуальных моделей хранения электронных ресурсов на основе выработанных корпоративных стандартов

Концептуальная онтологическая модель ИС поддержки научно-образовательной деятельности основана на стандартах ЭБ RM OAI (Open Archival Information System) и DELOS DLRM (The DELOS Digital Library Reference Model Foundations for Digital Libraries). Понятия, составляющие онтологию ЭБ, условно делятся на предназначенные для:

- описания контента предметной области;
- формирования тезауруса предметной области;
- описания тематических коллекций;
- описания задачи интеграции контента библиотеки с данными источников из

LOD (Linked Open Data).

Между группами понятий онтологии ЭБ определены семантически значимые связи [13-15].

Введем следующие формальные определения, необходимые для описания онтологии.

Определение 1) *Контент библиотеки* $C = \langle IR, A, IO \rangle$ определяется типами ее информационных ресурсов, описанных связанными с ними наборами атрибутов A и набором входных данных, определяющих информационные объекты IO , которые являются непосредственно объектами, хранящимися в библиотеке.

Определение 2) *Тезаурус библиотеки* $TH = \langle T, R \rangle$ определяется терминами T и связями R между ними. Набор терминов T , составляющих описание предметной области, строго задан.

Определение 3) *Семантические метки* $M = \{M_i\}$ информационного объекта — это термины, которые не попали в тезаурус, но являются необходимыми для специфицирования тематики информационного объекта. Семантические метки не связаны, в отличие от терминов тезауруса, связями между собой или с терминами тезауруса, но дают возможность дополнительного тематического разделения информационных объектов в рамках предметной области.

Определение 4) *Задача интеграции данных библиотеки* $IT = \langle DS, R, A, M, D, D_s \rangle$ с внешними источниками DS определяется типами ресурсов библиотеки и набором их атрибутов A , отображением M ресурсов R на схему источника данных S и набором связей D_s с данными из источника.

Определение 5) Коллекция информационных объектов $C = \langle IO, T, M, DS \rangle$ представляет собой набор объектов, объединенных на основе совокупности признаков:

- 1) по их термину тезауруса предметной области;
- 2) по семантическим меткам;
- 3) по источнику данных, из которого поступили объекты.

Определение 6) *Семантически значимыми связями библиотеки* $P = \{P_i\}$ назовем связи, определенные между контентом библиотеки, ее предметной областью (тезаурусом), семантическими метками и объектами источника данных.

Авторами данного исследования выделены следующие основные связи:

$P_1(t, io)$ – термин тезауруса – информационный объект;

$P_2(io, t)$ – информационный объект – термин тезауруса;

$P_3(r, s)$ – информационный ресурс – класс объектов источника, где информационный ресурс — это общее определение для информационных объектов, хранящихся в системе; таким образом, фактически информационные объекты являются экземплярами информационных ресурсов;

$P_4(a, s_a)$ – атрибут информационного ресурса – свойство класса источника;

$P_5(io, o_s)$ – информационный объект – экземпляр класса из источника данных;

$P_6(m, io)$ – семантическая метка – информационный объект;

$P_7(io, m)$ – информационный объект – семантическая метка.

Разработка модели представления метаописаний электронных ресурсов

Разработана модель представления метаописаний электронных ресурсов ЭБ. Инфраструктура для представления и обмена метаданными.

Информация о документах системы, их атрибутах и методах поддерживается сервером метаданных, содержащим метаописания системы и метаописания отдельных коллекций документов в основном каталоге метаданных (служебные метаданные, пользовательские интерфейсы, ассоциативные связи между документами и права доступа к документам). Документы Основного каталога содержат помимо описания схемы метаданных QDC описание служебных метаданных, которые описывают структуру объектов, пользовательские

интерфейсы, ассоциативные связи между документами и права доступа к документам.

Сервер метаданных является отдельной частью системы, содержащей описание информационной модели предметной области, параметров настройки стандартных функций системы. По информации сервера метаданных осуществляется динамическая генерация схем базы данных системы и ведение программного служебных баз данных, в которых хранятся данные, обеспечивающие поддержку стандартных функций системы, динамически определяемые отношения между документами и динамическое распределение ресурсов системы. Для заполнения основного каталога метаданных в соответствии с созданными схемами метаданных используются контролируемые словари из справочного блока системы.

Инфраструктура для представления и обмена метаданными.

При размещении в ЭБ материалы снабжаются унифицированным метаописанием. В онтологии ИС для поддержки научно-образовательной деятельности выделены сущности: Персоны, Факты, Документы, Публикации. В настоящий момент для заполнения отдельных полей были созданы следующие контролируемые словари:

- персоны – авторитетные файлы;
- классификационные индексы;
- списки учреждений, организаций и предприятий;
- типы документов;
- географические названия;
- типы содержимого документов.

В дальнейшем их список будет расширен.

Сервер метаданных поддерживает следующий набор служб:

- публикация/регистрация новых наборов данных;
- база метаданных для поиска данных по атрибутам;
- доступ к гетерогенным ресурсам посредством брокера ресурсов;
- контроль аутентификации и доступа;
- мониторинг информационных ресурсов и ресурсов ввода/вывода;
- распределенное исполнение служб.

Соответствующие службы могут регистрироваться в объектно-ориентированной базе данных и вызываться для обработки любого набора данных, хранящегося в репозитории. Комбинация возможности доступа к данным через базу метаданных, служб каталогов, зарегистрированного набора методов обработки данных позволяет решать задачи для создания среды обработки данных в соответствии с рисунком 3.

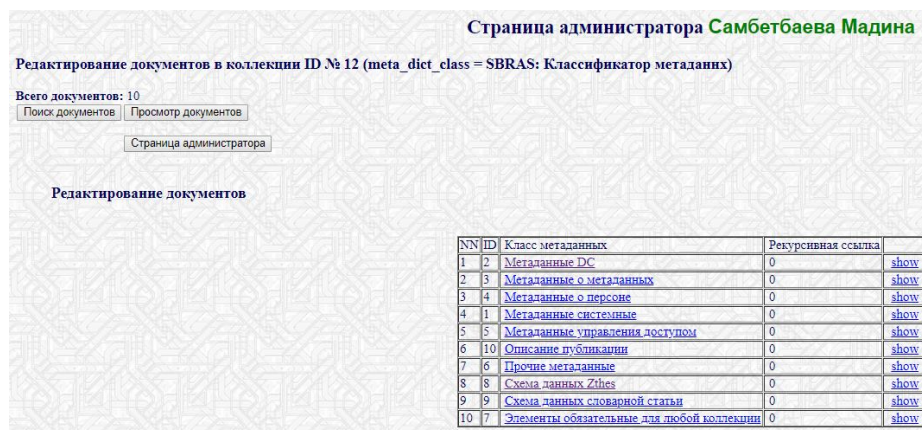


Рис. 3 Web-страница представления метаописаний электронных ресурсов

Заключение

В статье изложены технологии поддержки больших информационных хранилищ и организация разграниченного пользовательского доступа к этой информации. Разработка концептуальных модели документального и фактографического поиска в интеллектуальных информационных системах, работающих с документами достаточно произвольной структуры. Предложена модель классификации документов интеллектуальной информационной системы, основанная на использовании онтологического проектирования. Показано, что при создании фактографических информационных систем целесообразно следующее понимание факта: содержащаяся в тексте и метаданных документа совокупность связей между сущностями, описываемыми в онтологии информационной системы.

Литература

1. Михайлов, А. И., Черный А. И., Гиляревский Р. С. Научные коммуникации и информатика. — М.: Наука, 1976.
2. Арский, Ю. М., Гиляревский Р. С., Туров И. С., Черный А. И. Инфосфера: Информационные структуры, системы и процессы в науке и обществе. — М.: ВИНТИ, 1996.
3. Шокин, Ю. И., Федотов А. М., Барахнин В. Б. Проблемы поиска информации. — Новосибирск: Наука, 2010.
4. Марчук, А. Г. О распределенных фактографических системах // Труды 10 Всероссийской научн. конф. «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» (RCDL'2008). — Дубна, 2008. С.93–102.
5. Berners-Lee, T., Hendler, J., Lassila, O. The Semantic Web // Scientific American. 2001. Vol. 284. No. 5. P. 34–43.
6. Ранганатан, Ш. Р. Классификация двоеточием. Основная классификация / пер. с англ. — М.: ГПТНБ СССР, 1970.
7. Леонова, Ю.В., Федотов А.М. Создание прототипа системы управления информационными ресурсами // Совместный выпуск по матер. междунар. конф. «Вычислительные и информационные технологии в науке, технике и образовании» (СITech-2018) и Вестник ВКГТУ им. Д. Серикбаева, Вычислительные технологии. — Усть-Каменогорск – Новосибирск. - 2018. — Т.1, Ч.1. — С.47-56.
8. Yu. V. Leonova, A.M. Fedotov Creation of a prototype of a scientific information system for managing information resources // Proceedings of the XX International Conference “Data Analytics and Management in Data Intensive Domains”

(DAMDID/RCDL'2018). – М.: Lomonosov Moscow State University. - 2018. – P. 362-363.

9. Шрейдер, Ю.А. Об одной модели семантической информации // В сб.: Проблемы кибернетики. – М.: Наука.- 1965. – №13 – С. 233-240.

10. Sambetbayeva, M.A., Fedotova, O.A., Fedotov, A.M., Batyrkhanov, A.G. Conceptual model of information system for resources storage // Совместный выпуск по матер. междунар. конф. «Вычислительные и информационные технологии в науке, технике и образовании» (CITech-2018) и Вестник ВКГТУ им. Д. Серикбаева, Вычислительные технологии. – Усть-Каменогорск – Новосибирск. - 2018. – Т.1, Ч.3. – С.13-27.

11. M.A. Sambetbayeva, O.A. Fedotova, A.M. Fedotov Multilingual Thesaurus in Information System for Scientific and Educational Activity Support // Proceedings of the XX International Conference “Data Analytics and Management in Data Intensive Domains” (DAMDID/RCDL'2018). – М.: Lomonosov Moscow State University. - 2018. – P. 360-362.

12. ISO 25964-1:2011 Information and documentation – Thesauri and interoperability with other vocabularies. - Geneva: International Organization for Standardization. – 2011. – P.1. – 152 p.

13. ANSI/NISO. Z39.19: 2005 Guidelines for the construction, format and management of monolingual controlled vocabularies. – NISO Press: Bethesda. - 2005. – 135 p.

14. Когаловский, М.Р. Метаданные, их свойства, функции, классификация и средства представления // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Труды XIV Всерос. науч. конф. RCDL-2012. – Переславль-Залесский: Университет города Переславля. – 2012. - С. 25-36.

15. Fedotov, A.M., Tusupov, J.A., Sambetbayeva, M.A., Sagnayeva, S.K., Baranov, A.A., Nurgulzhanova, A.N., Yerimbetova, A.S. Using the thesaurus to develop it inquiry systems // Journal of Theoretical and Applied Information Technology. - 2016. - Vol.86. - Issue 1. - P.44-61.

МОДИФИЦИРОВАННЫЙ АЛГОРИТМ ОПТИМИЗАЦИИ СЕРЫХ ВОЛКОВ ДЛЯ SMART-ТЕХНОЛОГИИ ПРОГНОЗИРОВАНИЯ И УПРАВЛЕНИЯ СЛОЖНЫМИ ОБЪЕКТАМИ

Самигулина Г.А., Самигулина З.И.

galinasamigulina@mail.ru, zarinasamigulina@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Казахстанско-Британский технический университет, Казахстан

Аннотация. *Статья посвящена разработке модифицированного алгоритма на основе оптимизации серых волков (Grey Wolf Optimization, GWO) для реализации Smart –технологии прогнозирования и управления сложными объектами на базе многофункциональной искусственной иммунной системы. Рассмотрены*

перспективы применения GWO для решения задачи выделения информативных признаков при обработке многомерных производственных данных с целью снижения размерности и улучшения работы моделей распознавания образов. Осуществлён сравнительный анализ эффективности использования GWO для различных классификаторов. Получены результаты моделирования разработанного модифицированного алгоритма для многофункциональной искусственной иммунной системы.

Современные промышленные предприятия работают на основе сложных распределённых систем управления, которые осуществляют мониторинг и управление технологическими процессами, занимаются сбором, аналитикой и архивированием производственных данных. Подобные системы накапливают данные об объекте и позволяют своевременно корректировать технологические параметры, а также прогнозировать износ оборудования и перераспределять нагрузку между узлами. Ключевую роль в работе таких систем занимает обработка производственных данных в реальном времени. При функционировании сложных производственных процессов возникает переизбыток данных, большая часть которых не может быть обработана в реальном времени. В связи с этим актуальна разработка современных Smart – технологий на основе инновационных эвристических алгоритмов искусственного интеллекта (ИИ) для прогнозирования и управления сложными объектами.

Одним из перспективных методов ИИ является алгоритм оптимизации серых волков, который впервые был предложен Mirjalili S. и др. в 2014 г. [1]. Данный метод широко используется в области управления сложными объектами. В работе [2] представлен модифицированный алгоритм GWO для решения глобальных инженерных задач оптимизации. Статья [3] посвящена разработке ПИД регулятора на основе алгоритма оптимизации серых волков. В исследованиях [4] представлена оптимальная настройка параметров ПИД регулятора для управления роботом. В работе [5] алгоритм GWO используется для контроля частоты нагрузки взаимосвязной энергосистемы. Статья [6] посвящена разработке улучшенного алгоритма оптимизации серых волков. Представлен нечёткий ПИД регулятор на основе GWO для регулирования частоты энергосистемы. В работе [7] метаэвристический алгоритм GWO был применён к взаимосвязной гидротермальной энергетической системе для автоматического управления процессом генерации. Полученные результаты сравниваются с методами оптимизации роя частиц (Particle Swarm Optimization, PSO) и дифференциальной эволюции (Differential evolution, DE). Статья [8] посвящена применению алгоритма оптимизации серых волков для разработки системы управления двигателем постоянного тока. В работе [9] представлен алгоритм GWO для нескольких ПИД-контроллеров с целью улучшения работы ветряной турбины при низком напряжении и для отслеживания точки максимальной мощности. В статье [10] алгоритм GWO используется для оптимального проектирования гибридной системы возобновляемой энергии. Таким образом, алгоритм оптимизации серых волков является перспективным для исследований в области управления сложными объектами.

Постановка задачи формулируется следующим образом: необходимо разработать модифицированный алгоритм на основе метода оптимизации серых волков для многофункциональной искусственной иммунной системы в рамках реализации Smart-технологии прогнозирования и управления сложными

объектами. Осуществить сравнительный анализ полученных результатов с другими алгоритмами классификации.

Алгоритм оптимизации серых волков используется для решения задачи предварительной обработки производственных данных и редукции неинформативных параметров. В основе алгоритма GWO лежит следующая иерархия волков [1]: α – волки лидеры; β – волки советчики, помогающие α в принятии решений; δ – подчиняются группе α и β , но доминируют над последним уровнем; ω – последний уровень иерархии, подчиняется всем вышестоящим уровням. Метод оптимизации серых волков моделирует поведение охоты и состоит из этапов:

- отслеживание, преследование и подход к жертве;
- преследование, окружение до тех пор, пока жертва не перестает двигаться;
- нападение на жертву.

Математическая модель поведения волков во время охоты имеет следующий вид:

$$\begin{aligned} \vec{D} &= |\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)| \\ \vec{X}(t+1) &= \vec{X}_p(t) - \vec{A} \cdot \vec{D} \end{aligned} \quad (1)$$

где t – текущая итерация; \vec{A}, \vec{C} – векторы коэффициенты, рассчитываемые по формуле: $\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a}$; $\vec{C} = 2 \cdot \vec{r}_2$, значение \vec{a} линейно уменьшается от 2 до 0, согласно порядку итерации и \vec{r}_1, \vec{r}_2 – случайные векторы $[0,1]$; \vec{X}_p – позиция вектор добычи, \vec{X} – позиция вектор волков. Процесс охоты инициируется α , в то время как β и δ могут помогать. В математической модели (1) показано, что α , β и δ представляют собой наилучшее решение относительно потенциального расположения добычи. Первые три лучших решения сохраняются и другие агенты (agents) обязаны обновлять свои позиции согласно позиции агентов наилучшего поиска (best search agents) на основе уравнений вида:

$$\vec{D}_\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}|, \vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \vec{D}_\delta = |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}| \quad (2)$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{D}_\alpha), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot (\vec{D}_\beta), \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot (\vec{D}_\delta) \quad (3)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (4)$$

где вектор \vec{A} – представляет собой случайное значение, лежащее в диапазоне $[-2a, 2a]$, а значение параметра a уменьшается от 0 до 2, согласно текущей итерации.

Рассмотрим эффективность применения алгоритма GWO для решения задачи выделения информативных признаков на примере нескольких алгоритмов классификации. Для исследования выбрана база данных (БД) DB_D302 суточных замеров показаний с датчиков абсорбера среднего давления технологической Установки 300 по очистке газа от кислых компонентов предприятия ТенгизШевройл (например, LIC31053–датчик уровня; TI31050–датчик

температуры; PDI31003–перепад давления; FIC31005–датчик расхода; QRAH31001– анализатор H_2S в газе и т.д.). Размерность базы данных R=11x700, 7700 экземпляров данных. Для решения задачи классификации выделено 3 класса: 1 класс – приоритет тревоги «аварийный», сигнализация alarm включена; 2 класс – приоритет тревоги «высокий», сигнализация alarm выключена; 3 класс – приоритет тревоги «низкий», сигнализация alarm выключена.

На рис. 1 и рис. 2 показан сравнительный анализ использования алгоритма GWO для предварительной обработки данных в сочетании с различными классификаторами: Наивный Байесовский алгоритм (Naive Bayes, NB); Глубокое обучение (Deep Learning, DL); Деревья решений (Decision Tree, DT), Случайный лес (Random Forest, RF), Метод опорных векторов (Support Vector Machine, SVM).

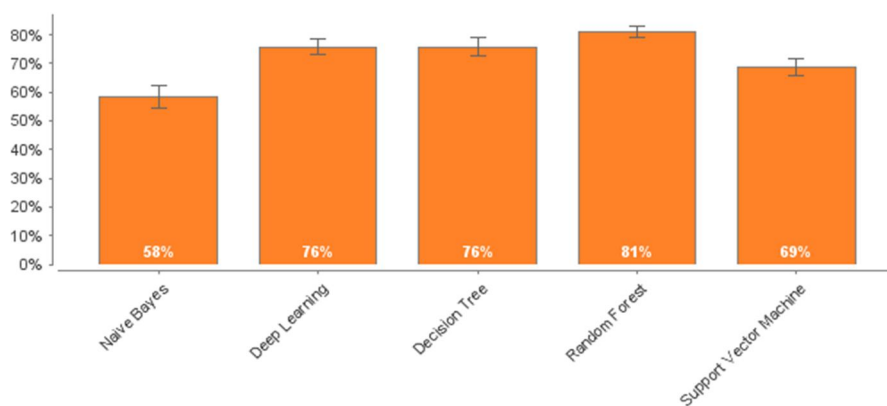


Рис. 1 Эффективность моделирования базы данных DB_D302 без предварительной обработки данных на основе GWO

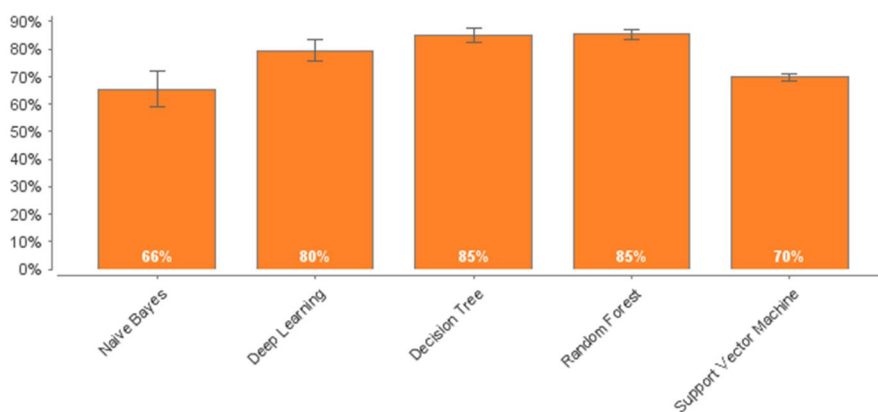


Рис. 2 Эффективность моделирования базы данных DB_D302 после предварительной обработки данных на основе GWO

Сравнительный анализ результатов моделирования (Таблица 1) осуществлён по двум основным параметрам: эффективность (accuracy) и время моделирования (run time). Представленные алгоритмы классификации значительно улучшили точность прогнозирования после использования предварительной обработки данных на основе GWO.

Таблица 1. Сравнительный анализ эффективности применения алгоритма GWO

Алгоритм распознавания образов	Распознавание образов без процедуры выделения информативных признаков на основе GWO		Эффективность распознавания образов после выделения информативных признаков на основе GWO	
	Эффективность	Время моделирования	Эффективность	Время моделирования
NB	58%	722 мс	65,5%	6 сек
DL	76%	3 сек	80%	12 мин 51 сек
DT	76%	885 мс	85%	16 сек
RF	81%	15 сек	85,4%	12 мин 20 сек
SVM	69%	10 сек	69,8%	2 мин 21 сек

Для разработки Smart технологии управления сложными объектами [11] эффективность представленных моделей ИИ (при тестировании на реальных производственных данных) является недостаточной. Актуальна разработка модифицированных алгоритмов для многофункциональной искусственной иммунной системы. Данный подход хорошо зарекомендовал себя при решении задачи распознавания образов на границе классов, обладает памятью, самоорганизацией и высокой адаптивностью. Модифицированный алгоритм на основе GWO состоит из следующих основных шагов.

Алгоритм.

1. Подключение к хранилищу данных распределённой системы управления предприятием. Формирование БД параметров сложного объекта, состоящей из показаний с датчиков.

2. Предварительная обработка данных на основе алгоритма оптимизации серых волков. Редукция неинформативных параметров. Формирование оптимальной БД параметров, описывающих поведение сложного объекта.

3. Решение задачи классификации на основе алгоритмов искусственных иммунных систем.

3.1 Моделирование на основе алгоритма распознавания искусственной иммунной системой (Artificial Immune Recognition System, AIRS).

3.2 Моделирование на базе алгоритма клональной селекции (Clonal Selection, CLONALG).

3.3 Иммунносетевое моделирование (Artificial Immune System, AIS) [11].

5. Оценка эффективности моделирования на базе алгоритмов 3.1, 3.2, 3.3.

6. Выбор модифицированного алгоритма на основе лучшего прогностического результата.

7. Прогноз и принятие решений по управлению сложным объектом.

Результаты моделирования на основе данного подхода представлены на рис. 3. На рис. 3 показана эффективность применения алгоритмов AIRS (62%), CLONALG (63,8%) и AIS (85%) до предварительной обработки данных на основе GWO и результат моделирования модифицированных алгоритмов GWO-AIRS (73,2%), GWO-CLONALG (75,4%) и GWO-AIS (93,6%) после редукции неинформативных параметров на основе оптимизации серых волков. Наилучший прогностический результат показал модифицированный алгоритм GWO-AIS.

Таким образом, разработка модифицированных алгоритмов для многофункциональной ИИС при реализации Smart-технологии управления сложными объектами является перспективным направлением исследований и

позволяет повысить точность прогнозирования моделей за счёт подбора такой модификации алгоритмов, которая будет наиболее эффективна для анализа конкретной базы данных.

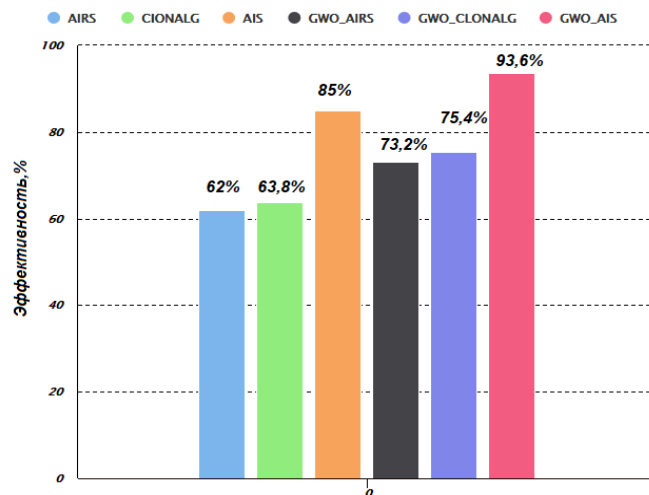


Рис. 3 Эффективность применения модифицированных алгоритмов искусственных иммунных систем на основе оптимизации серых волков

Работа выполнена по гранту №AP05130018 КН МОН РК по теме: "Разработка когнитивной Smart-технологии для интеллектуальных систем управления сложными объектами на основе подходов искусственного интеллекта" (2018-2020 гг.).

Литература

1. S. Mirjalili, S.M. Mirjalili, A. Lewis. Grey Wolf Optimizer // Advanced in Engineering Software - 2014. – Vol. 69. - P. 46-61.
2. N. Mittal, U. Singh., B.S. Sohi. Modified Grey Wolf Optimizer for Global Engineering Optimization // Applied Computational Intelligence and Soft Computing. – 2016. – P. 1-16.
3. P.B. Oliveira, H. Freire, E.J. Pires. Grey wolf optimization for PID controller design with prescribed robustness margins // Soft Computing. –Springer, 2016. – P. 1-13.
4. A.M. Şen, M. Kalyoncu. Optimal Tuning of PID Controller Using Grey Wolf Optimizer Algorithm for Quadraped Robot // Balkan Journal of Electrical and Computer Engineering. – 2018. – Vol. 6. - №1. – P.29-35.
5. D. Guha, B.C. Roy, P.K. Roy, S. Banerjee. Grey wolf optimization to solve load frequency control of an interconnected power system: GWO used to solve LFC problem // International Journal of Energy Optimization and Engineering. – 2016. – Vol.5. - №4. – P. 62-83.
6. B.P. Sahoo, S. Panda. Improved grey wolf optimization technique for fuzzy aided PID controller design for power system frequency control // Sustainable Energy, Grids and Networks. – 2018. – Vol.16. – P. 278-299.
7. D. K. Lai, B.M. Tripathy. Grey Wolf Optimizer Algorithm Based Fuzzy PID Controller for AGC of Multi-area Power System with TCPS // Procedia Computer Science.–2016.–Vol.92.–P.99-105.
8. A. Madadi, M.M. Motlagh. Optimal Control of DC motor using Grey Wolf Optimizer Algorithm // Technical Journal of Engineering and Applied Sciences.–2014.–Vol.4.–№4.–P. 373-379.

9. M.H. Qais., H.M. Hasanien., S. Alghuwainem. A Grey Wolf Optimizer for Optimum Parameters of Multiple PI Controllers of a Grid-Connected PMSG Driven by Variable Speed Wind Turbine // IEEE Access. – 2018. – Vol.6. – P.44120-44128.

10. A. Yahiaouia, F. Fodhilb, K. Benmansoura, M.Tadjinec, N.Cheggagab. Grey wolf optimizer for optimal design of hybrid renewable energy system PV-Diesel Generator-Battery: Application to the case of Djanet city of Algeria //Solar Energy. – Elsevier, 2017. – Vol.158. – P.941-951.

11. Samigulina G.A., Samigulina Z.I. Development of Smart technology for complex objects prediction and control on the basis of a distributed control system and an artificial immune systems approach // Advances in Science Technology and Engineering Systems Journal (ASTESJ).–2019, Vol 4. - Issue № 3. – P.75-87.

РАЗРАБОТКА МОДИФИЦИРОВАННОГО КООПЕРАТИВНОГО АЛГОРИТМА РОЯ ЧАСТИЦ С ВЕСОМ ИНЕРЦИИ ДЛЯ МНОГОФУНКЦИОНАЛЬНОЙ ИСКУССТВЕННОЙ ИММУННОЙ СИСТЕМЫ

Самигулина Г.А., Масимканова Ж.А.

galinasamigulina@mail.ru, masimkanovazh@gmail.com

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Казахский Национальный Университет им. аль-Фараби, Казахстан

Аннотация. *Статья посвящена разработке модифицированного кооперативного алгоритма роя частиц с весом инерции для многофункциональной Smart-системы прогнозирования и управления сложными объектами. Предварительная обработка данных по выделению информативных дескрипторов выполняется на основе кооперативного алгоритма роя частиц с весом инерции (Cooperative Particle Swarm Optimization with Inertia Weight, CPSOIW). Разработанный алгоритм позволяет более детально и эффективно исследовать пространство поиска, избежать ранней сходимости и предотвращает попадание в локальные оптимумы. Разработано программное обеспечение CPSOIW на языке программирования Python. Осуществлено сравнение CPSOIW алгоритма и алгоритма роя частиц с весом инерции (Inertia Weight Particle Swarm Optimization, IWPSO) на основе стандартных тестовых наборов данных.*

Существующая тенденция широкого использования интеллектуальных методов приводит к необходимости разработки и внедрения инновационных Smart-технологий прогнозирования и управления сложными нелинейными динамическими объектами с неопределённостями параметров. Особенностью таких систем является необходимость обработки огромного потока многомерных данных в режиме реального времени, поэтому актуально решение задачи выделения информативных дескрипторов. Построение оптимального набора данных представляет собой задачу многоцелевой оптимизации, при которой

минимальное количество выбранных дескрипторов обеспечивает максимальную точность классификации [1].

Анализ современных публикаций в области искусственного интеллекта показывает активное развитие алгоритмов роевого интеллекта (РИ), которые основаны на коллективном поведении биологических существ. Примерами являются: муравьиный алгоритм (Ant Colony Optimization, ACO) [2], алгоритм роя частиц (Particle Swarm Optimization, PSO) [3], пчелиный алгоритм (Artificial Bee Colony, ABC) [4], алгоритм кукушки (Cuckoo Search, CS) [5], алгоритм светлячка (Firefly Algorithm, FA) [6], алгоритм летучих мышей (Bat Algorithm, BA) [7], алгоритм серых волков (Grey Wolf Optimizer, GWO) [8], алгоритм эхолокации дельфинов (Dolphin Echolocation (DE) [9], алгоритм китов (Whale Optimization Algorithm, WOA) [10], алгоритм дрозофилы (Fruitfly Optimization Algorithm, FOA) [11] и др. Существующие вышеперечисленные алгоритмы РИ не являются универсальными. В зависимости от характера и объёма используемых данных возникает необходимость в создании эффективных модифицированных алгоритмов, которые обеспечивают параллельную обработку в режиме реального времени. Например, в статье [12] предлагается новый гибридный алгоритм на основе муравьиного и пчелиного алгоритма для выделения информативных дескрипторов и классификации. Оценка предложенного алгоритма с использованием 13 тестовых наборов данных показывает эффективность его применения. В работе [13] предлагается совместное использование бинарного алгоритма роя частиц (Binary Particle Swarm Optimization algorithm, BPSO) и алгоритма дерева решений для обработки многомерных данных и выбора информативных дескрипторов. Результаты моделирования на 11 наборах данных показывают, что данный метод превосходит наивный Байесовский классификатор и метод опорных векторов. В исследовании [14] описывается кооперативный алгоритм роя частиц с нечётким весом инерции и с улучшенным комплексным обучением (Enhanced comprehensive learning cooperative particle swarm optimization with fuzzy inertia weight, ECLCFPSO-IW) для решения проблемы ранней сходимости и предотвращения попадания в локальные оптимумы. В статье [15] для решения оптимального потока мощности в энергосистемах рассматриваются PSO алгоритмы с тремя разными весами инерции. Используются алгоритмы с постоянным весом инерции (Constant Inertia Weight, CIW), изменяющимся во времени весом инерции (Time-varying Inertia Weight, TVIW) и глобально-локальным лучшим весом инерции (Global-local best inertia weight GLbestIW)), которые позволяют анализировать влияние веса инерции на сходимость PSO алгоритма.

Постановка задачи формулируется следующим образом: необходимо на основе разработанного модифицированного кооперативного алгоритма роя частиц с весом инерции (Cooperative Particle Swarm Optimization with Inertia Weight, CPSOIW) решить задачу выделения информативных дескрипторов для многофункциональной Smart-системы прогнозирования и управления сложными объектами и протестировать CPSOIW алгоритм на стандартных наборах данных.

Модифицированный CPSOIW алгоритм разработан с использованием кооперативного алгоритма роя частиц CPSO (Cooperative particle swarm optimization) и алгоритма роя частиц с весом инерции IWPSO (Inertia Weight Particle Swarm Optimization).

Кооперативный алгоритм представляет собой сотрудничество агентов основного роя и подчинённых роев. Агенты подчинённых роев ориентированы на интенсивный локальный поиск и на ускорение скорости сходимости. Агенты

основного роя сосредоточены на глобальном поиске и на сохранении разнообразия решений.

Для предотвращения попадания в локальные оптимумы и ускорения скорости частиц вес инерции обновляется от итерации к итерации. На рисунке 1 представлена кооперация агентов, где каждый подчинённый рой представляет собой IWPSO алгоритм.

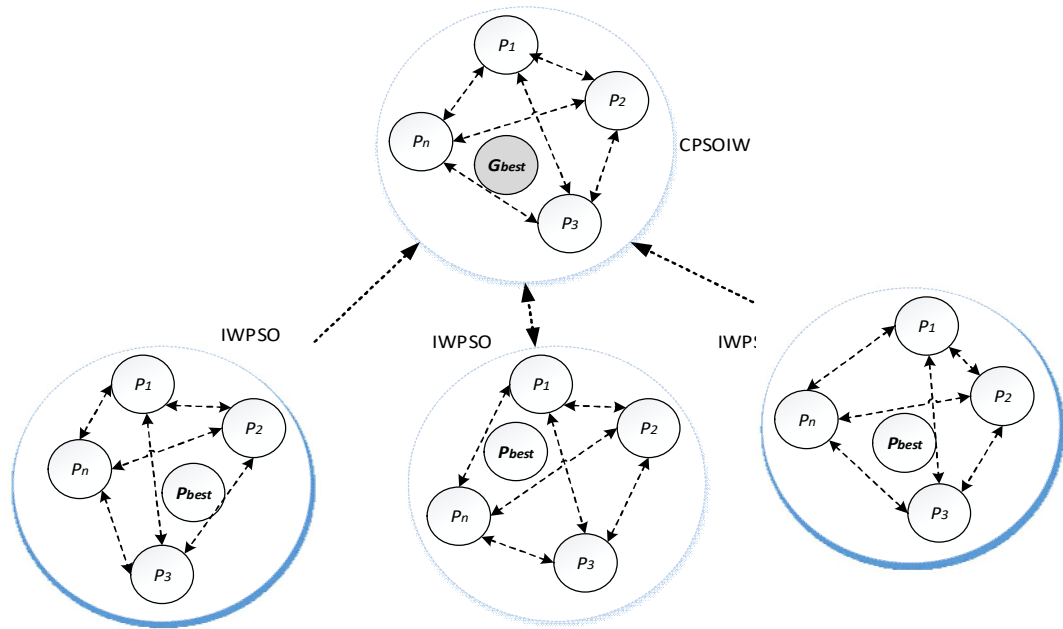


Рис. 1. Кооперация агентов CPSOIW алгоритма

Модифицированный CPSOIW алгоритм:

1. Ввод переменных $c_1, c_2, w_{max}, w_{min}$ и k_{max} , где c_1 и c_2 – коэффициенты ускорения, w_{max} – максимальное значение веса инерции, w_{min} – минимальное значение веса инерции, k_{max} – максимальное количество итераций.

2. Случайная генерация начальных положений x_i и скорости v_i частиц основного M и подчинённых S роев, где $i = 1, 2, \dots, n$.

3. Вычисление выбранной фитнес-функции для подчинённых S роев $f(x_i^s(t+1))$ и сохранение лучшего значения фитнес-функции $f(p_g^s)$.

4. Сравнение текущего значения фитнес-функции $f(x_i^s(t+1))$ и лучшего значения фитнес-функции $f(p_g^s)$ подчинённых роев.

5. Сравнение лучшего значения фитнес-функции подчинённых S роев $f(p_g^s)$ и глобального лучшего значения $f(p_g^M)$ основного роя, вычисленного на основе лучших значений подчинённых S роев.

6. Вычисление веса инерции w для всего роя. Вес инерции меняется от итерации к итерации и вычисляется по формуле:

$$w = w_{max} - \frac{w_{max} - w_{min}}{k_{max}} k, \quad (1)$$

где k – текущая итерация.

7. Обновление скорости частиц основного M роя v_i^M по формуле:

$$v_i^M(t+1) = wv_i^M(t) + r_1c_1 + r_2c_2(p_g^M - x_i^M(t)) + r_3c_3(p_g^s - x_i^M(t)), \quad (2)$$

где t – итерация, w – вес инерции, r_1, r_2, r_3 – равномерно распределённые случайные числа в интервале $[0,1]$, p_g^M – лучшее значение частиц основного роя, p_g^s – лучшее значение частиц подчинённых роев, x_i^M – положение частицы основного роя.

8. Обновление положения частиц основного M роя x_i^M по формуле:

$$x_i^M(t+1) = x_i^M(t) + v_i^M(t), \quad i = 1, 2, \dots n. \quad (3)$$

9. Проверка критерия останова $k < k_{max}$.

10. Выполнение 2.1 -2.4 шагов до окончания количества итераций.

11. Построение оптимального набора дескрипторов на основе лучших значений частиц основного M роя.

Для оценки эффективности предложенного алгоритма используются тестовые наборы данных UCI (UCI Machine Learning Repository), которые представлены в таблице 1. Выбранные наборы данных содержат различное количество дескрипторов и экземпляров.

Таблица 1. Описание базы данных

	База данных	Количество дескрипторов	Количество экземпляров
1	Sonar	61	208
2	Vehicle	19	846
3	Waveform	41	5000
4	Letter	17	20000

Выделение информативных дескрипторов на основе IWPSO алгоритма выполняется с использованием программного продукта WEKA [16].

Для построения оптимального набора дескрипторов на основе модифицированного CPSOIW алгоритма разработано программное обеспечение (ПО) на языке программирования Python. Данное ПО является модулем многофункциональной Smart-системы управления и прогнозирования сложными объектами.

В таблице 2 приведены параметры модифицированного алгоритма.

Таблица 2. Параметры модифицированного CPSOIW алгоритма

Параметры	Значения	Обозначение
P	20	количество частиц в рое
w_{min}	0.33	минимальное значение веса инерции
w_{max}	0.9	максимальное значение веса инерции
c_1	0.33	коэффициент ускорения
c_2	0.34	коэффициент ускорения
k_{max}	100	максимальное количество итераций

Сравнение модифицированного CPSOIW алгоритма с IWPSO алгоритмом приведено в таблице 3.

Таблица 3. Результаты моделирования

База данных	IWPSO алгоритм		CPSOIW алгоритм	
	Значение фитнес-функции	Количество выделенных дескрипторов	Значение фитнес-функции	Количество выделенных дескрипторов
Sonar	0,221	19	0,397	13
Vehicle	0,303	11	0,813	6
Waveform	0,252	16	0,026	15
Letter	0,478	11	0,837	4

Модифицированный CPSOIW алгоритм позволяет выделить информативные дескрипторы и построить оптимальный набор данных для дальнейшего решения задачи распознавания на основе искусственных иммунных систем.

Работа выполнена по гранту КН МОН РК по теме: №AP05130018 "Разработка когнитивной Smart-технологии для интеллектуальных систем управления сложными объектами на основе подходов искусственного интеллекта" (2018-2020 гг.).

Литература

1. Hegazy A., Makhloufa M.A., El-Tawelb Gh. S. Improved salp swarm algorithm for feature selection // Journal of King Saud University - Computer and Information Sciences. – 2018. – <https://doi.org/10.1016/j.jksuci.2018.06.003>. (Дата обращения: 31.05.2019)
2. Yaralidarani M., Shahverdi H. An improved Ant Colony Optimization (ACO) technique for estimation of flow functions from core-flood experiments // Journal of Natural Gas Science and Engineering. – 2016. – Vol. 33. – P. 624-633.
3. Huang S., Tian N., Wang Y., Ji Z. Particle swarm optimization using multi-information characteristics of all personal-best information // Springer Plus. – 2016. – Vol. 5 (1632). <https://doi.org/10.1186/s40064-016-3244-8>. (Дата обращения: 31.05.2019)
4. Chen X., Tianfield H., Li K. Self-adaptive differential artificial bee colony algorithm for global optimization problems // Swarm and Evolutionary Computation. – 2019. – Vol. 45. – P. 70-91.
5. Rajabioun R. Cuckoo optimization algorithm // Applied Software Computing. – 2011. – Vol. 11 (8). – P. 5508-5518.
6. Yang X.S. Firefly algorithm // Eng Optim. – 2010. – P. 221-230.
7. Yang X.-S. A new metaheuristic bat-inspired algorithm // Nature inspired cooperative strategies for optimization (NCSO 2010). – Springer, 2010. – P. 65-74.
8. Kohli M., Arora S. Chaotic grey wolf optimization algorithm for constrained optimization problems // Journal of Computational Design and Engineering. – 2018. – Vol. 5, Issue 4. – P. 458-472.
9. Kaveh A., Farhodi N. A new optimization method: dolphin echolocation // Advanced Engineering Software. – 2013. – Vol. 59. – P. 53-70.
10. Kaur G., Arora S. Chaotic whale optimization algorithm // Journal of Computational Design and Engineering. – 2018. – Vol. 5, Issue 3. – P. 275-284.

11. Pan W.-T. A new fruit fly optimization algorithm: taking the financial distress model as an example // Knowl Based Syst. – 2012. – Vol. 26. – P. 69-74.
12. Shunmugapriya P., Kanmani S. A hybrid algorithm using ant and bee colony optimization for feature selection and classification (AC-ABC Hybrid) // Swarm and Evolutionary Computation. – 2017. – Vol. 36. – P. 27-36.
13. Pradana A.C., Adiwijaya K., Aditsania A. Implementing binary particle swarm optimization and C4.5 decision tree for cancer detection based on microarray data classification // Journal of Physics Conference Series. – 2014. – Vol. 1192. – DOI: 10.1088/1742-6596/1192/1/012014. (Дата обращения: 31.05.2019)
14. Gholamian M., Meybodi M.R. Enhanced comprehensive learning cooperative particle swarm optimization with fuzzy inertia weight (ECLCFPSO-IW) // Proceedings of 2015 AI & Robotics (IRANOPEN). – 2015. – DOI: 10.1109/RIOS.2015.7270730. (Дата обращения: 31.05.2019)
15. Umapathy P., Venkatashehaiah C., Senthil Arumugam M. Particle Swarm Optimization with Various Inertia Weight Variants for Optimal Power Flow Solution // Discrete Dynamics in Nature and Society. – 2010. – <http://dx.doi.org/10.1155/2010/462145>. (Дата обращения: 31.05.2019)
16. <https://www.cs.waikato.ac.nz/ml/weka/index.html>. (Дата обращения: 03.04.2019)

РАЗРАБОТКА ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ СИСТЕМЫ РАЙОНОВ ПАДЕНИЙ КОСМОДРОМА БАЙКОНУР НА ОСНОВЕ ГИС-ТЕХНОЛОГИЙ ДЛЯ ВЫБОРА ОПТИМАЛЬНЫХ КООРДИНАТ ТОЧЕК ПРИЗЕМЛЕНИЯ ПРИ УПРАВЛЯЕМОМ СПУСКЕ ОТРАБОТАВШИХ СТУПЕНЕЙ РАКЕТ-НОСИТЕЛЕЙ

**Суйменбаев Б.Т.¹, Трушляков В.И.², Ермолдина Г.Т.¹,
Суйменбаева Ж.Б.¹, Бапышев А.М.¹**
e-mail: gulerm@mail.ru

¹*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

²*Омский государственный технический университет, Россия*

***Аннотация.** Рассмотрены научно-методические подходы к минимизации техногенного воздействия пусков перспективных ракет-носителей с жидкостными ракетными двигателями в районах падения космодрома Байконур при управляемом спуске отработавшей ступени после отделения от ракет-носителя в выделенный участок падения с точностью, не превышающей размеры выделенного оптимального участка. В качестве базового направления снижения техногенного воздействия отработавшего ускорителя первой ступени ракет-носителя в районах падения принято определение оптимальных участков для падения отработавших ступеней в выделенных зонах района падения, обладающей наибольшей устойчивостью и минимальной стоимостью работ на восстановление почвогрунтов участка падения отработавших ступеней в*

исходное состояние на основе ГИС-технологий. Для решения задач определения оптимальных зон, обладающих наибольшими параметрами устойчивости, предлагается создание дополнительной информационно-аналитической системы района падения. Выделены основные критериальные оценочные параметры, определяющие устойчивость экосистем района падения на основе ГИС-технологий.

***Ключевые слова:** район падения, отработавшая ступень, пожаробезопасность, критерий устойчивости, управляемый спуск, информационно-аналитическая система.*

1 Введение

Экологическая безопасность эксплуатации районов падений (РП) отработавших ступеней (ОС) – одна из основных стратегических задач, конечной целью которой является охрана здоровья населения, сохранение биоразнообразия, предотвращение загрязнения, устойчивое функционирование экологических систем, воспроизводство и рациональное использование природных ресурсов.

Влияние запусков ракет-носителей (РН) как один из видов техногенного воздействия на экосистемы районов падений обладает рядом специфических особенностей, делающих ее принципиально отличной от других видов хозяйственной деятельности. Это обусловлено нерегулярной периодичностью запусков, трансграничностью загрязнения, а также многовекторным воздействием практически на все компоненты геосферы - поверхность Земли, приземный слой атмосферы, озоновый слой, ионосферу, околоземное космическое пространство.

Многолетний мониторинг экологических последствий запусков РН в РП свидетельствует о негативной техногенной нагрузке на экосистемы, их уязвимости и долгосрочном восстановлении. В процессе постоянной эксплуатации РП ОС сформировались природные комплексы с почвенными контурами, загрязненными КРТ и продуктами их трансформации. Даже в старых местах падения отделяющихся частей ракет-носителей по истечению 10-15 лет обнаруживаются компоненты ракетного топлива (КРТ) в концентрациях, многократно превышающих ПДК.

Отработавшие первые ступени отделяются на высотах 60-90 км. Скорость движения ступени при вхождении в плотные слои атмосферы недостаточна для её разрушения, которое может произойти в результате аэродинамических перегрузок или вследствие взрыва при перегреве баков остатками топлива. В случае падения первых ступеней, содержащих остатки КРТ, при ударе об грунт возможен взрыв, в результате чего фрагменты ступени разлетаются в стороны от места падения, а продукты сгорания и не прореагировавшие остатки топлива попадают в атмосферу и почву [1]. КРТ, оставшиеся в незначительных объемах в элементах конструкции, проливаются на почву и испаряются в атмосферу. В некоторых случаях может происходить воспламенение проливов. Таким образом, при падении первых ступеней нарушается почвенно-растительный покров, происходит загрязнение приземного слоя атмосферы, почвы, растительности остатками КРТ и продуктами их сгорания и трансформации, а также возгорание растительного покрова и засорение территории фрагментами ступени [1].

В настоящее время перспективы развития ракетостроения направлены на спасение, способы возвращения и управляемый спуск отработавших первых ступеней РН [2-7].

В соответствии с [8] предлагается инновационная технологическая схема, направленная на кардинальное сокращение территории РП и снижение

техногенной нагрузки, в основе которой лежит управляемый спуск ОС после отделения от РН, обеспечиваемый за счет отработки невырабатываемых жидких остатков топлива в баках ОС в выделенные участки РП, обладающих наибольшими параметрами устойчивости, иными словами падение практически «сухих» ОС в заданные точки прицеливания. Для решения задачи выбора оптимальных точек прицеливания предлагается создание информационно-аналитической системы (ИАС) района падения (ИАС_{рп}), входящей в состав системы экологического мониторинга космодрома Байконур (СЭМК).

2 Экологическое состояние районов падений ОС

Согласно Договору аренды комплекса «Байконур» на территории Казахстана выделены 22 зоны падения с 46 районами падения отделяющихся частей ракет-носителей с общей площадью 41 364,7 км² в Кызылординской, Карагандинской, Костанайской, Акмолинской, Северо-Казахстанской, Восточно-Казахстанской, Павлодарской областях [1, 9.]. Кроме того, в соответствии с [10] предусматривается дополнительное выделение земельного участка площадью 5,8 тысячи гектаров в Костанайской области в качестве нового района падения отделяющихся частей РН типа "Союз-2" при пусках с космодрома Байконур в северном направлении. В этой связи проблема обеспечения экологической безопасности территорий РП ОС РН приобретает все большую актуальность благодаря возрастающим требованиям национальных законодательств об охране окружающей среды и претензиями общества.

Основными факторами техногенной нагрузки в РП являются [11]:

- механическое загрязнение ОС и их фрагментами (разрушение конструкции и последующий разброс частей);
- химическое загрязнение проливами невырабатываемых остатков жидких компонентов ракетного топлива (КРТ и продуктами его трансформации);
- пирогенно-термическое воздействие (пожары, взрывы).

Механическое воздействие является следствием падения ОС РН и связано с образованием воронок, разбросом металлическими обломками и фрагментами ОС, повреждением почвенно-растительного покрова. В зависимости от массы ОС и количества КРТ, этот процесс может сопровождаться повреждением гумусового слоя почв и возгоранием [1, 12], нарушением проективного покрытия фитоценозов и изменением флористического состава.

Химическое загрязнение жидкими остатками КРТ, приводящие к высокой вероятности пожаров, связано с высоким содержанием гарантийного невыработанного топлива, является наиболее агрессивным фактором воздействия. Величины жидких остатков КРТ для различных типов РН могут достигать до 3% от начальной заправки КРТ и более.

Загрязнение почвы КРТ приводит к значительным физико-химическим превращениям, выражающимся в изменении микроэлементного состава почвы, ее воздушного и окислительно-восстановительного режимов. Грунтовые и поверхностные воды являются преемниками КРТ и продуктов его распада и трансформации. При попадании КРТ в воду, происходит его окисление кислородом, содержащимся в воде, а также активное разложение под влиянием солнечного излучения и в присутствии химически активных примесей.

Химическое загрязнение почвенных профилей и грунтовых вод зависит от концентрации загрязнителей на поверхности, и чем больше водная нагрузка при детоксикации загрязненных почв, тем глубже проникают в почву токсические соединения - продукты химической трансформации горючего. Так, при проливе на

подзолистую почву через 2 месяца после обезвреживания загрязненной территории НДМГ и продукты его окисления могут проникать на глубину 5-70 см. Высокая фильтрационная способность песка способствует его более глубокому проникновению, глинистые почвы препятствуют глубинной миграции горючих веществ. На вертикальную миграцию углеводородного горючего сильно влияют сорбционные, физико-химические, водно-физические, фильтрационные свойства почвенных пород, минералогический состав, интенсивность процессов разбавления грунтовыми водами, количество углеводородов, попавших в почву. Основное количество нефтепродуктов сосредотачивается в верхнем горизонте; ниже их содержание уменьшается на порядок. При просачивании углеводородного горючего в нижнюю часть профиля образуется достаточно устойчивый ареал загрязнения, оказывающий неблагоприятное длительное воздействие на грунтовые и поверхностные воды [1].

Аккумуляция КРТ растениями зависит от комплекса факторов: семейства растений, геохимических условий мест их произрастания, близости источников поступления техногенных загрязнителей, количества поступившего на поверхность почвы и растений КРТ. Малые концентрации КРТ не оказывают существенного воздействия на цитоморфологическую структуру растений. Действие больших концентраций вызывают у растений морфологические и анатомические изменения [1].

Основным фактором, оказывающим негативное воздействие на окружающую среду, прежде всего на животных, являются пирогенные и термические нарушения растительного и почвенного покрова РП, возникающие при падении ОС. Пожары, инициированные падением ОС РН, уничтожают кормовую базу и в связи с этим способствуют исчезновению растительных млекопитающих, насекомоядных пресмыкающихся и зерноядно-насекомоядных птиц, а также уничтожению энтомофауны и беспозвоночных животных. Практика эксплуатации РН показала, что возгорания и пожары, повреждения растительного покрова из-за падения ОС происходят преимущественно в РП отработавших ускорителей первых ОС с кислородно-керосиновыми маршевыми жидкостными ракетными двигателями (ЖРД) [13].

Наиболее вероятными причинами возникновения пожаров являются жидкие остатки КРТ в баках ОС, которые воспламеняются при разрушении топливных отсеков ОС в момент удара о поверхность грунтов в РП. Так, в результате падения ОС РН «Союз-2.1а» в 2017 году вспыхнул пожар, длившийся 12 часов, погибли люди. Техногенный характер возгорания усугубился условиями сложных погодных условий (жаркой погодой, ветром), в силу чего площадь распространения степного пожара достигла более десяти тысяч гектаров [14].

Вероятность возникновения пожаров зависит от высотного поясного расположения участка территории благодаря продолжительности периода со снежным покровом, характера подстилающих субстратов, угла наклона и экспозиции, которые во многом определяют перераспределение осадков на территории; характера собственно растительности (наличие или отсутствие ветоши, ее высота, густота, возможность «усыхания») [6].

2.1 Параметры оценки

Анализ характера падения отделяющихся частей ракеты-носителя в отведенные районы падения показал, что техногенное воздействие ракетно-космической деятельности в районах падений носит локальный характер, различается для отдельных участков по интенсивности и плотности территориального

распределения, а приземление баков первых ступеней приводит к развалу конструкции, разливу остатков КРТ, возникновению пожаров и, как следствие, требует больших затрат на полную ликвидацию техногенных последствий. Следовательно, проектирование заданных точек прицеливания с устойчивыми характеристиками экосистемы должно базироваться на следующих положениях:

1) ограничение зоны рассеивания и характера разброса фрагментов отработавших ускорителей I ступеней современных РН расчетным эллипсоидом для уточнения области техногенного воздействия;

2) применение ГИС-технологий и использование картографического материала для создания массива данных $IAC_{РП}$;

3) определение оценочных геолого-геоморфологических, гидрографических, показателей, биопродуктивности;

4) определение сезонности и климатических условий;

7) описание хозяйственной деятельности в РП ОС и на сопредельных территориях;

8) определение экологического (экономического) ущерба;

9) определение совокупности оценочных показателей и их пороговых значений, фиксирующих принадлежность зон к приемлемым для прогнозируемых точек прицеливания, что позволит осуществить отбор искомых зон.

Для решения проблемы обеспечения экологической безопасности территории РП при управляемом спуске ОС в заданные точки прицеливания, необходима реализация следующих основных направлений исследования:

– определение оптимальных критериев Q оценки устойчивости и уязвимости к основным факторам техногенного воздействия пусков РН с маршевыми ЖРД;

– выделение оптимальных операционно-территориальных единиц (ОТЕ) (зон) для падения ОС в выделенные зоны S_{opt}^{ws} , обладающие наибольшей устойчивостью к техногенному воздействию пусков РН с маршевыми ЖРД, на основе выявленных критериев;

– определение экономической оценки ОТЕ РП S_{opt}^{ws} и стоимости работ на восстановление экосистем $\min C_{\Sigma}^{S_{opt}^{ws}}$ участка падения в S_{opt}^{ws} ;

Как известно, уязвимость и устойчивость биогеоценоза в РП в результате действия негативных факторов техногенного характера определяется состоянием основных ее компонентов – экотопа (почва, климат, вода) и биотопа (растительности).

Следовательно, для реализации первого направления рассматриваются факторы, определяющие устойчивость РП к техногенному воздействию: а) пожаробезопасность растительности и сопутствующих метеорологических условий; б) состояние почвенного покрова, в том числе угнетение плодородия и деградации почвенного покрова; в) состояние растительности.

2.1.1 Критерии пожаробезопасности

Пожарную опасность определяют такие пожарно-технические характеристики, как горючесть; воспламеняемость; скорость распространения пламени по поверхности; дымообразующая способность; токсичность [15-16].

Устойчивость экосистем к пирогенному воздействию определяется, прежде всего, их способностью к воспламенению и горючестью. Образование горючей среды обусловлено наличием в нем достаточного количества горючего вещества - растительной биомассы, обладающей различными показателями горения [17].

Опасность возникновения пожара определяется комплексом взаимосвязанных метеоэлементов (осадки, влажность воздуха, его температура и др.).

Пожароопасным считается период с момента схода снежного покрова до наступления устойчивой дождливой осенней погоды или образования снежного покрова. Опасность возникновения пожаров увеличивается с северо-востока данного района на юго-запад вслед за увеличением засушливости климата. В течение пожароопасного сезона выделяют периоды пожарных максимумов, когда число пожаров превышает их среднее число.

Следует отметить, что температура воздуха определяется на уровне 2 м над поверхностью земли. В нижнем, приповерхностном слое воздуха температура может достигать существенно больших значений. Нагрев поверхности почвы в дневные часы велик - средний максимум температуры поверхности колеблется от 48 до 59⁰ С в июле, абсолютный максимум достигает 76⁰ С.

2.1.2 Критерии оценки почвенного покрова

Почвенный покров РП включает в свой состав десятки почвенных разновидностей, обладающих различными физико-химическими свойствами. Поэтому в целях упрощения работы по сбору информации и ее обработке используются известные справочные данные с применением почвенной карты данного региона.

Для оценки и градации выделения каждого участка целесообразным представляется определение бонитета почв [18]. Бонитировка почв района падения должна проводиться согласно стандартных методик бонитировки почв. Балл бонитета каждой почвенной разновидности рассчитывается по процентному содержанию гумуса в полуметровом слое, а на все остальные свойства вводятся поправочные коэффициенты. Полученная таким образом бонитировочная шкала накладывается на почвенную карту.

В качестве показателей сопротивляемости поверхностного слоя и трудоемкости ее разрушения для определения допустимой нагрузки к механическому воздействию тяжелых фрагментов, преимущественно двигательных установок и топливно-насосных агрегатов, способствующих образованию воронок разной размерности - диаметром 1-20 м, глубиной 0,3-1,5 м [1] используется учет механических свойств пород по шкале крепости [19].

2.1.3 Критерии оценки растительного покрова

Потенциал растительности в самоочищении экосистем, участвующей в процессах перераспределения и накопления КРТ, зависит от комплекса таких факторов, как разновидность растительного сообщества, геохимические условия мест их произрастания, близость к источникам химического загрязнения КРТ, количество загрязняющих веществ, поступивших на поверхность растительного покрова, флористический состав и структура [20].

Ответная реакция растительности на техногенное воздействие запусков РН выражается в оценке устойчивости растительных сообществ [1, 21]. Потенциальная устойчивость растительных сообществ определяется эколого-биологическими свойствами самих растений [1, 20]. Восстановление некоторых сообществ происходит относительно быстро. В то же время, отдельные сообщества характеризуются низким ежегодным приростом фитомассы, для их самовосстановления требуется более длительный срок. Отличительной чертой всех

сообществ является их слабая устойчивость к пожарам и вследствие этого длительный сукцессионный период восстановления.

Для оценки устойчивости фитоценозов к специфическому техногенному воздействию предлагается учет следующих параметров: а) показатели густоты (структура) растительного сообщества; б) проективное покрытие (соотношение площади в ландшафте); в) фитомасса.

2.2 Оценка экономических затрат на восстановление экосистем

Методология оценки экологического ущерба базируется на последовательной реализации двух основных процедур: определение отрицательных экологических изменений и установление его экономического эквивалента, т.е. стоимости экологического ущерба [11].

Согласно методике определения платы, за загрязнение окружающей природной среды (ОПС) экологический ущерб оценивается согласно базовым нормативным платам за выбросы и сбросы загрязняющих веществ. Размер оплаты за загрязнение ОПС в РП для каждого случая штатного падения ОС определяется конкретными условиями воздействия: падение с разрушением или без разрушения, с проливом или без пролива топлива, объемы пролива на грунт, сброс в водные объекты, выброс в атмосферу и т.д.

Суммарные затраты на экологическое обеспечение пусков и возмещение экологического ущерба включают расходы: плату за загрязнение ОПС, проведение экологического мониторинга, очистку территории, компенсацию снижения продуктивности сельскохозяйственного производства и траты по линии здравоохранения [11].

В ценах 1996 года [11] плата за загрязнение ОПС составила:

- для керосина Т-1 (отходы IV класса токсичности) - 70000 руб./т

- для несимметричного диметилгидразина (отходы I класса токсичности) - 490000 руб./т.

В случае очистки территории, включающей сбор, детоксикацию и захоронение остатков КРТ и фрагментов ОС, нейтрализацию проливов, расходы на ликвидацию последствий существенно увеличиваются.

За все воздействия предусмотрены достаточно большие выплаты природоохранительным органам. К примеру, в ценах 1995г. затраты на проведение контроля в РП составляли ~107 млн.руб. на 1 пуск [11]. В случае попадания фрагментов ОС в населенный пункт последствия всех этих воздействий носят более критичный характер, требуют более оперативных мер по ликвидации их последствий и, возможно, эвакуации населения и проведения ему различных выплат за материальный и моральный ущерб.

Эколого-экономические затраты на РП ОС для РН типа «Протон» могут составить от ~3% до ~35% от стоимости пуска (исходя из стоимости пуска, оцениваемой в ~ 300 млн.руб. в ценах до 17.08.98), причем эти затраты существенно зависят от возможного варианта платы за загрязнение ОПС и природопользование [11, 22]. Для РН типа «Зенит» и «Союз», использующие в качестве топлива керосин Т-1, затраты могут составить от 2 до 21 млн.долларов за один пуск в ценах 2017-2018 гг. [23].

Таким образом, оценка экономических затрат в РП на проведение мероприятий по минимизации техногенных последствий воздействия ОС является обязательным условием и требует учета при выделении ОТЕ.

2.3 Анализ существующей ИАС космодрома

Одним из ключевых моментов в обеспечении экологической безопасности ракетно-космической деятельности является организация системы экологического мониторинга космодрома – СЭМК [11, 24].

Информационное обеспечение СЭМК представлено разработанными базами экологических данных, являющихся составной частью данной системы, и реализуется посредством геоинформационных систем, связанных с моделированием и визуализацией географического пространства и решением пространственных задач. Информационно-аналитическая система (ИАС) СЭМК ориентирована на формирование оперативной информационно-аналитической оценки, возникающих экстремальных ситуаций (аварии, проливы и т.д.) и принятие оперативных и досрочных решений по локализации и ликвидации (компенсации) экологического ущерба от деятельности космодрома Байконур. ИАС-СЭМК состоит из следующих подсистем: подготовка исходных данных, архивации и документирования; планирование, учет и контроль, анализ и регулирование; математическое и геоинформационное моделирование; обработка информации, анализ и принятие решений; нормативно-правовое обеспечение.

Существующая ИАС направлена на мониторинг, оценку состояния территорий, сопровождение полетов и принятие управленческих решений по устранению последствий.

Возможности ИАС СЭМК ориентированы в основном на мониторинг и анализ химического загрязнения территорий РП. Основными преимуществами действующей ИАС являются: картографическое обеспечение, наполняемая база данных, мониторинг загрязнения во временном разрезе, многостороннее выявление факторов воздействия, оперирование ГИС-инструментарием, связанным с моделированием и визуализацией географического пространства и решением пространственных задач. В качестве наглядной визуализации подсистемы ИАС-СЭМК выступает территориальное расположение обследованных ключевых участков и создание векторных слоев, которые имеют в структуре поля, заполненные информацией: номер РП, координаты места и даты падения, координаты точек отбора проб, максимальные значения ПДК, превышение нормативных показателей, используемая методология физико-химической диагностики и др. Основными параметрами устойчивости в разработанной системе критериев потенциальной устойчивости к техногенному воздействию являются показатели химического загрязнения, физико-химической трансформации КРТ, уменьшение биоразнообразия и состояние растительности на организменном и экосистемном уровнях [6].

Предлагаемая к реализации модернизированная ИАС включает мониторинг пирогенно-термического и механического воздействия и нацелена на информационное взаимодействие БД с применяемыми технологическими, схемными и проектно-конструкторскими решениями, направленными на исключение остаточного топлива, снижение экологической нагрузки и химического загрязнения компонентами ракетного топлива экосистем и соответствующее уменьшение районов падения, и сбор данных об экономической стоимости и проводимых мероприятиях по нейтрализации разлитого топлива и восстановлении нарушенных земель после приема ОС РН.

2.3.1 Обоснование создания ИАС_{рп}

Процесс выбора оптимальных участков для падения ОС в выделенные зоны РП предлагается поводить на основе существующей системы СЭМК с введением в

её состав разрабатываемой ИАС_{рп}, которая должна анализировать данные о состоянии почвенно-растительного покрова, рельефе местности, наличии водоисточников, метеопараметрах; классифицировать однородные группы объектов биогеоценоза; уточнять область воздействия штатных падений ОС.

ИАС_{рп} должна выполнять следующие функции:

- учет информации о литогенной основе, рельефе, типах почвогрунтов (подсистема «Почвогрунты»);
- подготовку информации о состоянии растительного покрова на основе архивных картографических данных и космоснимков (подсистема «Растительность»);
- подготовку информации о сезонности и анализ климатических данных (подсистема «Метеофакторы»);
- анализ справочных данных о пожарной безопасности по показателям горючести и воспламенения (подсистема «Пожаробезопасность»);
- учет и экономическая оценка мероприятий по ликвидации последствий падений ОС в РП;
- классификацию однородных групп объектов и анализ результатов зонирования РП.

3. Постановка задачи создания ИАС района падения (ИАС_{рп})

На основании проведенного анализа постановку задачи создания ИАС_{рп} ОС можно сформулировать следующим образом:

1. Обоснование выбора мест приземления ОС, что позволит ограничить саму область и наблюдение за воздействием ракетно-космической техники на окружающую среду РП, а также минимизировать расходы на ликвидацию последствий техногенного воздействия при каждом пуске РН.

2. Определение устойчивых к техногенным нагрузкам точек прицеливания ОС в отчуждаемых РП целесообразно осуществлять в рамках СЭМК, обеспечивающей локальный мониторинг объектов повышенного риска (объекты космодромов; районы падения; территории, прилегающие к районам падения, территории вдоль трасс пусков). Выбор прогнозируемых точек падения ОС РН полагается осуществлять на основе районирования (зонирования) данного РП.

3. Для наиболее эффективной организации экологического мониторинга и оценки внедрения технологий и схемных и проектно-конструкторских решений по управляемому спуску первых ступеней РН предлагается рассматривать информационное объединение основных систем - СЭМК, системы принятия проектно-конструкторских решений (СПКР) разработчиком РН и отработавшую ступень [25].

4. В рамках создаваемой ИАС_{рп} должно осуществляться согласование и координация деятельности организаций и ведомств, участвующих в экологическом мониторинге территорий районов падений ОС, а также разработчиками РН. Например, предлагаемая система ИАС_{рп} в части экологического мониторинга ориентирована на использование разнородной информации о состоянии окружающей природной среды (состояние почв, флоры, фауны, рек и атмосферы); а ИАС_{РН} направлено на внедрение технологий и проектно-конструкторских решений и обеспечение тактико-технических характеристик РН.

Основной целью создания ИАС_{рп} является прогнозирование точек прицеливания, сравнение прогнозируемых координат точек падения с координатами разрешенной зоны падения, оценка величины остатков топлива и рекомендации по корректировке точек падения ОС РН.

4 Основные принципы построения ИАС-М

Для выбора единиц районирования по территориям падения ОС РН и учета всех экологических факторов, влияющих на окружающую природную среду, представляется целесообразным проведение экологического зонирования района эксплуатации РКТ. Данная задача относится к задаче типологической классификации многомерных данных [26], результатом которой должна стать «мозаика» минимальных биогеографических областей (геомеров) с однородными компонентами окружающей среды. Среди такой мозаики может быть выбран некоторый индикаторный геомер [27], определена степень его устойчивости к техногенному воздействию, дана экономическая и экологическая оценка, и выработаны рекомендации к приемлемости и применимости его использования в дальнейшем в качестве контрольной точки.

Объектом классификации, как правило, являются *ОТЕ* — операционно-территориальные единицы [28]. В качестве *ОТЕ* могут выступать ячейки регулярной или нерегулярной сетки, наложенной на исследуемую территорию, ячейки раstra [29-30].

Классификация *ОТЕ* (классов *ОТЕ*) заключается в получении определенного количества предварительно отобранных групп. *ОТЕ* могут быть однородными либо отличаться по ряду признаков. В пространственном отношении *ОТЕ* фиксируется в пространстве, т.е. имеют свой набор географических координат. Типологическая классификация *ОТЕ* дает нам матрицу «объект-признак», для которой рассчитывается критерий оценки и устойчивости. Целью классификации является получение набора групп объектов (классов объектов).

Основными критериями оценки территории и пригодности земельных участков для конкретных точек прицеливания являются выраженные в количественных и качественных показателях свойства и признаки конкретной почвы, геоморфологические и климатические условия местности, эталонные и расчетные показатели урожайности многолетних трав, а также затраты на восстановление и детоксикацию земель. Поэтому земельные участки необходимо оценивать не по отдельным признакам, а по комплексу условий на основе многофакторного анализа.

Для решения поставленной задачи рассматриваются следующие основные компоненты — характеристика почв, растительности, наличие поверхностных и грунтовых вод, климатические условия, а также хозяйственная деятельность человека. Почвенный покров является основным фактором, и тем фокусом, в котором концентрируются и отражаются климатические, геоморфологические и другие факторы. В качестве критериев оценки почв приняты природные свойства, которые определяют условия роста и развития одного из компонентов экосистемы — растительности, как наиболее чувствительного индикатора экологического состояния окружающей среды.

4.1 Структура ИАС_{РН}

Структура ИАС в целом является многопользовательской развивающей средой, идентифицирующей объекты, территории влияния объекта, системная оболочка которой состоит из геоинформационной системы открытого типа, сопровождающая пакет задач, прием и обработку данных со всех источников, а также имеющий интерфейс для работы с внешними пользователями.

Основным принципом построения ИАС является информационное единение разнородной информации, составляющую базу данных. Структура *БД* определяет перечень таблиц базы данных (*ТБД*), межтабличные связи, состав и характеристики

полей отдельных таблиц. БД данного ИАС является реляционной, то есть состоящей из отдельных двумерных таблиц, связанных отношениями.

4.2 Реализация М-ИАС_{РП}

Исполнение М-ИАС_{РП} стандартно, проводится с использованием геоинформационных программных продуктов. Технологическая схема включает сбор данных, их первичной обработки и архивации, формирования картографического материала, и решения прикладных задач мониторинга. Основными ее компонентами являются картографическое отображение, управление базами данных и пространственный анализ данных.

Технологическая схема выглядит следующим образом:

1) Создание цифровой ситуационной (обзорной) расчетной карты района падения на основе топографического, тематического материала, отражающего состояние исследуемой территории с точки зрения различных характеристик окружающей среды; статистических данных и космоснимков для получения наиболее достоверного результата с отображением различных факторов окружающей среды [31].

2) Создание тематических слоев: «рельеф», «дороги», «гидрография», «линии электропередач», «землепользователи», «населенные пункты», «точки падения ракет-носителей»,

3) Оцифровка почвенных, геоботанических и климатических карт района. Создание схемы районирования должно основываться на бонитировке территории, т.е. ценности и экономической оценки земель. Для этого выбранные показатели переводятся в оценочные категории по определенной шкале, на основании которой выводится общая оценка для каждого района.

3) Создание банка данных по экологическим характеристикам штатных запусков ракет-носителей. Данные формируются посредством измерения параметров отдельных компонентов окружающей среды (воздуха, почв, поверхностных и подземных вод, растительного и животного мира, здоровья населения, проживающего в прилегающих местах) с использованием соответствующих методик, инструментальных средств, и др.

4) Оперирование данными для отображения и визуализации объектов из набора данных, моделирование прогнозных ситуаций.

Разрабатываемая ИАС_{РП} в предлагаемом контексте с применением данных по характеристикам растительности и наложении метеопараметров дает оценку состояния окружающей природной среды с учетом фактора пожаробезопасности территории и климатических особенностей.

Основная задача проектирования заданных точек прицеливания относится к типологической классификации многомерных данных [26] и может быть представлена в следующем виде:

Обозначим систему OTE , которую нужно получить в результате экологического зонирования, символом S_n , где n – множество OTE . Каждая OTE_i , будет получена при разбиении на участки с различными ландшафтными характеристиками наложением нерегулярной сетки с указанием географических координат в цифровом формате. Система OTE представляет собой «мозаику» областей с одинаковым составом экосистемы, для которых определены показатели устойчивости.

В качестве набора характеристик экосистемы выступают:

- земельно-оценочная информация, выводимая на основе кадастровой оценки (по плодородию почв и технологическим свойствам) Качество почв оценивается

методом бонитировки по основным свойствам, влияющим на величину плодородия;

- биопродуктивность, видовой состав растительности;

- геолого-геоморфологическая характеристика (холмистость и наличие локальных понижений в равнинном рельефе, мягкость и сыпучесть почвогрунтов предполагают преимущество выноса техногенных загрязнений над процессами их аккумуляции, обеспечивая тем самым устойчивость экосистемы к загрязнению КРТ [32];

- сезонность и климатические условия для мониторинга пожароопасности и миграции токсичных компонентов ракетного топлива.

Произведенный расчет границ ОТЕ дает нам матрицу «объект-признак», для которой рассчитывается устойчивость объекта (показатель резистентности $r_i (i = 1, \dots, N)$) и экономическая оценка затрат на ее восстановление - стоимость показателя - $c_i (i = 1, \dots, N)$. Далее строится временной ряд резистентности ОТЕ $Z(i) = f(b, c)$, $i = 1, \dots, N$. В качестве показателей выступают бонитет типа почв, растительности и т.д. ОТЕ могут интерпретироваться как матрица различия «объект – объект» по набору признаков.

Целью типологической классификации является получение устойчивых групп ОТЕ в М-мерном признаковом пространстве. После проведения типологической классификации дается интерпретация каждому классу-типу, т.е. выделяются диапазоны изменения каждого показателя на ОТЕ этого класса. Путем сравнения находим $Z(m) = \min \{Z(i)\}, i = 1, \dots, N$. Для $Z(m)$ определяют оптимальные координаты точки прицеливания ОЧ $R_{ин}(x, y)$, при которых масса полезного груза выводимого РН на заданную орбиту максимальна $m_{ин}^{max}$.

5 Выделение устойчивых типов операционно-территориальных единиц на примере зоны Ю-24

Алгоритм зонирования районов падения отработавших ступеней на основе выделения ядер заключается в следующем:

1. Уточняют область воздействия отработавшей ступени на окружающую среду. Область воздействия штатных процессов эксплуатации РН на окружающую среду должна быть заведомо известной, т.е. зонировается заданный район падения - эллипсоид рассеивания.

2. На эллипсоид, границы которого заданы расчетным путем, накладываются данные инвентаризации и таксации, физико-географические и природно-ресурсные карты, данные дистанционного аэрокосмического зондирования, формируют множество значений природных и техногенных признаков зонирования.

3. Проводят идентификацию объектов, заключающуюся в установлении взаимно-однозначного соответствия между объектами внешней таблицы атрибутивных признаков анализа и объектами слоя векторной карты

4. На основе пространственного анализа данных выделяют ядра классов и получают матрицу «схожести» для определения признакового пространства.

5. Производится расчет оценочной шкалы по набору природных и техногенных признаков.

6. Находят типичные ОТЕ и классифицируют их по предварительно разработанным достоверным классификационным признакам.

7. Проводят подробную детализированную классификацию с учетом рельефа местности, пожароустойчивости, схем землепользования и климатических условий для исключения миграции компонентов ракетного топлива и пожаров.

8. На основе проведенной классификации выделяют приемлемые для точек прицеливания области, в которых отсутствуют условия для биологического воспроизводства.

9. Визуализируют и документируют результаты классификации в виде карт и текстовых описаний.

В результате проведенного районирования и выделения типичных ОТЕ получается конкретная схема размещения потенциальных точек прицеливания, организации, режима функционирования и воздействия участков районов неблагоприятного антропогенного воздействия (рисунок 1). Такой подход позволяет минимизировать затраты при проведении мониторинговых исследований и дать рекомендации по проектно-конструкторским решениям.

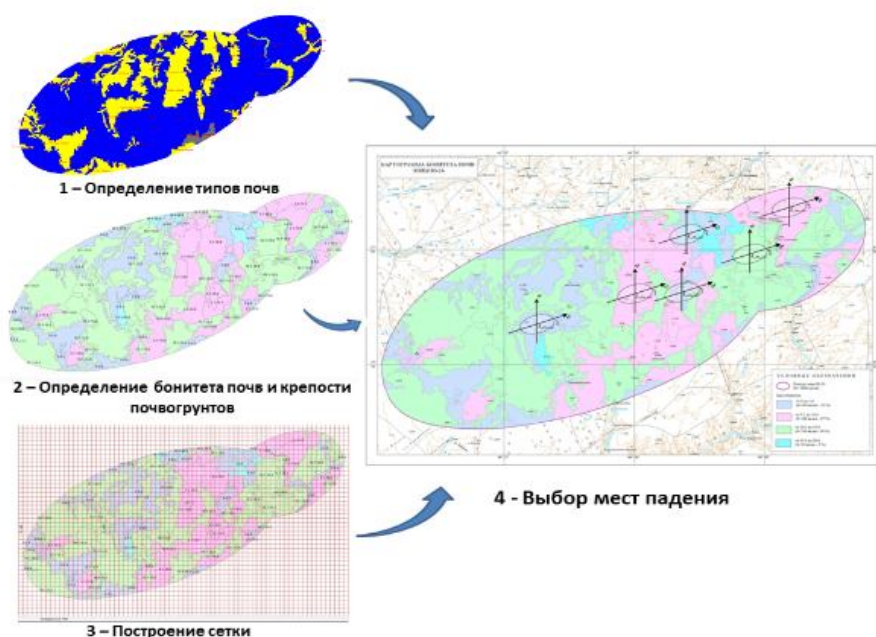


Рис. 1 Алгоритм определения потенциальных точек прицеливания в заданном РП на примере зоны Ю-24

6 Выводы

1. Проведен анализ экологического состояния РП космодрома Байконур, выявлены основные негативные факторы влияния запусков РН. Техногенные факторы воздействия пусков РН проявляются в химическом, механическом и пирогенно-техническом загрязнении на окружающую природную среду РП.

2. Для снижения техногенной нагрузки, повышения эффективности экологического обследования, мониторинга и контроля РП, уменьшения экономических затрат на ликвидацию последствия падений ОС и восстановление земель приняты:

– определение оптимальных участков для падения ОС в выделенных зонах РП, обладающей наибольшей устойчивостью к техногенному воздействию, соответственно, минимальной стоимостью работ на восстановление почвогрунтов участка падения ОС в исходное состояние;

– управляемый спуск ОС после отделения от РН в выделенный участок падения с точностью, не превышающей размеры выделенного оптимального участка.

3. Для решения задач, связанных с выбором мест приземления ОС предлагается создание дополнительной ИАС_{рп}, входящей в состав системы экологического менеджмента космодрома Байконур.

4. В рамках разработки ИАС_{рп} разработаны критериальные оценки, характеризующие основные экологические показатели исследуемого РП: пожароопасность, почвенный покров, растительность для последующего включения в ИАС_{рп}. Показаны объективная необходимость и основные положения по созданию ИАС_{рп}, являющаяся составляющей системы экологического менеджмента космодрома.

Исследования проведены при поддержке Комитета науки Министерства образования и науки Республики Казахстан в рамках гранта №AP05131162 «Теоретические и экспериментальные исследования инновационной технологии создания РН с улучшенными экологическими характеристиками на примере перспективных РН, запускаемых с космодрома Байконур».

Список литературы

1. Экологическая безопасность деятельности космодрома «Байконур» / Под ред. д.т.н., академика МАНЭБ Ж. Жубатова. – Алматы: Изд. Алматы, 2011. - 555 с.
2. Trushlyakov V., Shatrov Ya., Sujmenbayev B., Baranov D., The designing of launch vehicles with liquid propulsion engines ensuring fire, explosion and environmental safety requirements of worked-off stages// Acta Astronautica - 2017. - Vol. 131. - P. 96-101.
3. Баранов Д.А., Макаров Ю.Н., Трушляков В.И. Шатров Я.Т. Проект создания автономной бортовой системы увода отработавших ступеней ракет-носителей в заданные области/ Космонавтика и ракетостроение, №50 (84). - 2015. - С. 76 – 82.
4. Казаков Р.Р., Мингалиев Э.Р. Анализ способов улучшения экологических характеристик ракет-носителей. // Электронный журнал «Труды МАИ». Выпуск № 79.
5. Арсеньев В.Н., Фадеев А.С., Казаков Р.Р. Обеспечение падения отработавших частей ракеты-носителя в заданные районы при пусках с новых стартовых площадок // Электронный журнал «Труды МАИ». Выпуск № 58.
6. А.Д. Кондратьев, Н.С. Касимов, П.П. Кречетов, Т.В. Королева, О.В. Черницова, А.В. Шарапова. Экологическая безопасность ракетно-космической деятельности. - М.: Издательство «Спутник +», 2015. – 280 с.
7. П.П. Поляков. Управление отделяемыми частями ракет-носителей с целью сокращения районов падения //Лесной вестник, 3/2015. - С. 90-94.
8. В.Т. Suimenbayev, V.I. Trushlyakov, G.T. Yermoldina, Zh.B.Suimenbayeva, A.M. Vapyshev. Business-process development of the information analytical Systems of the Baikonur cosmodrom and launch vehicle design for ecological safety Improving in the impact areas of the worked-off stages. //News of The National Academy of Sciences of The Republic of Kazakhstan. Physico-Mathematical Series. ISSN 1991-346X. - Volume 1. - Number 323 (2019) - 5 – 13 <https://doi.org/10.32014/2019.2518-1726.1>
9. Договор аренды комплекса «Байконур» между Правительством Российской Федерации и Правительством Республики Казахстан (Москва), 10 декабря 1994 г.
10. Постановление Правительства Республики Казахстан от 23 мая 2019 года № 313 «О подписании Соглашения между Правительством Республики Казахстан и Правительством Российской Федерации о сотрудничестве по осуществлению

пусков с космодрома «Байконур» ракет-носителей типа «Союз-2» для запусков космических аппаратов в северном направлении на приполярные орбиты».

11. Шатров Я.Т. Обеспечение экологической безопасности ракетно-космической деятельности. Учебно-методическое пособие. - Королев Моск.обл., 2009г. - 759 стр.

12. Экологический мониторинг ракетно-космической деятельности. Принципы и методы / Под. ред. Н.С. Касимова, О.А. Шпигуна. - М.: Рестарт, 2011. - 472 с.

13. Кондратьев А.Д., Кречетов П.П., Королева Т.В. Обеспечение экологической безопасности при эксплуатации районов падения отделяющихся частей ракет-носителей. -М., изд. «Пеликан», 2006.

14. В Казкосмосе прокомментировали пожар при падении частей ракеты-носителя. - <http://gh-ecology.kz/v-kazkosmose-prokommentirovali-pozhar-pri-padenii-chastej-rakety-nositelya.html> (дата обращения 04.06.2019).

15. ГОСТ 12.1.004-91. Межгосударственный стандарт. Система стандартов безопасности труда. Пожарная безопасность. Общие требования. Введ. 1992-07-01. М.: Стандартиформ, 2006. 68 с.

16. ГОСТ 12.1.044-2018 Система стандартов безопасности труда. Пожаровзрывоопасность веществ и материалов. Номенклатура показателей и методы их определения. Введ. 2018-04-30. М.: Стандартиформ, 2018. 100 с.

17. В.Н. Ильина. Пирогенное воздействие на растительный покров //Самарская Лука: проблемы региональной и глобальной экологии. - 2011. - Т. 20, № 2. - С. 4-30.

18. Блисов Т.М. Земельный кадастр и оценка земель (методическое пособие). Костанай: КГУ им. А. Байтурсынова, 2003. - 42с.

19. ГОСТ 21153.1-75. Породы горные. Метод определения коэффициента крепости по Протодьяконову (с Изменением N1). Введ. 1976-07-01. 3 с.

20. Королева, Т.В.; Черницова, О.В.; Шарапова, А.В.; Кречетов, П.П.; Пузанов, А.В. Почвенно-геохимическая характеристика горно-тундровых ландшафтов районов падения отделяющихся частей ракет-носителей /Т. В. Королева [и др.] // Сибирский экологический журнал. - Новосибирск, 2014. - Т. 21, № 2. - С. 183-191.

21. В.В. Неронов, О.В. Черницова, Т.В. Королева, П.П. Кречетов. Современное состояние растительности космодрома Байконур и оценка ее потенциальной устойчивости к воздействию ракетно-космической деятельности //Аридные экосистемы. - 2012. - Том 18, № 3 (52). - С. 72-85.

22. Правила экономической оценки ущерба от загрязнения окружающей среды, утвержденные Постановлением Правительства Республики Казахстан от 27 июня 2007 года N 535. Доступ из информационно-правовой системы нормативных правовых актов Республики Казахстан «Әділет»

23. Твиттер Генерального директора Государственной корпорации по космической деятельности «Роскосмос» Рогозина Д.О. - https://twitter.com/Rogozin?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Cterm%5E1061549418724425728&ref_url=https%3A%2F%2Fwww.novayagazeta.ru%2Fnews%2F2018%2F11%2F11%2F146685-roskosmos-otsenil-stoimost-puska-novoy-rakety-soyuz-5-v-60-millionov-dollarov (дата обращения 04.06.2019).

24. Экологическая безопасность эксплуатации ракетно-космических комплексов. Алматы: Изд-во «Гигатрэйд», 2009. - 240с.

25. Я.Т. Шатров, Д.А. Баранов, Б.Т. Суйменбаев, В.И. Трушляков. Повышение пожаровзрывобезопасности при эксплуатации отработавших ступеней

ракет-носителей с жидкостными ракетными двигателями. //Пожаровзрывобезопасность. - 2016. - Том 25, № 4. - С. 30-42.

26. В.Ю. Ключников. Автоматизированное экологическое зонирование района эксплуатации ракетно-космического вооружения. //Вопросы электромеханики. - Т. 131, 2012. - С. 29-32.

27. Хорошев А.В. Полимасштабная организация географического ландшафта / Московский гос. ун-т им. М. В. Ломоносова, Географический фак. - Москва: Т-во науч. изд. КМК, 2016. - 416 с.

28. Трофимов В.Т. Принципы и признаки инженерно-геологического районирования // Теоретические основы инженерной геологии. М. - 1985.

29. Социально-экономическая география. Учебное пособие. Под ред. Г.И.Гладкевич и др. Изд-во: Геогр. ф-т МГУ Москва, 2016 г. - 234 с.

30. Тикунов В. С. Геоинформатика. Классификации. Операционно-территориальные единицы. Издательский центр «Академия», Москва, 2005 г. - 480 с.

31. Габдуллина Л.А. Научно-методические аспекты разработки и использования ГИС для геоэкологической оценки территории :на примере Ижевска. Диссерт. на соиск. степени канд.геогр.наук. 2007. /disserCat <http://www.dissercat.com/content/nauchno-metodicheskie-aspekty-razrabotki-i-ispolzovaniya-gis-dlya-geoekologicheskoi-otsenki-#ixzz5djT8s6Kt>. (дата обращения 04.06.2019).

32. Степанова Е.Ю., Бисариева Ш.С., Жубат К.Ж., Перменев Ю.Г. Ранжирование территории района падения первой ступени РН «Зенит» (РП 226) по уровням экологической устойчивости // Экологический вестник Северного Кавказа. – Краснодар, Издательство: Кубанский государственный аграрный университет имени И.Т. Трубилина, 2014. – Том 10, № 4. – С. 35-44.

НОВЫЕ ПОДХОДЫ К ИСПОЛЬЗОВАНИЮ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В УЧЕБНОМ ПРОЦЕССЕ

**Сулейменов^{1,2} И.Э., Бакиров¹ А.С., Матрасулова¹ Д.К.,
Тасбулатова¹ З.С., Гришина³ А.В., Косцова⁴ М.В., Мун⁵ Г.А.**
e-mail: esenych@yandex.ru

¹Алматинский университет энергетики и связи

²Институт информационных и вычислительных технологий МОН РК

³Крымский федеральный университет им. В.И. Вернадского,
Севастопольский экономико-гуманитарный институт (филиал)

⁴Севастопольский государственный университет

⁵Казахский национальный университет им. аль-Фараби

Аннотация: Предложены принципы действия многофункциональной системы искусственного интеллекта, предназначенной для использования в учебном процессе (по таким дисциплинам как «Социально-политическое значение систем искусственного интеллекта», «История и философия науки» и т.д.). Система ориентирована на автоматическую обработку критических рецензий,

написанных обучающимися на псевдонаучные работы, в изобилии представленные в текущей периодике. Система позволяет перевести такую инновационную форму учебной работы как написание обучающимися критических рецензий в режим дистанционного обучения. Дополнительной функцией системы является тестирование обучающихся в целях выявления лиц, обладающих психологическим типом, соответствующему ученому в истинном значении этого слова.

Ключевые слова: индекс Хирша, псевдонаука, критическое мышление, пассионарность, нейронная сеть, искусственный интеллект.

В настоящее время в Казахстане имеет место явный переизбыток высших учебных заведений [1]. На начало 2017/2018 учебного года в республике действовало 127 высших учебных заведений при численности населения около 18 миллионов (или 1 вуз на чуть более 140 тысяч человек). Для сравнения, в странах, являющихся признанными лидерами в области высшего образования, на один вуз приходится 650-700 тысяч человек [2]. Избыток высших учебных заведений в Казахстане отчасти обусловлен объективными причинами, так как массовое сознание продолжает рассматривать высшее образование как социальный лифт.

Однако, массовость высшего образования, когда студентами университетов является более 50% лиц соответствующей возрастной группы [3] объективно приводит, во-первых, к резкому падению качества образовательных услуг, а, во-вторых, к существенному падению уровня университетской науки. Усилия Министерства образования и науки РК, к сожалению, во многих случаях вызывают, в том числе, и обратный эффект, по крайней мере, если говорить о результативности научно-исследовательской деятельности в казахстанских университетах. В частности, внедрение в широкое использование такого наукометрического показателя как индекс Хирша привело к резкому увеличению числа псевдонаучных работ, публикуемых преподавателями университетов ради формальной отчетности, ради членства в диссертационных советах, возможности претендовать на получение бюджетного финансирования и т.д. [4-6].

Повышение качества научных работ, выполняемых преподавателями университетов, в сложившихся условиях является весьма актуальной задачей, так как преимущественная ориентация на формальные показатели не только приводит к нарастанию числа псевдонаучных публикаций [6], но и к использованию соответствующих материалов как основы выпускных работ, магистерских и докторских диссертаций, что де-факто развращает студенчество и служит питательной средой для коррупционных проявлений.

Однако, как показано в [4,5], псевдонаучные работы, в изобилии представленные в текущей литературе в настоящее время, могут рассматриваться и как вполне определенный ресурс, в частности, для развития навыков критического мышления обучающихся. А именно, написание критических рецензий на публикации в открытой печати представляет собой перспективную форму учебной работы, особенно если принять во внимание, что в сложившихся условиях написание рефератов, по традиции остающихся важной частью учебного процесса, уже превратилось в профанацию [4].

Целью данной работы является разработка принципов действия системы искусственного интеллекта, обеспечивающей параллельное решение следующих задач: развитие критического мышления у студентов, противодействие нарастанию вала псевдонаучных публикаций и диссертаций (как магистерских, так и докторских), а также выявление среди обучающихся лиц, обладающих выраженной пассионарностью и другими личными качествами, необходимыми для

осуществления плодотворной научной деятельности в современных условиях. В цели работы входит также первичная отработка предложенного подхода на примере действующей нейронной сети, допускающей последующую модернизацию.

Предлагаемая система основана на использовании искусственной нейронной сети (ИНС), ориентированной на анализ содержания критических рецензий, написанных обучающимися на псевдонаучные работы. При разработке заданий акцент делался на критику именно псевдонаучных работ, в том числе, для большей наглядности постановки задачи, однако, нейронные сети, основанные на этом принципе, могут быть использованы и для анализа критических рецензий на публикации любого типа (т.е. в том числе истинно научные, но содержащие отдельные просчеты или элементы добросовестного заблуждения).

В качестве первичного материала для разработки ИНС использовались критические рецензии, написанные магистрантами Алматинского университета энергетики и связи (АУЭС) в рамках выполнения практических заданий по инновационной дисциплине «Искусственный интеллект как драйвер четвертой технологической революции». Выполнение именно таких заданий предусмотрено рабочей программой по данной дисциплине, разработанной параллельно с учебным пособием [7].

В порядке эксперимента задание выполнялось в дистанционном режиме, что позволило полностью охватить весь поток Института космической инженерии и телекоммуникаций АУЭС.

Анализ текстов рецензий позволил выбрать группы слов, словосочетаний и речевых оборотов, выражающих степень критического осмысления текста рецензируемой публикации. Анализ рецензий также показал, что магистранты отчетливо делятся на несколько групп по степени готовности критически осмысливать публикации в открытой печати, по степени готовности отстаивать свою точку зрения и т.д. Прямой анализ текстов также показал, что существует вполне определенное соответствие между характером используемых слов, словосочетаний и речевых оборотов и категориями, к которым можно отнести обучающихся.

Классификация по степени критического отношения к псевдонаучным текстам представлена в Таблице 1, там же представлены примеры индикативных слов и словосочетаний.

Таблица 1. Классификация индикативных слов и словосочетаний по степени критического отношения обучающихся к материалу псевдонаучной работы

Классификационный признак / Описание признака	Характеристика используемых выражений / Примеры индикативных словосочетаний
Понимание содержания работы при неприятии самой идеи критики /	Стремление автора сосредоточить внимание на положительных сторонах работы. <i>Материал изложен логично, рассмотрены результаты, были проведены исследования, результаты показывают, эти опции позволяют, выводы отражают, выводы обоснованы, актуальность не вызывает сомнений и т.д.</i>

<p>Уход от ответа / Выраженное стремление уйти от острого обсуждения недостатков; понимание важности критики недостатков при стремлении не осуществлять данную критику лично</p>	<p>Высказывания с положительной смысловой окраской доминируют над отрицательными (положительные имеют оттенок сомнения, отрицательные отражают стремление критиковать несущественные недостатки). - <i>Было бы, на мой взгляд, можно было бы, среднюю научную ценность, но авторы сделали, статья соответствует, тема актуальна, актуальность статьи не вызывает сомнений, соответствуют всем требованиям, проблемы в статье решены, статья структурирована, оформлена в соответствии с требованиями, материал изложен логично, статья написана на понятном и доступном языке, можно применять как теоретический материал и т.д.</i> - <i>тема не раскрыта, не достаточно информативно орфографические ошибки, грамматические ошибки, цели не достигнуты, статья требует доработки, вывод несодержательный, не раскрыта полностью, отсутствует новизна, в качестве замечания, не указаны ссылки, требует серьёзной доработки, не рекомендуется к публикации, неточности переводов, отдалённое отношение к тексту, выводов нет, отрицательные стороны, не раскрыта до конца и т.д.</i></p>
<p>Мягкая критика / Готовность к критическим оценкам недостатков в мягкой и сглаженной форме</p>	<p>Превалируют негативные выражения, направленность на критику несущественных недостатков. Позитивные оценки присутствуют тоже, но встречаются менее часто. <i>Не раскрыта полностью, отсутствует новизна, в качестве замечания, не указаны ссылки, не указана литература, использовать статью сложно, полностью не соответствует содержанию, язык статьи не понятен, тема не раскрыта, статья требует доработки, вывод несодержательный, цели не достигнуты, отсутствует новизна, автор не смог, не раскрывают всю научность, расчеты не обоснованы, требует серьёзной доработки, не рекомендуется к публикации, отдалённое отношение к тексту, выводов нет, отрицательные стороны, не раскрыта до конца и т.д.</i></p>
<p>Жесткая критика / Готовность к жесткой критике недостатков, готовность отстаивать свое мнение</p>	<p>Рецензии, содержат ТОЛЬКО негативные выражения: <i>Не рекомендуется, непонятно, неясно, не раскрыта проблема, не соответствует, нелогичный, не раскрывает сути, требованиям не удовлетворяет, публиковать не рекомендую, легкое читиво, странная манера изложения, ряд замечаний, отсутствует логический вывод, отсутствует главное, статья требует самоанализа автора, статья требует полной содержательной переработки и т.д.</i></p>
<p>Непонимание сути рецензируемой работы</p>	<p>Прямой пересказ теста статьи отсутствие индикативных слов и словосочетаний</p>

Выявленное соответствие позволяет использовать хорошо известные методы для построения и обучения ИНС [7,8]. Для обучения нейронной сети использовались рецензии, полученные в ходе описанного выше эксперимента. Параллельно в целях контроля проводилось анкетирование студентов,

обеспечивающее дополнительное подтверждение их способности к выработке критических оценок.

Пример одной из анкет, представлен в Таблице 2. В преамбуле анкеты респондентам (магистрантам) была предоставлена следующая информация, которая дополнительно разъяснялась на практических занятиях.

На сайте Интернет-ресурса «Зеркало недели» (Украина) была опубликована статья «Академическое недостоинство 2018». В ней отражены результаты «антиконкурса», проведенного украинскими коллегами. Данный «антиконкурс» базируется на идее выдающегося физика С. Шарапова, предложившего отличать академически недобропорядочных ученых, давая им своеобразную «черную метку». Тем самым научное сообщество приобретает способность контролировать «собственную территорию». Идея была воплощена в жизнь украинскими активистами инициативы «Дисергейт»; первое такое «награждение» состоялось в 2016 году.

Респондентам было предложено выбрать только один из представленных вариантов ответов, каждый из которых соответствует «черным меткам», присуждаемым украинскими коллегами в перечисленных ниже номинациях.

Таблица 2. Перечень вопросов в анкете, предложенной магистрантам при опросе и абсолютное число магистрантов N , выбравших данный вариант ответа

	Вариант ответа	N
1	«Плаггиатор года» – черную метку высылают коллегам и чиновникам, которые наиболее беззастенчиво использовали чужие результаты	9
2	«Псевдочученый года» – номинация для тех, кто публиковал и отстаивал наиболее вредоносный бред	5
3	«Мурзилка года» – «награждается» журнал, который наиболее рьяно публиковал псевдонаучные работы (чаще всего – за наличные)	7
4	«Скандал года» – за наиболее громкий скандал в научной среде	1
5	«Посіпака года» – присуждается организациям (например, ученым советам) за «значительный вклад» в продуцирование плаггиаторов, фальсификаторов и фабрикаторов	3
6	Казахстан – это не Украина, нам такие «антиконкурсы» уж точно не нужны	8
7	Я не считаю себя вправе выносить суждение по данному вопросу, магистрантам не следует даже думать о том, чтобы критиковать старших	5

Распределение полученных ответов в процентах представлено на рис. 1. Обученная нейронная сеть позволяет классифицировать обучающихся по

признакам, определяющим их готовность к научной работе. Подчеркиваем, что в современных условиях одним из наиболее существенных признаков является готовность преодолевать препятствия (создаваемые как консерватизмом значительной части профессорско-преподавательского состава, так и бюрократической рутинной) и другие черты характера, формирующие в совокупности пассионарный психотип [9].

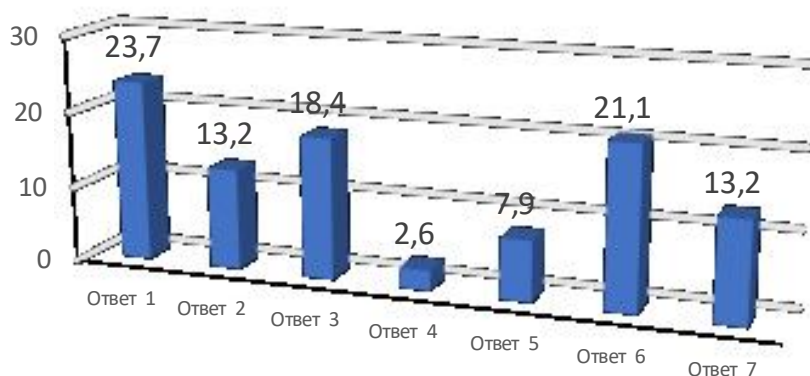


Рис. 1. Процентное распределение ответов по вариантам (Таблица 2)

Напомним, что по Л.Н. Гумилеву [10] под пассионарностью понимается характерологическая доминанта, необоримое внутреннее стремление (осознанное или, чаще, неосознанное) к деятельности, направленной на осуществление какой-либо цели (часто иллюзорной).

Как показывает опыт работы с молодыми научными кадрами, в современных условиях пассионарность зачастую имеет куда более серьезное значение, нежели формальные компетенции (так, существует сколько угодно примеров, когда подающий надежды молодой ученый бросал науку при возникновении тех или иных препятствий), впрочем, значительную роль фактора пассионарности для достижения успеха в науке демонстрирует вся ее история [11].

Для реализации искусственной нейронной сети предложенного типа был использован язык программирования Python, с открытой нейросетевой библиотекой для глубокого обучения Keras. Она представляет собой надстройку над фреймворками DeepLearning4j, TensorFlow и Theano. Данная библиотека нацелена на оперативную работу с сетями глубинного обучения, при этом спроектирована так, чтобы быть компактной, модульной и расширяемой. Она была создана как часть исследовательских усилий проекта ONEIROS (англ. Open-ended Neuro-Electronic Intelligent Robot Operating System) [12]. В качестве бэкенда использовался TensorFlow, представляющий собой открытую программную библиотеку для машинного обучения, разработанную компанией Google для решения задач построения и тренировки нейронной сети с целью автоматического нахождения и классификации образов, достигая качества человеческого восприятия. Применяется как для исследований, так и для разработки собственных продуктов Google [13].

Разрабатываемая система ориентирована на работу с текстовыми выборками (текстами рецензий), поэтому осуществлялось предварительная обработка обучающей выборки для создаваемой глубокой искусственной нейронной сети в удобном виде. Имеющаяся выборка рецензий была разбита на 5 групп по степени

понимания рецензируемого материала в соответствии с классификационными признаками, отраженными в Таблице 1.

Рецензии обучающей выборки были помещены в один csv-файл. Структура csv-файла представляет собой строки, состоящие из двух столбцов, разделенных знаком «собачки». В первом столбце в одну линию записана сама рецензия, во втором указана группа, к которой относится данная рецензия. В дальнейшем средствами библиотеки Keras осуществлялся препроцессинг текста: при помощи функции Tokenizer (модуль tokenize представляет собой лексический сканер для исходного кода на Python, реализованный на Python. Сканер этого модуля возвращает комментарии как токены; слова в выборке заменялись на цифры, с которыми уже может работать искусственная нейронная сеть.

Реализованная сеть представляет собой последовательную рекуррентную искусственную нейронную сеть, состоящую из 3 слоев. Схематически такая нейронная сеть показана на рис. 2.

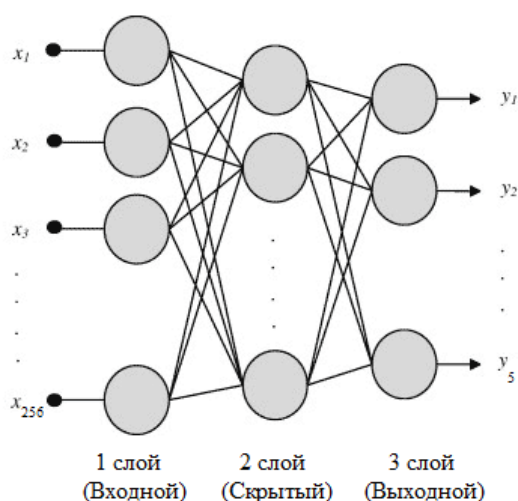


Рис. 2. Схема первичного варианта искусственной нейронной сети

Первым входным слоем является полносвязный слой, состоящий из 256 нейронов, с типом активации ReLU. Данная функция активации имеет название «выпрямитель» (rectifier, по аналогии с однополупериодным выпрямителем в электротехнике). Нейроны с данной функцией активации называются ReLU (rectified linear unit). ReLU имеет следующую формулу $f(x) = \max(0, x)$ и реализует простой пороговый переход в нуле.

Отметим, что вычисление сигмоиды и гиперболического тангенса требует выполнения ресурсоемких операций, таких как возведение в степень, в то время как ReLU может быть реализован с помощью простого порогового преобразования матрицы активаций в нуле. Кроме того, ReLU не подвержен насыщению. Кроме того, применение ReLU существенно повышает скорость сходимости стохастического градиентного спуска (в некоторых случаях до 6 раз) по сравнению с сигмоидой и гиперболическим тангенсом. Считается, что это обусловлено линейным характером и отсутствием насыщения данной функции [14].

Второй слой построенной сети представляет собой слой регуляции с вероятностью отключения нейрона 50%. Выходной слой состоит из 5 нейронов, соответствующих количеству определяемых категорий. Тип активации для

выходного слоя Softmax, общая сумма выходов всех 5 нейронов выходного слоя равна 1, что отвечает процедуре классификации.

Функция SoftMax (функция мягкого максимума) часто используется в нейронных сетях в качестве функции активации при решении задачи классификации. SoftMax задается следующей формулой: $\sigma(z)_i = \frac{e^{z_i}}{\sum_{k=1}^N e^{z_k}}$, где z_i – значение на выходе из i -го нейрона до активации, а N – общее количество нейронов в слое.

В качестве функции ошибки при компиляции модели искусственной нейронной сети была выбрана категориальная кроссэнтропия*, которая хорошо подходит для случая, когда на выходе искусственной нейронной сети имеются вероятности появления класса [15]. Оптимизация модели осуществляется посредством метода адаптивной оценки момента (Adam), в качестве метрики модели выбрана точность.

В обучающей выборке имелось 20 рецензий. Размер мини-выборки для каждой эпохи обучения составил 9 рецензий, количество эпох обучения 15. Также была выбрана возможность автоматического разделения выборки на обучающую и проверочную. Размер проверочной выборки составил 10% от основной.

По итогам 15 эпох обучения точность определения класса рецензии на тестовых данных составила 94,4%, на обучающей выборке 50% (рис.3).

```

Train on 18 samples, validate on 2 samples
Epoch 1/15
18/18 [=====] - 2s 96ms/step - loss: 2.0089 - acc: 0.1111 - val_loss: 1.5148 - val_acc: 0.0000e+00
Epoch 2/15
18/18 [=====] - 0s 445us/step - loss: 1.4680 - acc: 0.3889 - val_loss: 1.8863 - val_acc: 0.0000e+00
Epoch 3/15
18/18 [=====] - 0s 389us/step - loss: 1.1337 - acc: 0.4444 - val_loss: 2.4155 - val_acc: 0.0000e+00
Epoch 4/15
18/18 [=====] - 0s 389us/step - loss: 0.9287 - acc: 0.6667 - val_loss: 2.7458 - val_acc: 0.0000e+00
Epoch 5/15
18/18 [=====] - 0s 445us/step - loss: 0.6838 - acc: 0.8889 - val_loss: 3.0558 - val_acc: 0.5000
Epoch 6/15
18/18 [=====] - 0s 445us/step - loss: 0.6847 - acc: 0.8889 - val_loss: 3.3143 - val_acc: 0.5000
Epoch 7/15
18/18 [=====] - 0s 445us/step - loss: 0.5302 - acc: 0.8889 - val_loss: 3.5213 - val_acc: 0.5000
Epoch 8/15
18/18 [=====] - 0s 445us/step - loss: 0.4458 - acc: 0.9444 - val_loss: 3.6991 - val_acc: 0.5000
Epoch 9/15
18/18 [=====] - 0s 445us/step - loss: 0.3468 - acc: 0.9444 - val_loss: 3.8344 - val_acc: 0.5000
Epoch 10/15
18/18 [=====] - 0s 445us/step - loss: 0.3257 - acc: 0.9444 - val_loss: 3.9366 - val_acc: 0.5000
Epoch 11/15
18/18 [=====] - 0s 389us/step - loss: 0.2584 - acc: 1.0000 - val_loss: 4.0200 - val_acc: 0.5000
Epoch 12/15
18/18 [=====] - 0s 445us/step - loss: 0.2476 - acc: 1.0000 - val_loss: 4.0929 - val_acc: 0.5000
Epoch 13/15
18/18 [=====] - 0s 389us/step - loss: 0.2384 - acc: 1.0000 - val_loss: 4.1628 - val_acc: 0.5000
Epoch 14/15
18/18 [=====] - 0s 389us/step - loss: 0.2074 - acc: 1.0000 - val_loss: 4.2180 - val_acc: 0.5000
Epoch 15/15
18/18 [=====] - 0s 389us/step - loss: 0.2080 - acc: 0.9444 - val_loss: 4.2882 - val_acc: 0.5000

```

Рис. 4. Результаты обучения искусственной нейронной сети

Низкое значение точности определения класса на обучающей выборке вызвано малым размером исходной выборки, однако полученные результаты показывают, что разработанный инструмент является вполне работоспособным и может быть далее доведен до уровня реального практического использования за

* Кросс-энтропия - одна из многих возможных функций потерь (другой популярной является потеря петли SVM). Эти функции потерь обычно записываются как $J(\theta)$ и могут использоваться в градиентном спуске, что является итеративной основой перемещения параметров (или коэффициентов) к оптимальным значениям.

счет увеличения обучающей выборки. Таким образом, существует реальная возможность перевести значительную часть практических занятий по таким дисциплинам как «История и философия науки» [11] в дистанционный режим, параллельно тестируя обучающихся на наличие склонностей к самостоятельной научной работе с целью подбора кадров для выполнения научно-технических проектов и программ, в том числе, междисциплинарного характера.

Полностью аналогичный инструмент может быть использован также и для стимулирования междисциплинарного сотрудничества, ориентированного на специальности, связанные с радиоэлектроникой и инфокоммуникационными технологиями. В этом случае обучающимся предлагается писать критические рецензии на характер использования измерительных методов (в части радиоэлектронного обеспечения, например), а также в части использования инфокоммуникационных методов в смежных (и не только) областях знания.

Критика методик, реально используемых в работах по химии, нанотехнологии, экологии и т.д. позволит обучающимся увидеть точки приложения своих усилий, что в свою очередь, является залогом для плодотворной проектной деятельности в рамках концепции «треугольника знаний».

Литература

1. Шырын М.К. О некоторых проблемах высшего образования в Казахстане // Мат. конф. «Глобальные вызовы и современные тренды развития высшего образования» - 2013 – Алматы, <https://articlekz.com/article/9466>
2. Estimated number of universities worldwide as of July 2018, by country <https://www.statista.com/statistics/918403/number-of-universities-worldwide-by-country/>
3. Национальный доклад о состоянии и развитии системы образования Республики Казахстан (по итогам 2017 года) http://iac.kz/sites/default/files/nacionalnyy_doklad_za_2017_god_s_oblozhkami_dlya_sayta.pdf
4. Мун Г.А. Экология и альтернативная энергетика – поле битвы информационной войны // Корё Ильбо, №49 (7.12.2018), с.12-13, режим доступа <http://www.arirang.ru/archive/kore-ilbo/2018/KI.07.12.2018.pdf>
5. Suleimenov I., Gabrielyan O., Egemberdyeva Z. Kopyshhev E., Tasbolatova Z. IMPLEMENTATION OF EDUCATIONAL INFORMATION TECHNOLOGY TO DEVELOP CRITICAL THINKING SKILLS //Известия научно-технического общества «Кахак». – 2019. – №. 1 (64) – С. 63-71.
6. Мун Г.А., Тасбулатова З.С., Сулейменов И.Э. Псевдонаука как ресурс: нестандартные подходы в образовательных информационных технологиях //Известия Научно-Технического Общества «КАХАК». –2019. – №. 1 (64). – С. 43
7. Калимолдаев М.Н., Мун Г.А., Пак И.Т., Бакиров А.С., Байпакбаева С.Т., Сулейменов И.Э. Искусственный интеллект как драйвер четвертой технологической революции. Учебное пособие для магистрантов // Алматы: Изд-во ТОО «Полиграфкомбинат», 2018г. -313с.
8. Хайкин С. Нейронные сети: полный курс, 2-е издание. – Издательский дом Вильямс, 2008.
9. Савельева О. В., Маслова М. В. Пассионарность как мера активности человека и общества //Философия образования. – 2008. – №. 3. – С. 163-167
10. Гумилев Л. Н. Этносфера: история людей и история природы; Этногенез и биосфера Земли //М.: Эксмо. – 2012. – С. 1056.

11. Сулейменов И. Э., Габриелян О.А., Седлакова З.З., Мун Г.А. История и философия науки. — Алматы, Изд-во КазНУ, 2018, 406 с.
12. Keras: The Python Deep Learning library [Электронный ресурс] / Официальный сайт библиотеки «Keras». — Режим доступа : <https://www.keras.io/> (23.05.2019).
13. Abadi M. et al. Tensorflow: A system for large-scale machine learning //12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16). – 2016. – С. 265-283.
14. Li Y., Yuan Y. Convergence analysis of two-layer neural networks with relu activation //Advances in Neural Information Processing Systems. – 2017. – С. 597-607.
15. Zhang Z., Sabuncu M. Generalized cross entropy loss for training deep neural networks with noisy labels //Advances in Neural Information Processing Systems. – 2018. – С. 8778-8788.

ОБЗОР НЕЙРОСЕТЕВЫХ МЕТОДОВ РАЗРЕШЕНИЯ АНАФОРИЧЕСКИХ СВЯЗЕЙ

**Терехов А.Г., Егембердиева З.М., Джаксылыкова А.Б.,
Бисаринов Б.Ж.**

e-mail: aterekhov1@yandex.ru, egem.zm@mail.ru, aselya17.89@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан
Казахский Национальный Университет имени аль-Фараби , Казахстан*

***Аннотация.** В статье представлены результаты экспериментов по построению общей системы разрешения анафоры для сбалансированных по роду корпуса. Основная идея проведенных экспериментов заключается в использовании преимуществ механизма ядра векторизации глубоких слов по отношению к различным семантическим и синтаксическим особенностям. Мы покажем, что это оказывает положительное влияние на производительность системы по сравнению с традиционным подходом.*

1. Введение

Разрешение Coreference Resolution (CR) является критически важным шагом в конвейере многих высокоуровневых приложений обработки естественного языка (NLP), а именно извлечения фактов, машинного перевода, суммирования текста и т.д. Задача CR начинает свою историю с первых попыток разрешить связь между именами местоимения и именами [1], в настоящее время такая задача называется Anaphora Resolution (AR). Впоследствии на конференции по пониманию сообщений были решены некоторые частичные подзадачи [2]. Традиционно, алгоритмы, в качестве первого шага, генерировали набор упоминаний сущностей. После этого каждая пара местоимений и упоминаний была классифицирована с помощью алгоритма дерева решений. В конце обработки положительно классифицированные пары были объединены в кластер, который представляет собой выделенную цепь между связанными контекстами упоминаний и именными фразами. В работе [3] была представлена модификация порождающей грамматики,

а именно «Теория правительства и связывания» (GB). Значительная часть теории была посвящена синтаксической анафоре, то есть анализу связи между местоимением и другими частями речи.

Следующая эпоха разрешения контрольных точек начинается с конференции «ConLL-2011/2012» и общедоступной задачи моделирования неограниченных базовых элементов, в результате чего появилось новое разнообразие подходов [4]. Отношение - это связь между упоминаниями (существительными), которые относятся к одной и той же сущности. По аналогии с GB существует подход, который использует мультифильтрацию, где каждый уровень фильтров устанавливает или запрещает связь между словами. В работе [5] указывалось, что при выполнении основной задачи по существительному словосочетанию подход к обучению достигает точности, сравнимой с нелинейными подходами, более того, вместо того, чтобы рассматривать задачу AR как локальную задачу оптимизации, авторы предложили новое рассмотрение задачи. как глобальная оптимизация среди всех пар в документе.

Существует большое разнообразие ресурсов и корпусов, помеченных вручную. Тем не менее, большинство из них не фиксируют неоднозначные местоимения в достаточном объеме или разнообразии, чтобы точно указать практическую применимость моделей. Кроме того, существует нехватка сбалансированных корпусов из-за гендерной предвзятости в существующих корпусах и системах, благоприятствующих мужским сущностям. Чтобы решить эту проблему, был представлен новый корпус - сбалансированный пол (GAP) из 8908 неоднозначных пар местоимений и имен, отобранных для обеспечения разнообразного охвата задач, возникающих в реальном тексте. Существует ряд базовых показателей, которые демонстрируют сложность задачи, лучшее достижение - всего 66,9% балла F1 [6].

2. Вложения слов: BERT

Вложения слов - это форма вычислительной семантики распределения для определения значения слова «от компании, которую оно хранит» [7], то есть со словами, с которыми оно встречается. Самая популярная модель недавно была представлена Двухнаправленным кодировщиком от преобразователя (BERT). В отличие от других встраиваний слов [8,9], BERT предназначен для предварительной подготовки глубоких двухнаправленных представлений путем совместной подготовки как левого, так и правого контекста во всех слоях. В результате, упомянутая модель хорошо подходит для широкого круга задач обработки естественного языка, таких как ответы на вопросы и разрешение основных вопросов и другие [10].

Модель BERT имеет несколько общих характеристик с ELMo и OpenAI GPT. Это многослойный двухнаправленный преобразователь, основанный на оригинальной реализации, описанной в работе [11]. В настоящее время реализация преобразователя настолько популярна, что мы опускаем исчерпывающее описание архитектуры модели и отсылаем читателей к Vaswani. [11].

3. Kernel трюк

Kernel трюк - это математический инструмент, который можно применить к любому алгоритму, который состоит из точечного произведения между двумя векторами. Kernel трюк утверждает, что для вычисления точечных произведений векторов в пространстве более высокой размерности можно напрямую использовать функцию ядра, используя векторы более низкой размерности.

Следовательно, каждое скалярное произведение может быть заменено функцией ядра. Таким образом, функция ядра может быть определена как функция, которая принимает в качестве входных векторов исходное пространство и возвращает скалярное произведение векторов в пространстве признаков. Пространство объектов обычно представляет собой пространство более высокого измерения, в котором данные могут быть линейно разделены. Чтобы функция была действительной функцией ядра, она должна соответствовать «теореме Мерсера». Согласно теореме Мерсера, каждая положительно определенная симметрическая функция может рассматриваться как функция ядра.

4. Классификатор данных и дополнительные функции

Задача состоит в том, чтобы определить цель местоимения в текстовом отрывке. Исходный текст взят из статей Википедии. В наборе данных есть метки местоимения и два имени кандидатов, на которые может ссылаться местоимение. Алгоритм должен быть способен решить, относится ли местоимение к имени А, имени В или ни к одному из них. Набор данных также доступны в GitHub Repo.

Существуют следующие столбцы для анализа:

- ID - уникальный идентификатор для примера (соответствует Id в формате выходного файла);
- Text - текст, содержащий неоднозначное местоимение и два имени кандидата (длиной около абзаца);
- Pronoun - целевое местоимение (текст);
- Pronoun - символное смещение местоимения в тексте;
- A - первое имя кандидата (текст);
- A-offset - смещение символа имени А в тексте;
- B - второе имя кандидата (текст);
- B-offset - смещение символа имени В в тексте;
- URL - URL-адрес исходной страницы Википедии для примера;
- Дополнительные функции

$$w_{ab}^{\rho} = \rho \cdot \rho_a \cdot b \quad (1)$$

где ρ_a ρ_b - векторные представления кандидатов А и В; \cdot - поэлементное умножение.

$$w_{pp}^{\rho} = \rho \cdot \rho \cdot \rho \quad (2)$$

где ρ ρ - векторные представления местоимения; \cdot - поэлементное умножение.

$$w = \exp \left(\left\| \rho_{pp}^{\rho} - \rho_{ab}^{\rho} \right\| \right) \quad (3)$$

Мы проверили и доказали, что дополнительные функции, такие как расстояния между кандидатами А и В и местоимением; определитель (его или ее) обычно ссылается на субъект, поэтому были добавлены дополнительные функции синтаксических ролей; Прикладные функции ядра (3) улучшают производительность классификации. Всего 4000 образцов с размером 6149.

5. Модели и эксперименты

Наша нейронная сеть - это сеть с прямой связью, которая состоит из 2 скрытых слоев размером 2048 с выпадением 0,7 и 254 с выпадением 0,2, функция активации использует «relu». Последний слой - softmax - вероятностное распределение классов, см. Рисунок 1.

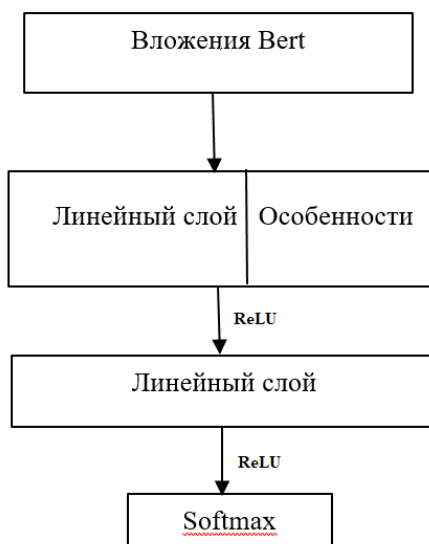


Рис.1 Архитектура классификатора разрешения анафоры

3. Результаты

Задача состоит в том, чтобы определить цель местоимения в текстовом отрывке. Вместе с различными архитектурами векторизации были протестированы различные дополнительные функции. Вы можете увидеть сравнительный анализ в Таблице 1, в первом столбце дано название модели и дополнительные функции для улучшения классификации. Второй столбец показывает ошибку в виде значения кросс-энтропии.

Таблица 1. Особенности модели в сравнении с функцией потери журнала

Модель	Подтверждение Потеря
База BERT	0.63
Большой BERT	0.58
Большой BERT + Особенности	0.52
Большой BERT + Kernel трюк + Особенности	0.45
Большой BERT + [18,20] слои + Особенности	0.42
Большой BERT + все	0.33

Мы благодарим Институт Информационных и Вычислительных Технологий за финансовую поддержку. (Грант AP05132760, Казахстан).

Литература

1. Hobbs, J.R. Resolving pronoun references / J.R. Hobbs // *Lingua*, 1978. – Vol. 44, N. 4. – P.311-338.
2. Seventh Message Understanding Conference: Proceedings of a Conference Held in Fairfax, Virginia, USA, MUC 1998, April 29 - May 1, 1998. ACL 1998
3. Chomsky, N. Lectures in Government and Binding / N. Chomsky. - Studies in generative grammar 9. Dordrecht: Foris, 1981. – 371 p.
4. Pradhan, S., Moschitti, A., Xue, N., Uryupina, O., Zhang, Y. CoNLL-2012 shared task: Modeling multilingual unrestricted coreference in OntoNotes / S. Pradhan, A. Moschitti, N. Xue, O. Uryupina, Y. Zhang // In Proceedings of the Sixteenth Conference on Computational Natural Language Learning (CoNLL 2012), 2012. – P. 1-40.
5. Soon, W.M., Ng, H.Y., Lim, D.C.Y. A machine learning approach to coreference resolution of noun phrases / W.M. Soon, H.Y. Ng, D.C.Y. Lim // *Computational Linguistic (special Issue on Computational Anaphora Resolution)*, 2001. – Vol. 27, N. 4. – P.521-544.
6. Webster, K., Recasens, M., Axelrod, V., Baldrige, J. Mind the GAP: A Balanced Corpus of Gendered Ambiguous Pronouns / K. Webster, M. Recasens, V. Axelrod, J. Baldrige // *Transactions of the Association for Computational Linguistics*, 2018. – Vol. 6. – P. 605-617.
7. Firth, J.R. Applications of General Linguistics / J. R. Firth // *Transactions of the Philological Society*, 1957. – Vol. 56. – P. 1-14. doi:10.1111/j.1467-968X.1957.tb00568.x
8. Peters, M., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., Zettlemoyer, L. Deep contextualized word representations / M. Peters, M. Neumann, M. Lyyer, M. Gardner, C. Clark, K. Lee, L. Zettlemoyer // *Proceedings of NAACL-HTL*, 2018. – P.2227-2237.
9. Radford A, Narasimhan V, Salimans T, Sutskever I. Improving language understanding with unsupervised learning. Technical report, OpenAI. - https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf (01.02.2019).
10. Devlin J, Chang M-W, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. - <https://arxiv.org/pdf/1810.04805.pdf> (10.02.2019).
11. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser L, Polosukhin I. Attention is all you need. - <https://arxiv.org/pdf/1706.03762.pdf> (14.03.2019).

СПУТНИКОВЫЙ МОНИТОРИНГ ВОДНЫХ ОБЪЕКТОВ СИНЬЦЗЯН-УЙГУРСКОГО АВТОНОМНОГО РАЙОНА КНР

Терехов А.Г., Пак А.А.

e-mail: aterekhov1@yandex.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** Проект ГФ МОН РК (2018-2020 гг.) № АР 05134241 посвящен задачам организации и проведения спутникового мониторинга объектов гидротехнической инфраструктуры (каналы, водохранилища), природных водных объектов (реки, озёра) в пределах бассейнов трансграничных рек (р. Иле, р. Кара-Ертис) и зависимых территорий (Джунгарская равнина) Синьцзян-Уйгурского автономного района (СУАР) КНР. Обеспеченность водой территории Казахстана в трансграничных бассейнах рек Иле и Ертис сильно зависит от водопользования, расположенных выше по течению рек, территорий СУАР КНР и сезонной водности года. Состояние крупных водных объектов на сопредельной территории СУАР КНР диагностируют, как параметры водопользования (например, наполненность водохранилищ), так и погодную компоненту (площадь естественных озёр). В данной работе на основе спутникового мониторинга LANDSAT-8 и SENTINEL-2A 2019 года рассмотрена динамика площади водных зеркал озёр Эби-Нур и Манас, а также расчетный запас воды в рабочих частях двух доминирующих водохранилищ китайской части бассейна р. Иле (Капшагайское (р.Текес) и Жарынтайское (р.Каш)). Полученные результаты указывают, что состояние озёр имеет тренд на уменьшение площади, что, по всей видимости, связано с релаксацией гидрологии территорий после исключительно влажного 2016 года. Мониторинг водохранилищ (Капшагайское (р.Текес) и Жарынтайское (р.Каш)), в первую половину 2019 года показывает существенное истощение запасов воды, относительно среднего состояния на это время. Причиной такого режима стало наложение планового водопользования с холодной погодой апреля-мая 2019 года, что привело к маловодности горных рек Восточного Тянь-Шаня.*

Введение

Одной из основных целей проекта ГФ МОН РК (2018-2020 гг.) АР 05134241 «Проведение мониторинга территории Синьцзян-Уйгурского Автономного Района КНР в пределах бассейнов трансграничных рек и зависимых территорий на базе спутниковой информации» является получение оперативной информации, связанной с гидрологией верхней части трансграничного бассейна реки Иле и территории Джунгарской равнины, на глубину доступной спутниковой съёмки (30 лет).

Программа обмена гидрологическими данными между Казахстаном и КНР для трансграничных речных бассейнов определена в соответствующих межгосударственных документах [1]. В программу не входят, ни оперативная информация по естественным объектам, ни данные по состоянию водохранилищ в бассейне р.Иле в СУАР КНР. Недостаток гидрологической информации необходимой для эффективного администрирования в нижних, Казахстанских частях бассейнов трансграничных рек, вынуждает разрабатывать и использовать методики обработки данных дистанционного зондирования, в частности, спутниковую съёмку.

Водные объекты мониторинга первой половины 2019 года, в том числе, включали: естественные водные объекты Джунгарской равнины, озёра Эби-Нур и Манас; и искусственные резервуары китайской части бассейна р. Иле, Капшагайское (р.Текес) и Жарынтайское (р.Каш) водохранилища. Параметрами мониторинга выступали площади водных зеркал озёр, а также расчетные объёмы запасов воды в рабочих частях водохранилищ.

Методы

Спутниковые данные LANDSAT-8 (разрешение 30 м) и SENTINEL-2A (разрешение 20 м) доступны на свободной основе на Интернет порталах NASA\USGS [2]. Спектральные различия в каналах оптического (и ближнего ИК) диапазона между водной поверхностью и подстилающей поверхностью других классов весьма велики. Это значительно упрощает их распознавание и разделение на снимках, свободных от помех облачного или снежно-ледового покрова.

Спутниковые данные LANDSAT-8 и SENTINEL-2A в оптических каналах спектра подвергались кластеризации (классификации) для выделения водных зеркал. Методически использовалась автоматическая классификация (ISODATA), с экспертной коррекцией, в случае помех от облачного покрова или снежно-ледовых образований в холодное время года.

Исходные данные

В качестве исходной спутниковой информации в задаче мониторинга водных зеркал использовались снимки LANDSAT-8 (сцены WRS2: 146x30; 146x29; 144x28) и SENTINEL-2A (с 2015 года), доступные на свободной основе на интернет ресурсах NASA\ USGS [2]. Пространственное разрешение спутниковой съемки в 30 м (LANDSAT-8) и 20 м (SENTINEL-2A) [2] позволяет с высокой точностью определять изменения площади водных зеркал, в определенном диапазоне их размеров [3].

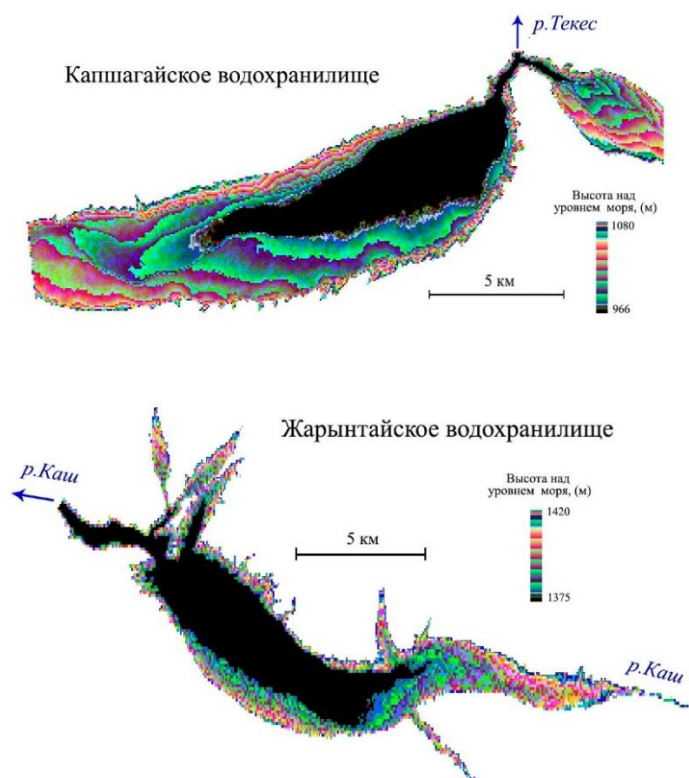


Рис. 1. Структура цифровой модели рельефа местности SRTM- v.2 для рабочих частей двух доминирующих водохранилищ в бассейне р.Иле (КНР)

Цифровые модели ложа Кашагайского и Жарынтайского водохранилищ, необходимые для расчета запасов воды в их рабочих частях [4], базировались на глобальной спутниковой 3D модели рельефа местности – SRTM [5], рис.1. Модель

SRTM была создавана на основе спутниковых данных, полученных в 1999 году, т.е. до начала строительства водохранилища (2004 г.). Поэтому, эта информация содержит данные для территории современного дна резервуара. В работе использовался вариант SRTM-2000 v.2, используемый в качестве DEM (Digital Elevation Model) в системе Google-Earth.

Полученные результаты

Малооблачные спутниковые снимки LANDSAT-8 (разрешение 30 м) и SENTINEL-2A (разрешение 20 м) из системы GLOVIS [2]: за 27 марта; 4, 15 апреля; 2, 17, 30 мая; и 1 июня использовались для дополнения имеющихся данных по мониторингу площадей водных зеркал озер Дунгарской равнины: Манас [6], см. рис.2, и Эби-Нур [7], см. рис.3. А также расчетным запасам воды в рабочих частях водохранилищ китайской части бассейна р. Иле: Капшагайское (р.Текес) [8], см. рис.4 и Жарынтайское, (р.Каш), [8], см.рис.5.

Обсуждение

Озеро Манас

Озеро Манас, расположенное в западной части Джунгарской равнины СУАР КНР, является конечным водным объектом для доминирующей одноименной реки (объем годового стока около 1 куб. км), берущей свое начало в ледниках хребта Ирен-Хабырга. В 60-х годах прошлого века, в связи с ростом водозабора на сельскохозяйственные нужды озеро Манас полностью пересохло [9]. Временное водное зеркало перестало формироваться даже в многоводные годы. Восстановление озера Манас началось в 1999 году. Причиной, по всей видимости, стал приход в его бассейн магистрального канала межбассейновой переброски воды - «Кара-Ертис - Карамай» [6, 10]. С 2006 года дополнительной поддержкой водного баланса озера стал второй магистральный водный канал - «Кара-Ертис - Урумчи», который, также оканчивается в его бассейне [11]. Таким образом, в настоящее время озеро Манас, фактически, стало искусственным резервуаром, наполненность которого обеспечивается, в большей степени водными каналами, чем естественным стоком региональных рек. Подтверждением этого является отсутствие естественной внутри сезонной, вариативности размера озера, включающей весенний максимум и осенний минимум.

По имеющимся данным (рис.2.), зависимость размера озера Манас от водности года все еще присутствует, но точный механизм пока не определен. Это либо увеличение объема воды транспортируемого из реки Кара-Ертис в бассейн озера в многоводные годы, либо реакция на повышенный уровень сезонного стока местных рек. Возможно, это следствие обеих причин. В настоящий момент, площадь озера Манас находится на спаде после ее значительного роста в 2016 году.

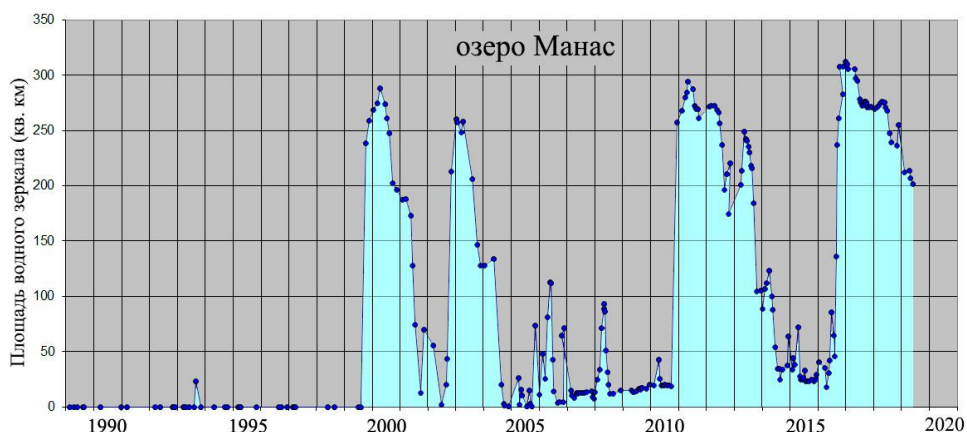


Рис. 2. Результаты спутникового мониторинга над размерами водного зеркала озера Манас (Джунгарская равнина, СУАР КНР). Данные до 2017 года взяты из [6]

Озеро Эби-Нур

Озеро Эби-Нур, относится к Балхаш-Алакольской группе озер, занимая крайнюю, юго-восточную позицию. Его размер, около 1000 кв. км, что в 3 раза превышает размер озера Манас, см. рис.2 и 3. Реакция размера озера на погодные условия года достаточно выражена. В отличие от оз.Манас, у оз.Эби-Нур присутствует значимая внутри сезонная вариативность площади зеркала, включающая весенний максимум и осенний минимум. В целом, озеро находится под растущим давлением сельскохозяйственного водопотребления. Значительный рост площади пашни [8] в бассейне оз. Эби-Нур формирует многолетнюю тенденцию на уменьшение его площади, которая сочетается с многолетней погодной вариативностью регионального увлажнения. Суммарно, эти два фактора формируют современную (1989-2019 гг.) динамику площади его водного зеркала, рис.2. Первая половина 2019 года продолжает тенденцию на уменьшение озера после его значительного роста, вызванного исключительно влажным 2016 годом.

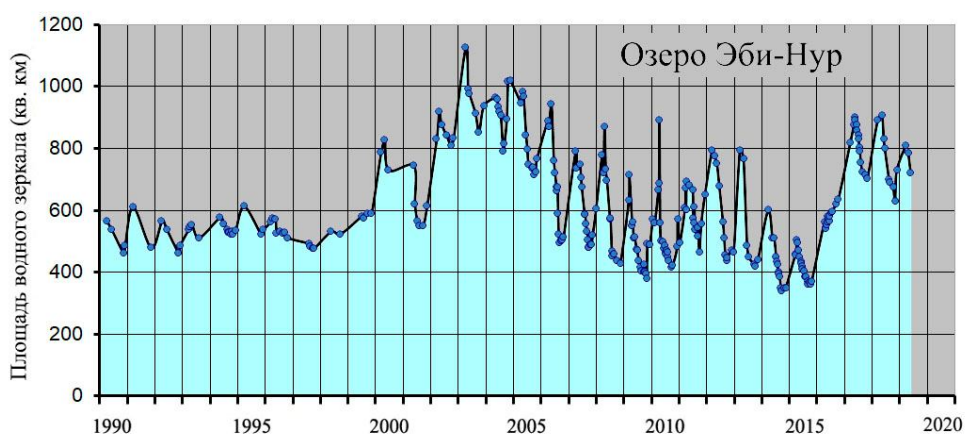


Рис.3. Результаты спутникового мониторинга размера водного зеркала озера Эби-Нур (Джунгарская равнина, СУАР КНР). Данные до 2017 года взяты из [7].

Капшагайское водохранилище на реке Текес

Капшагайское водохранилище, в составе соответствующего гидроузла, было построено в СУАР КНР в нижнем течение реки Текес в 2006 году [12]. Начиная с 2007 года, водохранилище работает в энергетическом режиме. Пополнение резервуара из стока р. Текес происходит в летний период, а сбратывание, для выработки электроэнергии, в холодный. Достаточно большой рабочий объем резервуара (1,4 куб. км), относительно объема годового стока реки Иле на границе КНР-Казахстан (12 куб. км) приводит к возможности существенного влияния работы водохранилища на расход воды в реке ниже по течению [13]. Этим обстоятельством объясняется особый интерес к работе этого объекта.

Первая половина 2019 года, см. рис.2, отмечалась, практически полным, опорожнением Капшагайского водохранилища. В период с 14 мая по 30 мая регистрировались исторические минимумы запаса воды. Причина - плановая работа резервуара на фоне холодной погоды апреля-мая 2019 года, с низкими расходами воды в горных реках Восточного Тянь-Шаня. В целом, режим работы резервуара в 2019 году очень близок к 2018 году. 2018 год отличался острым (неожиданным для специалистов) техногенным маловодьем в августе-сентябре. Повторение в 2019 году ситуации с маловодьем конца лета представляется вполне возможным. Отличием сезона 2019 года от прошлого, 2018 года, является больший запас снега, в зонах формирования стока (высота снеговой линии существенно ниже). Если фактор наличия больших запасов горного снега сможет существенно увеличить сток горных рек, маловодье августа-сентября в 2019 году может не повториться.

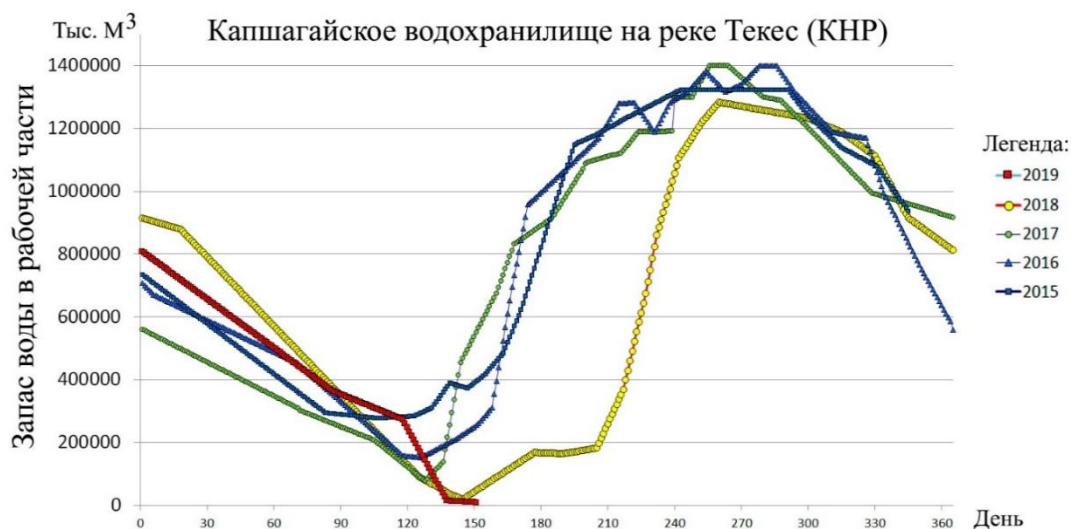


Рис. 4. Результаты спутниковых оценок запасов воды в рабочей части Капшагайского водохранилища (р.Текес, СУАР КНР). Данные до 2017 года взяты из [8]

Жарынтайское водохранилище на реке Каш

Жарынтайское водохранилище на реке Каш (второй по значимости приток р.Иле с объемом годового стока в 4 куб. км) было построено в 2006 году [8]. Особый интерес к водопользованию в бассейне реки Каш (КНР) связан с его географическим расположением. Узкие и высокие хребты восточного Тянь-Шаня - Боро-Хоро (высшая точка 4500 м над уровнем моря) и Ирен-Хабьрга (5500 м над

уровнем моря), формируют водораздел между водоизбыточной долиной реки Иле и вододефицитной Джунгарской равниной. Достаточно многоводная река Каш, длиной 350 км, собирает воду в ледниках хребта Ирен-Хабырга и впадает в р.Иле. Таким образом, задача обводнения земель Джунгарской равнины за счет стока реки Иле может решаться путем перенаправления туда части стока р. Каш. Технически для этого необходимы аккумулирующие воду гидротехнические сооружения и тоннельные водоводы. Наиболее крупное водохранилище бассейна р.Каш (Жарынтайское, рабочий объем 1,6 куб. км), теоретически, может быть задействовано в подобном проекте.

Мониторинг Жарынтайского водохранилища, см. рис.5, показывает практически полную идентичность режима 2019 года с режимами 2017 и 2018 гг. Отличительной чертой является несколько пониженный уровень запаса воды в апреле-мае 2019 года, вызванный пониженной из-за холодной погоды водностью горных рек региона.

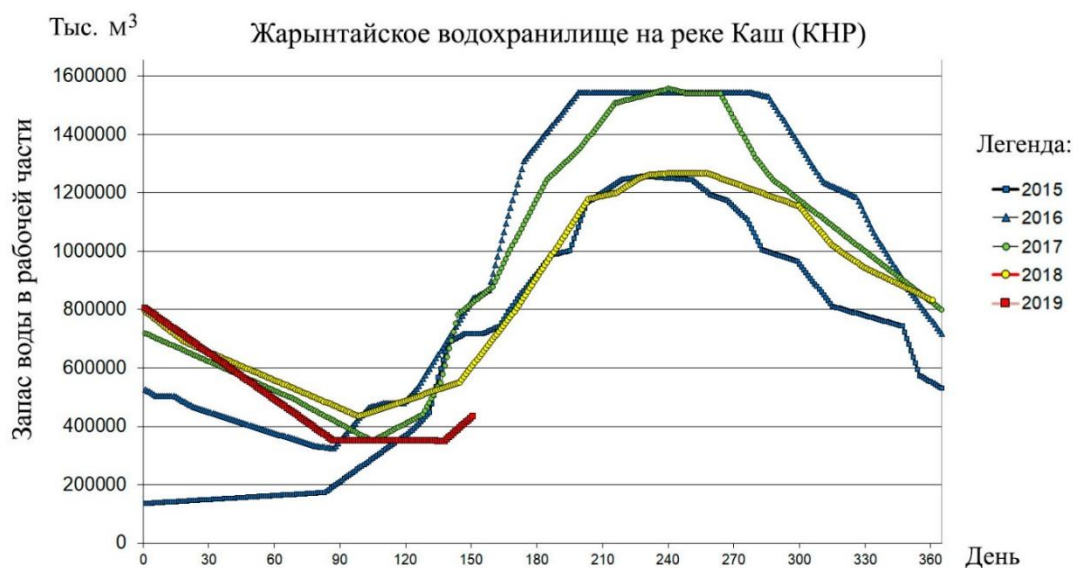


Рис. 5. Результаты спутниковых оценок запасов воды в рабочей части Жарынтайского водохранилища (р.Каш, СУАР КНР). Данные до 2017 года взяты из [8]

Заключение

Условия ограниченного информационного обмена оперативной гидрологической информацией по состоянию верхних частей трансграничных (КНР-Казахстан) речных бассейнов, относящихся к юрисдикции КНР, требуют развития объективных, дистанционных методов гидрологического мониторинга. Спутниковые данные, в этом случае, представляют собой практически единственный регулярный источник информации. Таким образом, оперативный спутниковый мониторинг сопредельных территорий СУАР КНР, в пределах трансграничных речных бассейнов (р.Иле и Кара-Ертис) и зависимых территорий (Джунгарская равнина), имеет стратегическое значение для заинтересованных государственных структур Казахстана.

Работа выполнена при поддержке грантового финансирования Министерства образования и науки Республики Казахстан, проект № AP 05134241.

Литература:

1. Сборник протоколов заседаний Казахстанско-Китайской совместной комиссии по использованию и охране трансграничных рек. 2001-2008. - http://www.cawater-info.net/library/rus/protokols_kaz-china.pdf (дата обращения: 16.05.2019).
2. USGS Global Visualization Viewer Data Portal. - <https://glovis.usgs.gov/app?fullscreen=0> (дата обращения: 4.06.2019)].
3. Kornienko S. G. Analysis of Errors in Estimating Changes in Water Body Areas by Satellite Data: Case Study of Thermokarst Lakes in Yamal Peninsula// Water Resources 2017, Vol.44, No.2, pp.180-191.
4. Терехов А.Г., Пак И.Т., Долгих С.А. Данные LANDSAT 5,7,8 и ЦМР в задаче мониторинга гидрологического режима Капчагайского водохранилища на реке Текес (китайская часть бассейна реки Иле) // Ж. Современные проблемы дистанционного зондирования Земли из космоса. – 2015.- Т.12.- № 6.- С. 174-182.
5. Rodriguez E., Morris C.S., Belz J.E., Chapin E.C., Martin J.M., Daffer W. An assessment of the SRTM topographic products // Technical Report JPL D-31639.Pasadena, California: Jet Propulsion Laboratory, 2005, pp.143.
6. Терехов А.Г. Мониторинг озера Манас (КНР) в период 1989-2017 годов по спутниковым данным LANDSAT// Гидрометеорология и Экология.- 2018.- № 2(89).- С.63-72.
7. Терехов А.Г., Долгих С.А. Спутниковый мониторинг водного зеркала оз.Эби-Нур (КНР) в период 1990-2017 гг.//Гидрометеорология и Экология – 2017. - №3 (86).- С.72-79.
8. Терехов А., Калимолдаев М., Пак И., Долгих С. Экспансия агроландшафта и состояние основных водных объектов на территории Синьцзян-Уйгурского автономного района (бассейны реки Иле и озера Эби-Нур) по данным спутниковой съемки / в кн. Новые методы и результаты исследований ландшафтов в Европе, Центральной Азии и Сибири: в 5 т. // под ред. В.Г.Сычёва, Л.Мюллера. – М.: Изд-во ФГБНУ «ВНИИ агрохимии», 2018. – Т. 3. – С. 219-223.
9. PetroChina Xinjiang Oilfield Emission Reduction and Afforestation Project UNFCCC /CCNUCC. 2012. - http://cdm.unfccc.int/filestorage/b/w/GIZSYQVOM3D49RL1NH8AJTPKEC7B52.pdf/PDD_PetroChina%20Xinjiang%20Afforestation.pdf?t=WDZ8cNBzOW1ifDA7MKXb1xtUAaUw702gDJD7 (дата обращения: 10.04.2019).
10. Терехов А.Г. Технические характеристики водного канала Кара-Ертис – Карамай: спутниковые оценки // Гидрометеорология и Экология – 2017. - №4 (87).- С.54-62.
11. Терехов А.Г. Технические характеристики водного канала Кара-Ертис – Урумчи: спутниковые оценки // Гидрометеорология и Экология – 2018. - №4 (91).- С.63-74.
12. Терехов А.Г., Долгих С.А., Никифорова Л.Н. Китайский сектор бассейна р. Иле: спутниковая диагностика изменений гидрографа реки Текес в результате строительства в 2006 году Капчагайского водохранилища (КНР) // Гидрометеорология и Экология – 2016. - №4 (83).- С.24-31.
13. Терехов А.Г. Потенциал влияния водохранилищ китайского сектора бассейна реки Иле на режим её водности на границе КНР-Казахстан // Гидрометеорология и Экология – 2016. - №4 (83).- С.63-71.

ХАОС ПУАНКАРЕ В НЕЙРОННЫХ СЕТЯХ

Тлеубергенова М.А., Жаманшин А.У.
e-mail: madina_1970@mail.ru, akylbek78@mail.ru

*Актюбинский региональный государственный университет
им. К. Жубанова,
Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** Для исследования хаоса в нейронных сетях с помощью дифференциальных уравнений, мы применяем результаты, начатые в работах [1,2], где впервые дано определение хаоса Пуанкаре. Наличие хаоса подтверждается присутствием непредсказуемых решений. Настоящее исследование рассматривает существование и единственность асимптотически устойчивого непредсказуемого решения для клеточной нейронной сети с шунтирующим торможением (КНС). Приводятся соответствующие пример с графиком, который подтверждает теоретические результаты.*

1. Введение и основные сведения

Наличие хаоса в нейронных сетях полезно для разделения сегментов изображения [3], обработки информации [4] и синхронизации [5]. Хаос может быть порожден самой нейронной сетью (эндогенный хаос), или хаотическое влияние вне нейронной сети может быть реализовано в ее выходе (экзогенный хаос). Эндогенный хаос в нейронных сетях широко изучался в литературе, но до сих пор не был эффективно рассмотрен. Причина заключается в отсутствии тщательно разработанного механизма ввода / вывода для этого явления.

В [1] было доказано, что непредсказуемая точка в квази-минимальном множестве приводит к существованию хаоса Пуанкаре. Таким образом, если показать существование непредсказуемого решения уравнения, то существует и хаос.

Поскольку основная часть результатов о хаотических движениях была сформулирована в терминах дифференциальных и разностных уравнениях, мы можем предположить, что все эти достижения должны быть включены и развиты в теории динамических систем или, более конкретно, в теории дифференциальных уравнений или гибридных систем.

В данной статье строго доказано существование непредсказуемых движений для КНС с шунтирующим торможением. Чтобы доказать существование непредсказуемых движений, мы используем топологию равномерной сходимости на компактах и принцип сжимающего отображения. Приведен пример, который подтверждает теоретические результаты.

КНС с шунтирующим торможением, который был представлен Боузердомом и Пинтером в работе [6], играет исключительную роль в психофизике, речи, восприятии, робототехнике, адаптивном распознавании образов, видении и обработке изображений. Модель КНС с шунтирующим торможением в своей первоначальной формулировке [6] представляет собой двумерную сетку ячеек обработки.

Пусть C_{ij} означает клетку в (i, j) положении решетки. Обозначим через $N_r(i, j)$ r -окрестность клетки C_{ij} , такое, что

$$N_r(i, j) = \{C_{kl}: \max\{|k - i|, |l - j|\} \leq r, 1 \leq k \leq m, 1 \leq l \leq n\},$$

где m и n фиксированные натуральные числа. В КНС соседние клетки осуществляют взаимные тормозные взаимодействия шунтирующего типа. Динамика клетки C_{ij} описывается следующим нелинейным обыкновенным дифференциальным уравнением,

$$\frac{dy_{ij}}{dt} = -a_{ij}y_{ij} - \sum_{C_{kl} \in N_r(i, j)} C_{ij}^{kl} f(y_{kl}(t))y_{ij} + v_{ij}(t)$$

где y_{ij} – это активность клетки C_{ij} ; $v_{ij}(t)$ внешний вход в ячейку C_{ij} ; постоянные a_{ij} представляют скорость пассивного распада активности клеток; $C_{ij}^{kl} \geq 0$ сила сцепления постсинаптической активности клетки C_{kl} передается в клетку C_{ij} ; и функция активации $f(y_{kl})$ является положительной непрерывной функцией, представляющей выходную скорость ячейки C_{kl} .

Основная новизна настоящей работы - существование, единственность и асимптотическая устойчивость непредсказуемых решений в КНС с шунтирующим торможением, динамика которых определяется уравнением (1).

На протяжении всей статьи R и N обозначают множество действительных и натуральных чисел. Также используется норма $\|u\|_1 = \sup_{t \in R} \|u(t)\|$, где $\|u\| = \max_{(i, j)} |u_{ij}|$, $u = (u_{11}, \dots, u_{1n}, \dots, u_{m1}, \dots, u_{mn})$, $u_{ij} \in R, i = 1, \dots, m, j = 1, \dots, n$.

Определение 1. Равномерно непрерывная и ограниченная функция $u: R \rightarrow R^{m \times n}$ называется непредсказуемой, если существуют положительные числа ε_0, δ и последовательности t_p, s_p которые расходятся и $u(t + t_p) \rightarrow u(t)$ при $p \rightarrow \infty$ равномерно на компактных подмножествах R и $\|u(t + t_p) - u(t)\| \geq \varepsilon_0$ для всех $t \in [s_p - \delta, s_p + \delta]$ и $p \in N$.

2. Основные результаты

Пусть B – множество функций $u_{ij} = (u_{11}, \dots, u_{1n}, \dots, u_{m1}, \dots, u_{mn})$, $u_{ij} \in R$ $i = 1, \dots, m, j = 1, \dots, n$, где $m, n \in N$, такие что:

(У1) функция $u(t)$ равномерно непрерывна так, что $\|u\|_1 < H$, где H положительное действительное число;

(У2) существует последовательность $t_p, t_p \rightarrow \infty$ при $p \rightarrow \infty$ такое, что для каждого $u(t) \in B$ последовательность $u(t + t_p)$ равномерно сходится к функции $u(t)$ на каждом замкнутом и ограниченном интервале вещественной оси.

Наше основное предположение заключается в следующем.

(П1) Функция $v = (v_{11}, \dots, v_{1n}, \dots, v_{m1}, \dots, v_{mn})$, $v_{ij} \in R, i = 1, \dots, m, j = 1, \dots, n$, в системе (1) принадлежит множеству B и она непредсказуемая так, что существуют положительные числа δ, ε_0 и последовательность $s_p, s_p \rightarrow \infty$ при $p \rightarrow \infty$ которые удовлетворяют неравенству $|v_{ij}(t + t_p) - v_{ij}v_{ij}(t)| \geq \varepsilon_0$ для всех $i = 1, \dots, m, j = 1, \dots, n, t \in [s_p - \delta, s_p + \delta]$, и $p \in N$.

Следующие условия необходимы для этой статьи.

(П2) $\gamma = \min_{(i, j)} a_{ij} > 0$;

(П3) существуют положительные числа m_{ij} такие, что $\sup_{t \in R} |v_{ij}| \leq m_{ij}$;

(П4) существуют положительные числа m_f такие, что $\sup_{|s| < H} |f(s)| \leq m_f$;

(П5) существует положительное число L такое, что $|f(s_1) - f(s_2)| \leq L|s_1 - s_2|$ для всех $s_1, s_2, |s_1| < H, |s_2| < H$;

$$(П6) a_{ij} > m_f \sum_{C_{kl} \in N_r(i,j)} C_{ij}^{kl};$$

$$(П7) \frac{m_f}{a_{ij} - m_f \sum_{C_{kl} \in N_r(i,j)} C_{ij}^{kl}} < H;$$

$$(П8) \frac{LH + m_f}{\gamma} \max_{(i,j)} \sum_{C_{kl} \in N_r(i,j)} C_{ij}^{kl} < 1.$$

Основным результатом нашей работы является следующая теорема.

Теорема 1. Предположим, что выполняются условия (П1) - (П8), тогда система (1) имеет единственное асимптотически устойчивое непредсказуемое решение.

3. Пример

В работе [7] показано, что наличие непредсказуемой функции неизбежно сопровождается хаосом Пуанкаре. Следовательно, мы можем искать подтверждение результатов для непредсказуемости, наблюдающей неравномерность в моделировании.

Рассмотрим логистическое дискретное уравнение

$$\lambda_{i+1} = \mu \lambda_i (1 - \lambda_i), \quad (2)$$

где $i \in Z$. Интервал $[0,1]$ инвариантен относительно итераций (2) для $\mu \in (0,4]$. В [7] доказано, что уравнение (2) имеет непредсказуемое решение. Пусть $\varphi_i, i \in Z$, является непредсказуемым решением уравнения (2) с $\mu = 3.93$ внутри единичного интервала $[0,1]$ и введем функцию

$$\theta(t) = \int_{-\infty}^t e^{-5(t-s)} \Omega(s) ds, \quad (3)$$

где $\Omega(t)$ кусочно-постоянная функция, определенная на вещественной оси уравнением $\Omega(t) = \rho_i$ для $t \in [i, i+1), i \in Z$. Стоит отметить, что $\theta(t)$ является единственным глобально экспоненциально устойчивым решением дифференциального уравнения

$$\frac{dv}{dt} = -5v(t) + \Omega(t). \quad (4)$$

В работе [8] доказано, что функция $\theta(t)$ является непредсказуемой.

Рассмотрим следующую КНС:

$$\frac{dx_{ij}}{dt} = -a_{ij}x_{ij} - \sum_{C_{kl} \in N_1(i,j)} C_{ij}^{kl} f(x_{kl}(t))x_{ij} + v_{ij}(t), \quad (5)$$

$$i, j = 1, 2, 3,$$

где $f(s) = 0.5 \arctan(s)$, $v_{11}(t) = -6\theta(t)$, $v_{12}(t) = 4\theta(t)$, $v_{13}(t) = -3\theta(t) + 2$, $v_{21}(t) = 8\theta(t) + 3$, $v_{22}(t) = 5\theta(t)$, $v_{23}(t) = 6\theta(t) - 1$, $v_{31}(t) = -7\theta(t) + 2$, $v_{32}(t) = 3\theta(t)$, $v_{33}(t) = -5\theta(t) + 3$ и

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 2 & 5 & 3 \\ 6 & 7 & 3 \\ 8 & 4 & 2 \end{pmatrix}, \quad \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{pmatrix} = \begin{pmatrix} 0.05 & 0.02 & 0 \\ 0.04 & 0.07 & 0.03 \\ 0.02 & 0 & 0.06 \end{pmatrix}.$$

Рисунок 1 показывает результат теоремы (1) и изображает непредсказуемое решение $x(t)$ системы (4) с начальными значениями $x_{11}(0) = 0.5432$, $x_{12}(0) = 0.1234$, $x_{13}(0) = 0.4678$, $x_{21}(0) = 0.8931$, $x_{22}(0) = 0.0132$, $x_{23}(0) = 0.7143$, $x_{31}(0) = 0.15247$, $x_{32}(0) = 0.1325$, $x_{33}(0) = -0.0014$.

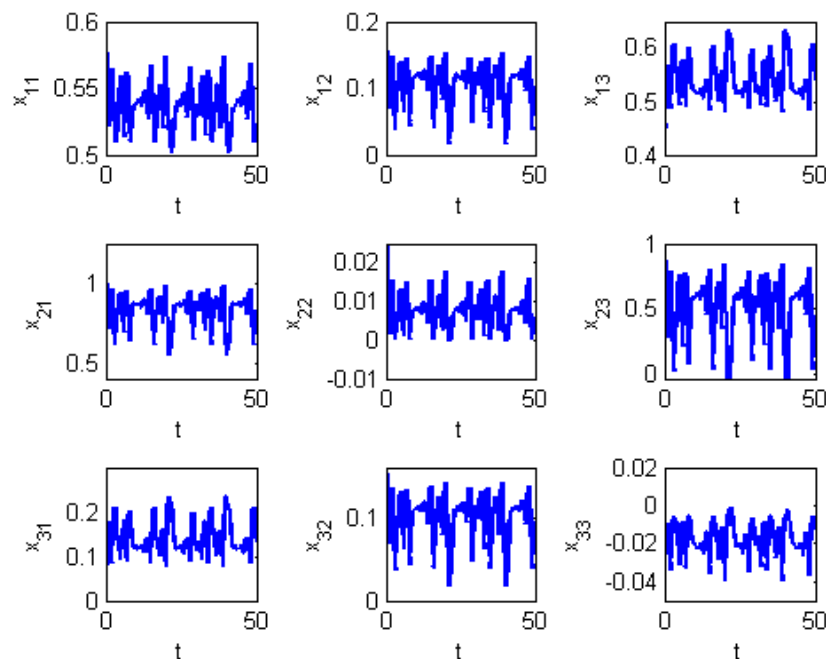


Рис.1 Непредсказуемое решение системы (5)

Данная работа выполнена в рамках грантового проекта AP05132573 МОН РК "Клеточные нейронные сети с непрерывным и дискретным временем и сингулярными возмущениями" (2018-2020).

Литература

1. Akhmet, M., Fen, M.O. Unpredictable points and chaos /M. Akhmet, M.O. Fen // -Commun. Nonlinear Sci. Numer. Simulat. 2016. – N. 40. -P. 1-5.
2. Akhmet, M., Fen, M.O. Existence of unpredictable solutions and chaos / M. Akhmet, M.O. Fen // -Turk. J. Math. 2016. - P. 1-5.
3. Shibasaki, M., Adachi, M. Response to external input of chaotic neural networks based on Newman-Watts model / M. Shibasaki, M. Adachi // - WCCI 2012 IEEE World Congress on Computational Intelligence, Brisbane, Australia. 2012. - P. 1-7.
4. Nara, S., Davis, P., Kawachi, M., Totsuji, H. Chaotic memory dynamics in a recurrent neural network with cycle memories embedded by pseudoinverse method / S.

Nara, P. Davis, M. Kawachi, H. Totsuji // - Int. J. Bifurcation Chaos. 1995. - N. 5. - P. 1205-1212.

5. Liu, Q., Zhang, S. Adaptive lag synchronization of chaotic Cohen-Grossberg neural networks with discrete delays / Q. Liu, S. Zhang // -Chaos. 2012. - N. 22. 033123.

6. Bouzerdoum, A., Pinter, R.B. Shunting inhibitory cellular neural networks: derivation and stability analysis / A. Bouzerdoum, R.B. Pinter // IEEE Trans. Circuits Systems I Fund. Theory and Appl. 1993. –N. 40. – P. 215-221.

7. Akhmet, M., Fen, M.O. Poincare chaos and unpredictable functions / M. Akhmet, M.O. Fen // - Commun. Nonlinear Sci. Numer. Simulat. 2017. - N. 48. – P. 85-94.

8. Akhmet, M., Fen, M.O. Non-autonomous equations with unpredictable solutions / M. Akhmet, M.O. Fen // - Commun. Nonlinear Sci. Numer. Simulat. 2018. – N. 59. – P. 657-670.

ИССЛЕДОВАНИЕ ПРИЗНАКОВ ДЛЯ АВТОМАТИЧЕСКОЙ ИДЕНТИФИКАЦИИ ЖАНРОВ ПУБЛИКАЦИЙ В СМИ

Уалиева И.М., Мусабаев Р.Р.

e-mail: i.ualiyeva@gmail.com, rmusab@gmail.com

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Аннотация. В статье мы описываем подход, основанный на исследовании и выделении признаков, для автоматического распознавания разных жанров публикаций. В частности, на небольшом корпусе, собранном вручную из разнотипных источников: спорт, конспирология, бизнес и наука, мы провели эксперимент по автоматическому распознаванию типов текстов. Первоначально мы исследовали признаки, присущие разным типам текстов, сформировали вектор признаков и далее применили его для распознавания. Результаты показали, что сформированные признаки позволяют распознать типы публикаций с достаточно высокой точностью.

Введение

Публикации в СМИ преследуют разные цели. Цель одних – только информировать читателя о событии: кто сделал, что сделал, где и когда оно произошло. Другие не только информируют о событии, но и добавляют к деталям справочную информацию, обобщают информацию, делая выводы. Такие публикации содержат и факты, и мнение автора. Третьи только отталкиваются от конкретного факта и пытаются трактовать события нужным им образом, возможно, манипулируя мнением читателя.

Мы предлагаем подход для автоматического распознавания таких типов публикаций, основанный на выделении признаков. Мы исследуем и выделяем ряд признаков, присущих трем типам публикаций, и на небольшом корпусе исследуем их применимость для автоматического распознавания текстов. Для исследования типов публикаций мы обратились к работам лингвистов, изучающим жанры [1, 2], для исследования признаков мы опирались на исследования [3-7].

Мы собрали свой небольшой корпус, достаточный для того, чтобы исследовать признаки и провести эксперимент. Источники имеют принципиально разную направленность: спортивный новостной сайт, конспирологический сайт, сайт с научно-популярными статьями, и сайт о бизнесе. Мы отобрали из этих источников статьи, в которых наиболее видна типизация публикаций: информативная публикация, аналитическая или артистическая. Каждый из типов мы обозначили как жанр. В работах лингвистов еще ведутся дискуссии по определению жанров, но так или иначе, такая типизация присутствует во многих их работах [1].

Предлагаемый нами подход можно использовать для разного рода задач. Например, для классификации фейковых новостей или поиска пропаганды, или поиска самых свежих новостей, новостей финансового рынка или другой информации. Недостатком такого подхода является необходимость иметь размеченный корпус, что для русского языка на данный момент довольно затруднительно.

В своей работе мы собрали корпус из открытых источников официальных и полуофициальных СМИ и разделили его по типам; исследовали и выделили признаки; исследовали с какой точностью признаки автоматически распознают типы.

Корпус

Корпус были собран из открытых публикаций новостных и аналитических порталов за 2018-2019 год. Было выбрано четыре разнотипных направления: *спорт, научно-популярные статьи, конспирология и бизнес.*

Спорт. Новостные статьи о спорте, репортажи, информируют читателя о результатах соревнований, о спортсменах. Такие тексты обычно короткие, выбор пал на те, длина которых не менее одиннадцати предложений. Такие тексты мы классифицировали как информативные.

Конспирология. Конспирология – это «концепция, объявляющая отдельные общественно значимые события, определённые исторические явления или даже ход истории в целом результатом заговора со стороны некоторой группы людей (государство в государстве, мировое правительство), управляющих этим процессом из корысти, амбиций или иных групповых, клановых и других интересов». Авторы этого направления «занимаются не сбором фактов, опровергающих или подтверждающих существование того или иного заговора, а интерпретируют любые факты в пользу существования заговора». Такого рода статьи только отгалкиваются от фактов или даже просто строят гипотезу, которую не подтверждают фактами, а свое мнение оспаривают простыми аргументами: «Пока в Интернете я этого не нашел, но скоро это откроется» или «Они не хотят пока открываться». Особенностью таких текстов является то, что авторы применяют фразы и предложения, которые привлекают внимание пользователей.

Бизнес. Сайт публикует аналитические статьи о развитии компаний, их владельцах, о процессах, происходящих на рынках и биржах. В таких статьях авторы обязательно интерпретируют факты, т.е. аргументированно пишут о своей точке зрения, пытаясь глубоко раскрыть тему путем обобщений, выводов, аргументов. Длина текстов от 100 до 200 предложений.

Наука. Научно-популярные – это авторские статьи, по структуре своей схожи с научными статьями, но в более простой подаче. Авторы статей обычно ученые-исследователи, статьи содержат большое количество фактов, исследователи

делятся в них своим опытом. По аргументации, глубине изложения они близки к статьям о бизнесе. Статьи более короткие по длине, чем конспирологические.

Конспирологические статьи мы отнесли к артистическому жанру, научно-популярные и статьи о бизнесе отнесли к аналитическому жанру, новостные к информативному жанру.

Артистические тексты более индивидуальные и персонифицированные. Восприятие и интерпретация читателями одной и той же артистической статьи могут быть разными и будет зависеть от индивидуальных характерологических особенностей читателя.

Артистические тексты отражают воображаемый мир автора, его собственную оценку, информативные статьи передают реальные факты и события в прямой и краткой манере, без интерпретаций, аналитические статьи содержат и факты и их интерпретацию, но они богаче стилистически.

Неартистические тексты влияют на мнение читателя косвенно, через стилистические приемы, образы, повторения слов, тогда как считается, что информативные тексты передают информацию прямо, ясно, без однозначности.

Это примеры некоторых предложений, взятых из статей разных жанров (тексты взяты не из основного корпуса):

Научно-популярная статья:

*Для некоторых видов мартышкообразных обезьян (*Cercopithecidae*, см. Old World monkey) показано наличие дифференцированных звуковых сигналов, обозначающих разные виды опасностей (см.: Мартышки произносят фразы из двух слов, «Элементы», 21.05.2006). Большие всего данных собрано по трем сигналам («леопард», «змея» и «орел»), характерным для обитающих в Восточной Африке верветок *Chlorocebus pygerythrus*.*

Конспирологическая статья:

Одной из спорных и противоречивых идей в творчестве Блаватской является концепция об эволюционном цикле рас, часть которой изложена во втором томе «Тайной доктрины». Некоторые исследователи считают, что теория рас «от Блаватской» была взята за основу идеологами Третьего рейха.

«Инопланетяне сейчас живут среди нас и ведут за нами постоянное наблюдение»

«Человечество чудесным образом возвращается на десятки тысяч лет назад. Библия трактует потоп как наказание за людские пороки.»

По прошествии 5 сезонов в «Игре престолов» среди главных героев остаются практически одни негодяи, насильники, развратники, изменники и предатели. Воспитательный вывод напрашивается сам собой – хорошие парни долго не живут и им уж точно не следует лезть в политику. По сути, фильм продвигает всё тот же лживый тезис «политика – грязное дело», внедрение которого в сознание масс предотвращает проникновение в сферу управления честных и порядочных людей.

Бизнес-статья

Постепенно интернет-популярность пенсионерки стала конвертироваться в реальные деньги: в 2016 году количество заказов «Ленбагета» начало расти, а по итогам 2017-го компания принесла первую ощутимую выручку – 2 млн рублей. Важную роль сыграло то, что ролики Лозовской стали смотреть за пределами России. «Благодаря YouTube обо мне узнали наши соотечественники по всему миру: люди стали присылать свои работы из Америки, Канады, Германии», — перечисляет Лозовская.

Новости

Их поединок состоялся 8 июня в Нью-Йорке (США) Madison Square Garden и завершился победой GGG нокаутом в четвертом раунде.

Признаки

Мы полагаем, что конспирологические и аналитические статьи могут отличаться от информативных такими статистическими признаками как среднее количество слов в предложении, количество слов в документе, количество предложений, среднее количество абзацев. Новостные статьи, в свою очередь, более информативные, содержат больше цифр и именованных сущностей. Также мы предполагали, что кавычки, как признак прямой речи, чаще всего встречаются в аналитических и конспирологических текстах. Лексические признаки включают частоту встречаемости глаголов, частоту встречаемости прилагательных и существительных в документе. Мы ожидали, что большое количество глаголов в тексте будет ассоциироваться с более информативными новостными текстами. Ранее в своих работах мы предполагали, что существительные во множественном числе и наличие слов-квантификаторов как *все, каждый, любой, всегда, вечно, никогда, постоянно, никто, ничего* – это признаки генерализованных текстов [6] и было интересно анализировать их распределение в жанрах. Окончательно было отобрано двенадцать стилистических и лексических признаков (см. таблицу 1).

Эксперименты и результаты

Мы применяем алгоритм кластеризации KMeans [8,9] для каждого признака и для совокупности признаков. Алгоритм KMeans разбивает множество элементов векторного пространства (матрица, размер которой $m * n$, где m - количество текстов, n - количество признаков) на заранее известное количество кластеров (типов публикаций). Для нормализации векторов мы использовали линейную нормализацию:

$$X_{normalized} = (X_{raw} - X_{min}) / (X_{max} - X_{min})$$

Оценку точности алгоритма рассчитывали на основе Rand index метрики.

В таблице 1 показаны результаты работы алгоритма для каждого признака. Как видно из таблицы 1 точность распознавания для одного признака достигает 75% для цифр и чисел.

Таблица 1. Точность классификации текстов по жанрам (один признак)

№ признака	Признак	Точность
0	Количество слов	0.62
1	Количество предложений	0.62
2	Частота встречаемости знаков препинания	0.46
3	Количество кавычек	0.49
4	Количество цифр, чисел	0.75
5	Частота встречаемости существительных во множественном числе	0.46
6	Частота встречаемости слов-квантификаторов в документе	0.49
7	Количество абзацев	0.61
8	Среднее количество слов в предложении	0.56

9	Частота встречаемости существительных	0.48
10	Частота встречаемости глаголов	0.52
11	Частота встречаемости прилагательных	0.46
	Все признаки	0.67

Затем мы экспериментировали с разными вариациями признаков, чтобы посмотреть какая совокупность признаков показывает наиболее высокую точность, мы тренируем модель для всех совокупностей признаков. Совокупность четырех признаков: количество кавычек, цифр и чисел, среднее количество слов в предложении, частота встречаемости глаголов и прилагательных показала самую высокую точность 85%. Для совокупности всех признаков алгоритм показал 67%. Таким образом совокупность четырех признаков в целом дает более высокую точность распознавания, чем один признак или все признаки. Результаты работы алгоритма на совокупности признаков показаны в таблице 2.

Таблица 2. Точность классификации текстов по жанрам (комбинации признаков с наилучшей точностью распознавания)

Количество признаков	Комбинации признаков с наилучшей точностью распознавания	Точность
2	{4, 5}	0.75
3	{4,8,10}	0.78
4	{0, 4, 8, 10}	0.78
5	{3, 4, 8, 10, 11}	0.85
6	{0, 1, 2, 4, 5, 7}	0.83
7	{0, 1, 2, 4, 5, 8, 11}	0.83
8	{0, 1, 2, 4, 5, 6, 8, 11}	0.84
9	{0, 1, 2, 3, 4, 5, 7, 10, 11}	0.82
10	{0, 1, 2, 3, 4, 5, 6, 7, 8, 11}	0.73
11	{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 11}	0.73
Все признаки	{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11}	0.67

Метод главных компонент показал, что пять признаков определяют 85% точности распознавания. Оценка важности признаков показала, что количество кавычек внесло в распознавание 35.03%, чисел - 22.42%, среднее количество слов в предложении - 14.63%, частота встречаемости глаголов - 7.37%, частота встречаемости прилагательных - 6.41%.

Дополнительные эксперименты по распознаванию источников текстов показали, что алгоритм распознает новостные статьи с точностью 88.5%, конспирологические статьи с результатом 100%, бизнес статьи с результатом 67.7% и научно-популярные с результатом 75%. Более высокие результаты автоматического распознавания новостных и конспирологических статей могут быть результатом того, что имеющиеся в них синтаксические признаки более выражены, на наш взгляд. Тогда как научно-популярные и бизнес статьи различаются больше по лексическому составу. Это требует дополнительного изучения и ввода новых признаков, например, как частота встречаемости уникальных слов.

Заключение

Мы предлагаем подход, основанный на выделении признаков, для разделения текстов по жанрам. Полагается, что такой подход не зависит от языка, но эксперимент был проведен для русского языка. В частности, мы использовали такие признаки как наличие пунктуации, кавычек, длину слов, предложений, грамматические признаки и т.д. проведенный нами эксперимент показал, что наша модель может разделить жанры с довольно высокой точностью.

В будущем мы планируем провести эксперименты с большим количеством признаков и расширить пространство признаков за счет признаков для поиска юмора или слухов [10]. Мы также планируем использовать разные алгоритмы кластеризации для улучшения качества распознавания [8]. Предложенный подход позволит нам динамически разделять публикации на артистические и неартистические при парсинге новостных порталов и сайтов, что ускорит подготовку корпуса для его дальнейшей разметки.

Поддержка исследований. Работа проводится в рамках ПЦФ «Разработка информационных технологий и систем для стимулирования устойчивого развития личности как одна из основ развития цифрового Казахстана».

Литература

1. Beghtol, Clare. (2005). The Concept of Genre and Its Characteristics. Bulletin of the American Society for Information Science and Technology. 27. 19 - 17. 10.1002/bult.192.
2. Hanganu, Eduard. (2015). Generic Structures and Their Functions. 10.13140/RG.2.1.2882.7046.
3. Dearden, E., Baron, A.: Fool's Errand: Looking at April Fools Hoaxes as Disinformation through the Lens of Deception and Humour. In: Proceedings of 20th International Conference on Computational Linguistics and Intelligent Text Processing, CICLing 2019, La Rochelle, France, (2019)
4. Yang, Y., Nenkova, A.: Combining Lexical and Syntactic Features for Detecting Content-dense Texts in News (2017). arXiv preprint arXiv:1704.00440
5. Yatsko, V. A., Starikov, M. S., Butakov, A. V.: Automatic Genre Recognition and Adaptive Text Summarization. Automatic Documentation and Mathematical Linguistics. 44(3), 111–120 (2010)
6. Krassovitsky, A.M., Ualiyeva, I.M., Meirambekkyzy, Zh., Mussabayev, R.R.: Generalization annotation for its automatic recognition in media texts. In: Proceedings of 20th International Conference on Computational Linguistics and Intelligent Text Processing, CICLing 2019, La Rochelle, France, (2019)
7. Hardalov, M., Koychev, I., Nakov, P.: In Search of Credible News. Artificial Intelligence: Methodology, Systems, and Applications. AIMS A 2016. Lecture Notes in Computer Science, vol 9883. Springer, Cham 172-180 (2016) DOI https://doi.org/10.1007/978-3-319-44748-3_17
8. Yang, D., Lavie, A., Dyer, C., Hovy, E.: Humor recognition and humor anchor extraction. In: Proceedings of 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015, Lisbon, Portugal, pp. 2367–2376 (2015)
9. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second Edition. Springer, 764, (2017)
10. Zaharia, M., Chowdhury, M., Franklin, M.J., Shenker, S., Stoica, I.: Spark: cluster computing with working sets. In: Proceedings of 2nd USENIX Conference on Hot Topics in Cloud Computing, HotCloud 2010, Boston, MA, p. 10 (2010)

ОБЕСПЕЧЕНИЕ ЗАЩИТЫ ИНФОРМАЦИОННОЙ СИСТЕМЫ С ПОМОЩЬЮ ДВУХФАКТОРНОЙ АУТЕНТИФИКАЦИИ

Усатова О.А., Нысанбаева С.Е.
e-mail: uoa_olga@mail.ru, sultashal@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК
Республика Казахстан.*

***Аннотация.** В статье описаны методы применения аутентификации на основе второго фактора. Рассмотрены известные алгоритмы и протоколы двухфакторной аутентификации. Подробно описан алгоритм двухфакторной аутентификации с использованием мобильных устройств в качестве идентификаторов и генерации временного пароля на основе хеш-функции стандартов шифрования. Для автоматизированной системы управления разработаны модель двухфакторной аутентификации и последовательный алгоритм генерации временного пароля с использованием тригонометрических функций. Реализация системы основана на программной платформе Node.js с использованием языка программирования JavaScript, а также фреймворков и подключенных системных библиотек. Использована документоориентированная система управления базами данных с открытым исходным кодом MongoDB, для хранения и обработки информации.*

Введение

В нашей стране электронная информация защищается в основном зарубежными программными и аппаратно-программными средствами. Поэтому создание казахстанских средств защиты информации (в том числе и криптографических) является очень актуальной. В связи с этим приняты нормативно-правовые документы по созданию и использованию средств криптозащиты, в том числе и Концепция кибербезопасности «Киберщит Казахстана» [1]. В ней особо оговаривается «Предоставление приоритета исследованиям и собственной школе прикладной математики, по разработке средств криптографической защиты информации, криптологии, разработок по программируемым логическим интегральным схемам, квантовой криптографии и разработке защиты систем передачи, обработки и хранения информации, а также систем информационной безопасности и преодолению проблемы невысокой востребованности отечественных разработок».

Одним из наиболее распространенных методов защиты информации в настоящее время является парольный доступ к данным. Однако, наряду с несомненными достоинствами, этот метод защиты данных обладает и определенными недостатками: пароль можно забыть, его могут «взломать». В данной работе описывается предлагаемая система защиты данных с использованием двухфакторной аутентификации, основанной на генерации одноразового пароля. Актуальность системы заключается в том, что защита данных от взлома является важной частью для любого ресурса, предполагающего работу с этими данными.

Известные алгоритмы и протоколы

Двухфакторная аутентификация – это алгоритм обеспечения безопасности, в котором пользователь предлагает два разных фактора аутентификации, что улучшает защиту доступа и к учетным данным пользователя, и к ресурсам пользователя [2]. Традиционные системы используют имя пользователя и пароль для аутентификации. Этот метод обеспечивает минимальный уровень безопасности, поскольку имена и пароли можно легко перехватить и даже угадать. Двухфакторная аутентификация обеспечивает более высокую степень защиты, чем однофакторная аутентификация, при которой пользователь предлагает только один фактор, обычно пароль и используется для контроля доступа к чувствительным системам и данным. Рассмотрим некоторые известные протоколы и алгоритмы авторизации:

OAuth – протокол авторизации, который позволяет выдавать определенному сервису или приложению права на доступ к ресурсам пользователя на стороннем сервисе или приложении [3]. Данный протокол предоставляет возможность избавления от необходимости доверия стороннему приложению данных для авторизации. Этот протокол позволяет выдавать некоторый определенный набор прав, а не все сразу.

Выполнение протокола OAuth основано на использовании таких базовых веб-технологиях, как HTTP-запросы и редиректы (перенаправление маршрута). Вследствие этого его использование возможно на различных платформах с доступом к интернету и браузеру. Общая структура работы приложения, использующего протокол OAuth, включает два этапа:

- получение авторизации;
- обращение к защищенным ресурсам.

Результатом авторизации является «токен доступа» - определенный ключ, предъявление которого открывает доступ к защищенным ресурсам. Обращение к ним в самом простом случае происходит по HTTPS с указанием в заголовках или в качестве одного из параметров полученного токена доступа.

В протоколе описано несколько вариантов авторизации для различных ситуаций:

- авторизация для приложений, имеющих серверную часть, чаще всего, это сайты и веб-приложения;
- авторизация для полностью клиентских приложений;
- авторизация по логину и паролю;
- восстановление предыдущей авторизации.

Общий принцип работы с протоколом OAuth заключается в следующих шагах:

- редирекция страницы авторизации;
- на странице авторизации у пользователя запрашивается подтверждение выдачи прав;
- в случае согласия пользователя, браузер осуществляет редирекцию на URL, указанный при открытии страницы авторизации, с добавлением в GET-параметры специального ключа — «кода авторизации»;
- сервер приложения выполняет POST-запрос с полученным authorization code в качестве параметра, в результате которого возвращается токен доступа.

Безопасность OAuth во многом основана на SSL. Это сильно упрощает жизнь разработчикам, но требует дополнительных вычислительных ресурсов и

администрирования. Это может быть существенным вопросом в высоконагруженных проектах.

НОТР (HMAC-Based One-Time Password Algorithm) – алгоритм защищенной аутентификации с помощью использования одноразового кода, основанного на SHA-1. Этот алгоритм является алгоритмом односторонней аутентификации, который подразумевает аутентификацию клиента на стороне сервера [4]. Алгоритм впервые формально описан командой IETF в декабре 2005 года. Параметром, отвечающим за динамику генерации паролей, является факт генерации или само событие. Каждый раз при создании нового пароля счетчик событий увеличивает значение на единицу. Именно такое монотонное возрастающее значение используется как основной параметр алгоритма. Вторым параметром, отвечающим за генерацию одноразовых паролей, является симметричный ключ, которому необходимо быть уникальным для каждого клиента и в тоже время закрытым от всех, кроме сервера и самого клиента. Системы защиты, построенные с использованием НОТР, обладают высокой степенью надёжности. Они, в основном, устойчивы к широко распространённым криптографическим атакам.

ТОТР (Time-based One-Time Password Algorithm) – алгоритм создания одноразовых паролей для защищенной аутентификации (2008 г.) [5]. Он также, как и НОТР, является алгоритмом односторонней аутентификации, в котором сервер удостоверяется в подлинности клиента. В отличие от НОТР в нем параметром, отвечающим за динамику генерации паролей, является время. Обычно используется не определенное указание времени, а текущий интервал с установленными заранее границами. Концепция одноразовых паролей вкупе с современными криптографическими методами может использоваться для реализации надежных систем удаленной аутентификации. ТОТР достаточно устойчив к криптографическим атакам, однако вероятность взлома есть. Например, возможен такой вариант атаки как «человек посередине».

Алгоритм двухфакторной аутентификации

Рассмотрим пример предложенной системы защиты информации с использованием комбинации двух факторов: постоянного и временного паролей [6]. Постоянный пароль (первый фактор) выбирается пользователем и используется при регистрации аккаунта (учетная запись). Перед авторизацией необходимо пройти регистрацию в приложении. После этого запускается приложение для ввода данных пользователя (логина и пароля), которые должны соответствовать зарегистрированным данным. Затем перейти в приложение на смартфоне и ввести начальные данные для генерации временного пароля (второго фактора). Одноразовый или временный пароль генерируется на сервере по предложенному алгоритму [6] и действует в течение конкретного отрезка времени для одного сеанса аутентификации. Преимущество одноразового пароля заключается в том, что пароль не используется повторно. Таким образом, злоумышленник, перехвативший данные из успешной сессии и аутентификации, не может использовать скопированный пароль для получения доступа к защищаемой системе. Генерация временного пароля возможна в режиме онлайн [7]. Этот пароль генерируется на сервере и отображается пользователю в дополнительном программном обеспечении на смартфоне со сроком действия 20 секунд. Генерация временного пароля осуществляется на основе результата выбранной тригонометрической функции, которая имеет ряд переменных параметров.

В качестве входной строки для хэш-функции используется комбинация учетных данных пользователя, текущего момента времени по Гринвичу и

дополнительной секретной строки. Результат хэш-функции разделяется на отдельные числа, которые будут являться индексами для выбора функции и её начальными данными. Секретная строка является обязательным полем, которая будет выбираться из массива случайным образом. Секретная строка при каждом обращении меняется, что затрудняет вскрытие начальной входной строки и усиливает защиту.

Реализация алгоритма

Для реализации системы используется Node.js [8]. Эта программная платформа отвечает за написание серверной части на языке программирования JavaScript. Вместе с ней поставляется пакетный менеджер npm, который используется для установки различных библиотек и фреймворков. Дополнительно необходимо установить MongoDB [9] - документоориентированную систему управления базами данных открытым исходным кодом, не требующая описания схемы таблиц. При возникновении сложных запросов они обычно решаются на стороне приложения, что позволяет облегчить работу с данными и ссылками на них. Использование этой системы управления базами данных (СУБД) обусловлено тем, что в данную систему встроена достаточно простая масштабируемость с использованием технологии шардинга, которая производит разделение (партиционирование) базы данных на отдельные части так, чтобы каждую из них можно было вынести на отдельный сервер [9].

Преимуществом использования СУБД MongoDB является:

- увеличение скорости разработки;
- отсутствие необходимости синхронизировать схему в базе данных и приложении;
- понятен путь к масштабируемости;
- простота предписанных решений.

Для работы с Mongo на диске «С» создается папка «data», а в ней - папка «db».

Для установки пакетов, необходимых для запуска определенных частей программы («Server», «devschacht», «dir») нужно открыть консоль в папке с файлом package.json и в ней выполнить команду «npm install». После установки всех пакетов программа запускается на выполнение по частям. В начале запускается «Server». Для этого нужно зайти в папку, открыть консоль и выполнить команду «npm start». После запуска появится сообщение «сервер запущен». Далее запускается мобильное приложение через папку «devschacht», в которой открывается консоль и запускается команда «expo start -- android». После этого запуска появится QR код. Также в браузере откроется вкладка с системной информацией и QR кодом. Поле с IP адресом над QR кодом необходимо запомнить.

Для работы мобильного приложения необходимо считывание QR кода, которое можно произвести при помощи установленного сканера на мобильном устройстве по умолчанию либо установки дополнительного программного обеспечения «ЕХРО». Для запуска смартфона и персонального компьютера производится настройка локальной сети таким образом, чтобы они находились в одной сети. Далее происходит запуск web сайта, открывается консоль и выполняется команда «npm run dev». Для работы web приложения используется порт 4000 «localhost:4000». Для просмотра содержимого базы данных нужно установить MongoDB Compass и произвести подключение нажатием кнопки «connect».

Защита баз данных (БД) является важнейшей частью в хранении пользовательских данных. Такие результаты хакерских атак как утечка

персональных данных и неавторизированный доступ к информации являются следствием ее недостаточной защиты. Для определения методов защиты нужно определить объекты защиты. Такими объектами могут выступать значения определенных полей, таблиц или записей, сами записи в таблицах, отдельные таблицы и целые базы данных. Для защиты этих объектов можно использовать такие методы как представления, триггеры и встроенные функции шифрования данных:

1. Представлением является динамическая выборка данных из нескольких таблиц. Результат этого можно назвать виртуальной таблицей, формирование которой происходит в зависимости от пользовательского запроса, направленного доступному представлению (непонятно). В результате пользователь получает доступ только к совокупности столбцов и записей, которые определены используемым представлением. Таким образом, можно ограничить доступные данные и контролировать набор информации с доступом для редактирования. Исходя из этого, метод представления можно использовать для защиты конфиденциальности и целостности данных.

2. Триггер — это слушатель, который вызывается во время исполнения определенных событий. Такими событиями являются добавление, удаление или редактирование данных в БД. Используя триггеры можно определять пользовательский доступ для взаимодействия с данными в базе. Также они позволяют фиксировать все события, связанные с модификацией данных в защищаемой таблице. Таким образом, данный метод необходимо использовать для эффективного контроля целостности данных и слежения за всеми событиями, возникающими во время изменения информации.

3. Функции шифрования доступны далеко не во всех системах управления базами данных. В СУБД MySQL для шифрования данных используют стандарты AES и DES. Для хэширования используются алгоритмы MD5 и SHA1. Использование доступных функций шифрования значительно облегчает задачу защиты данных в БД.

Реализация предложенного метода защиты базы данных MongoDB, в которой хранятся данные для входа в систему, секретные слова, тригонометрические функции и временная информация для организации аутентификации, используется также метод шифрования данных.

Важное значение имеет также защита базы данных от SQL инъекций [10] (это распространенный способ взлома сайтов и приложений). Целью такой атаки является нарушение работы или получение данных без права доступа к ней. Внедрение инъекции, в зависимости от используемой системы управления базами данных и условий внедрения, может позволить атакующему выполнение произвольных запросов к базе данных. Атака типа внедрения SQL может быть возможна из-за некорректной обработки входных данных, используемых в SQL-запросах. Разработчики прикладных программ, работающих с базами данных, должны знать о таких уязвимостях и уметь предотвращать возможность внедрения SQL инъекций.

Используемая в реализуемом проекте MongoDB защищена от SQL инъекций, так как это не реляционная система управления базами данных стема управления базами данных документоориентированного типа и использует JSON (JavaScript Object Notation) -подобные документы и схему базы данных.

Выводы

Новизна предложенной системы заключается в разработке модифицированного алгоритма формирования секретного кода. В этом алгоритме формирование хэш-функции SHA256 проводится на основании таких входных данных как логин, текущий момент времени и секретное слово. Секретное слово не формируется пользователем, оно случайным образом подбирается из соответствующего массива. После формирования хэш-функции на основании ее значения определяются индексы тригонометрической функции из соответствующего блока и переменные, которые выбранная функция будет использовать для расчета результата. Из результата выполнения функции формируется временной код, который отображается только в мобильном приложении пользователя, привязанном к логину on-line системы. Для успешной авторизации в системе, пользователю необходимо ввести полученный код после попытки авторизации с корректными данными по запросу.

Реализуется данный алгоритм на стороне сервера. Со стороны клиента необходимо лишь взаимодействие с API приложения. Для реализации клиентской части приложения была выбрана платформа для разработки мобильных приложений – ReactNative. Она позволяет использовать ряд веб-технологий привычных для разработки веб-приложений и нативные компоненты для написания универсальных мобильных приложений.

Отметим, что новым является генерация и последующее обновление блоков тригонометрических функций и секретных слов с помощью генераторов, реализованных на стороне сервера.

Основное отличие реализуемого проекта от Google Authenticator заключается в алгоритме генерации временного пароля. Authenticator использует такие подходы, как TOTP и HOTP, но реализует алгоритмы, указанные в RFC 6238 и RFC 4226. Данное приложение генерирует 80-битный ключ, который кодируется с помощью Base32 в 16-, 26- и 32-значную строку. Основываясь на полученной строке, на стороне клиента формируется HMAC-SHA1, из которого извлекается преобразованный шестизначный код.

Отказ от использования SMS для обработки второго фактора заключается в небезопасности данного подхода. Сотовые сети используют «SignalSystem 7» для взаимодействия между собой, но в этой системе обнаружены серьезные уязвимости, позволяющие перехватывать входящие вызовы и SMS абонентов.

Литература

1. Концепция информационной безопасности Республики Казахстан: утв. 10 октября 2006 года № 199.

2. Юрьев Д.Р., Рогова О.С. Сравнительный анализ двухфакторной аутентификации // Технические науки - от теории к практике: сб. ст. по матер. LXXI междунар. науч.- практ. конф. - № 6(66). – Новосибирск: СибАК, 2017. – С. 46-51.

3. Часть 2. Передача реквизитов клиента OAuth. - <https://www.ibm.com/developerworks/ru/library/se-oauthjavapt2/index.html> (Дата обращения 13 мая 2019).

4. Национальная библиотека им.Н.Э.Баумана - <https://ru.bmstu.wiki/>. (Дата обращения 13 мая 2019).

5. Двухфакторная аутентификация с TOTP <https://medium.com/@n.moretto/two-factor-authentication-with-totp-ccc5f828b6df>. <https://ru.bmstu.wiki/>. (Дата обращения 13 мая 2019).

6. Нысанбаева С.Е., Усатова О.А. «Двухфакторная аутентификация в автоматизированной системе управления» //III Международная научная

конференция «Информатика и прикладная математика»- Алматы, 2018, том № 2, 448, С. 239-242.

7. S. Nyssanbayeva, W. Wojcik, O. Ussatova «Algorithm for generating temporary password based on the two- factor authentication model» // Przegląd Elektrotechniczny, Poland, № 5, 2019 г., ISSN 0033-2097, R. 95, С. 101 – 106.

8. O. Ussatova, S. Nyssanbayeva, W. Wojcik « Software implementation of two-factor authentication to ensure security when accessing an information system» // Вестник КазНУ им. аль-Фараби, март 2019г, 136, С.87-95.

9. Введение в MongoDB. - <https://mkdev.me/posts/vvedenie-v-mongodb>

10. SQL Injection от А до Я -<https://www.ptsecurity.com/upload/corporate/ru-ru/analytics/PT-devteev-Advanced-SQL-Injection.pdf>.

ИССЛЕДОВАНИЕ ПРОБЛЕМ ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ ЭТАПОВ ЖИЗНЕННОГО ЦИКЛА ИННОВАЦИЙ

^{1,2}Утепбергенов И.Т., ³Бобров Л.К., ^{1,2}Буранбаева А.И.,
¹Тойбаева Ш.Д., ¹Коржаспаев А.Е.

e-mail: i.utepbergenov@gmail.com, l.k.bobrov@edu.nsuem.ru,
Aburanbaeva@gmail.com, shara_t@mail.ru, arman.korzhaspayev@gmail.com

¹Институт информационных и вычислительных технологий МОН РК,

²Университет Туран, Казахстан

³Новосибирский государственный университет экономики и управления
«НИНХ», Россия

***Аннотация.** Разработана модель жизненного цикла инноваций в привязке к внешним информационным ресурсам, необходимым для успешного выполнения работ на его этапах, и показаны особенности информационного обеспечения инновационной деятельности этапов жизненного цикла инноваций. Для исследования проблем информационного обеспечения инновационной деятельности в Казахстане по предложенной модели проведен многоаспектный анкетный опрос инновационных и академических организаций и экспертов. Результаты статистического анализа подтверждают рабочую гипотезу о целесообразности и важности унифицированных источников данных о новейших исследованиях и разработках для отечественных организаций и производств.*

Введение

Президент Республики Казахстан Н. Назарбаев 31 января 2017 г. свое послание народу Казахстана [1] посвятил третьей модернизации государства и глобальной конкурентоспособности. Президент определил цель, согласно которой к 2050 году Казахстан должен войти в число 30 развитых государств мира. Поставленная цель предполагает дальнейшее развитие инновационной экономики как приоритетного направления промышленной политики Казахстана, и укрепление инфраструктуры поддержки инноваций для становления новых

индустрий с учетом современных тенденций цифровизации инновационного развития.

Одним из способов диверсификации национальной экономики и переориентации от сырьевого материала к модели промышленного обслуживания является развитие цифровых технологий. 12 декабря 2017 года Указом Правительства Республики Казахстан № 827 была утверждена Государственная программа «Цифровой Казахстан» [2].

В соответствии с этой программой необходимо принять меры для повышения качества существующей инфраструктуры инновационного развития, а ключевым направлением развития отрасли ИКТ является обеспечение роста доли услуг в области информационных технологий. Планируется, что программные мероприятия будут осуществляться по пяти направлениям, одним из которых является «Создание инновационной экосистемы». Таким образом, инновационное и цифровое развитие рассматриваются в органическом единстве, что создает благоприятные предпосылки для реализации проектов, направленных на информационную поддержку инновационной деятельности.

Приведенные ниже подзаголовки ежегодно издаваемого Всемирной организацией интеллектуальной собственности Глобального инновационного индекса наглядно свидетельствуют о важности инноваций и внимании, которое следует уделять различным аспектам инновационной деятельности:

- The Human Factor in Innovation [3];
- Effective Innovation Policies for Development [4];
- Winning with Global Innovation [5];
- Innovation Feeding the World [6].

Процессы информационного обеспечения и предоставления информационной поддержки инновационной деятельности предприятия включают в себя несколько пунктов, такие как:

- обновление и создание банков данных, научно-технических результатов и потенциальных возможностей, а также научно-технических программ и проектов, касающихся перспективных направлений;
- отбор и поиск инновационных проектов, рассмотрение предложений по производству наукоёмкой продукции для физических и юридических лиц, которые заинтересованы в их финансировании;
- реализация доступа к базам данных и удаленным информационным центрам, включая зарубежные, с использованием современных телекоммуникационных систем;
- предоставление доступа заинтересованных организаций и лиц к информационным базам данных, а также к информационным ресурсам Интернета в инновационной сфере деятельности.

Использование современных информационных технологий в интересах успешного ведения инновационной деятельности обеспечивает почти мгновенное подключение к любым электронным информационным массивам (таким как базы данных, электронные справочники и энциклопедии, различные оперативные сводки, аналитические обзоры, законодательные и нормативные акты и т.д.), поступающим из международных, региональных и национальных информационных систем.

1. Модель жизненного цикла инноваций

Исходя из представленной ранее сущности инновации и детализируя содержание, направленность, а также приоритеты работ, приведем модель

жизненного цикла инноваций в привязке к внешним информационным ресурсам, необходимым для успешного выполнения работ на следующих этапах:

- 1) генерация и фильтрация идей;
- 2) проведение НИР;
- 3) проведение ОКР и опытное производство;
- 4) промышленное производство;
- 5) вывод на рынок, рост, насыщение, спад.

Данная модель информационной поддержки жизненного цикла инноваций предполагает комплексное использование документальных и фактографических источников (на что справедливо обращали внимание авторы работ и ряда других работ), отражающих следующие основные тематические и видовые аспекты:

- информацию о рынках;
- сведения о программах и направлениях исследований по практическому использованию результатов фундаментальных НИР;
- маркетинговую информацию;
- информацию о потенциальных инвесторах и конкурсах грантов;
- научно-техническую и патентную информацию, отчеты о НИР и диссертации;
- материалы научных конференций;
- прогнозно-аналитическую информацию;
- конъюнктурно-коммерческую информацию;
- правовые и нормативные акты, стандарты и справочники;
- экономическую информацию (включая менеджмент), и др.

Следует подчеркнуть, что для стадий генерации инновационной идеи и инициации бизнес-идеи свойственны: высокая неопределенность целей и задач, технологическая, семантическая и организационная разноформатность, а в ряде случаев и территориальная рассредоточенность задействованных в них ключевых участников и ресурсов инновационной деятельности.

Следовательно, информационно-аналитическое обеспечение инновации востребовано не только, как общепринято, для стадий маркетинга, производства, реализации и послепродажного сопровождения, но и для начальных стадий (генерации/инициации, НИОКР) их ЖЦ. Здесь для информационно-аналитического обеспечения начальных стадий ЖЦ инноваций требуется системный подход к исполнению бизнес-процесса создания инновационного продукта, учитывающий перспективы трансфера, коммерциализации, тиражирования и расширения рынка сбыта новшества [7].

1. На стадии генерации инновационной идеи и инициации бизнес-идеи основополагающей является внешняя информация, содержащая в результатах аналитических и маркетинговых исследований, источником которых, как правило, является внешнее окружение высокотехнологичного предприятия. Следует подчеркнуть, что сегодня большая часть новых инновационных решений становятся доступными профильным специалистам, в том числе, благодаря интеграции автономных информационных систем. В связи с этим в настоящее время прослеживается тенденция, заключающаяся в интеграции разобщенных информационных ресурсов и создании единых информационных пространств как на уровне предприятия, так и на мезо-, макро-, мега уровне.

Единое информационное пространство представляет собой совокупность баз и банков данных, технологий их ведения и использования, информационно-телекоммуникационных систем и сетей, функционирующих на основе единых принципов и по общим правилам, обеспечивающим информационное

взаимодействие организаций и граждан, а также удовлетворение их информационных потребностей, группировке информационных ресурсов в кластеры информации, данных, знаний и компетенций.

2. В свою очередь, на стадии фундаментальных НИР, целью которых является раскрытие новых связей между явлениями в конкретной предметной области, познание закономерностей в протекающих естественных и искусственных процессах в части их адресного использования, требуется извлечение из баз данных информации по данной проблематике и последующая ее систематизация. Для этой стадии жизненного цикла базы данных должны содержать:

- данные о результатах фундаментальных исследований в данной предметной области в части новых знаний, которые, в последующем могут быть основаниями для прикладных НИР;

- данные о полезных идеях, моделях, полученных в результате фундаментальных исследований в предметной области и в смежных областях.

Для стадии прикладных исследований предусмотрено использование новых знаний, ориентированных на достижение практико-ориентированных целей и решение конкретных задач, связанных, в том числе, с созданием моделей, изготовлением макетов, прототипов, пилотных образцов новой техники, перспективных новых технологий. На этой стадии выполняются специальные расчеты в целях анализа, оценки и возможной корректировки проводимых исследований, а при необходимости исключения бесперспективных идей. Итогом является разработка технического задания на проектирование искомого натурального объекта или образца изделия, технологические рекомендации, методики.

Стадия прикладных исследований требует извлечения из информационных ресурсов (например, баз данных) информации:

- о результатах поисковых и прикладных исследований (публикации), касающихся конкретной предметной области, в целях обоснования сгенерированных идей, замыслов, бизнес-моделей и др.;

- о практиках по трансформации результаты прикладных исследований в новые образцы коммерческих продуктов или их прототипы, модели, макеты, экспериментальные образцы (экспресс-информация, проспекты и др.);

- о процедурах и результатах защиты исследований и разработок в смежных областях (нормативно-правовые документы, методические рекомендации по составлению заявок на объекты интеллектуальной собственности, патенты, свидетельства о государственной программе для ЭВМ и др.).

3. Поскольку целью опытно-конструкторских разработок является разработка технической документации (технический, эскизный, рабочий проект, технологические карты и др.), а также предварительные, рабочие и серийные испытания образцов новых изделий для подтверждения их надежности, работоспособности, производственного и коммерческого использования, то для эффективного обеспечения данной стадии требуется извлечение из информационных ресурсов (например, баз данных) информации:

- о подготовке и выходе на рынок новых инженерно-технических решений (конструктивных, технологических и др.), технологий, продуктов (например, ИТ-решений), услуг;

- о системах управления инженерными данными;

- об уровне защищенности аналогичных инженерно-технических решениях и их правообладателях;

- об уже используемых схемно-компоновочных и других проектных решениях;

– о составных частях (деталях, узлах, комплектующих), технологиях их изготовления, сборки, монтажа, экспериментов, испытаний, материалах и других элементах, условиях их поставки, используемых при создании и использовании разрабатываемой продукции (изделия, технологии, продукта, услуги и др.);

– о возможных контрагентах и субподрядчиках, их компетенциях как на стадии ОКР, так и на последующих стадиях создания инновационного решения;

– о потенциальных источниках финансирования (бюджетные ресурсы различного уровня, гранты, собственные средства предприятия, финансовые средства участников команды инноваторов, кредитные ресурсы, венчуры, бизнес-ангелы и др.) и условиях их предоставления на различных стадиях ОКР.

4. Особое место занимает информационное обеспечение на стадии тиражирования инновации как инновационного продукта в рамках промышленного непрерывного производства (единичное, серийное, массовое). Здесь под промышленным производством понимается процесс, в ходе которого сырье, основные материалы или полуфабрикаты превращаются с использованием промышленного оборудования в инновационный продукт, востребованный на рынке.

На указанной стадии востребована информация об анализе и оценке рыночных перспектив нового инновационного решения, соответствия его действующим отечественным и международным стандартам, финансовых возможностей, информация о практиках проектирования и разработки технологических, организационных и управленческих процессов, подготовки всех компонентов производства высокотехнологичного продукта.

5. Ключевым условием эффективности информационного обеспечения управления инновационными бизнес-процессами предприятия (вывод на рынок, рост, насыщение, спад) является превышение скорости обобщения и систематизации информации над скоростью осуществления стадий ЖЦ инноваций, что достигается не столько совершенствованием механизмов и процессов обмена информацией между субъектами инновационных процессов, сколько методов преобразования информации в знания или интеллектуальный капитал инновационного предприятия.

Как уже было показано, каждая из стадий ЖЦ инноваций требует конкретного набора информации: научно-технической (документы и отчетности других процессов поиска в данной предметной или смежных предметных областях), патентной, коммерческой, маркетинговой, статистической, демографической, экологической, правовой, информации о конкурентах, информации о потенциальных потребителях инноваций.

2. Проведение многоаспектного анкетного опроса инновационных и академических организаций, а также интервьюирования экспертов в РК

Достижение поставленной цели данного этапа проекта предполагает сбор и статистическую обработку данных о работе инновационных предприятий с использованием разработанной методики формирования ранжированного реестра проблем информационного обеспечения инноваций в привязке к отдельным этапам их жизненного цикла. При организации любого статистического наблюдения необходимо решить ряд вопросов, объединенных в две группы: программно-методические и организационные вопросы наблюдения.

Программно-методологические вопросы.

1) Цель наблюдения – результат исследования, для которого проводится сбор данных.

- 2) Объект исследования – исследуемая статистическая совокупность.
- 3) Предмет исследования – единица исследуемой статистической совокупности.
- 4) Программа наблюдения – перечень вопросов, по которым будут собираться сведения, формуляр анкеты.

Путем изучения Интернет-ресурсов нами были выявлены 138 предприятий и организаций, ведущие инновационную деятельность в различных отраслях экономики Республики Казахстан. Часть из них входит в ассоциацию СЭЗ ПИТ «Алатау», для которых проводилось анкетирование отдельно.

На основе проведенного анализа вышеприведенных статистических форм Казахстана и России была разработана многоаспектная анкета для анкетирования инновационных предприятий (Приложение Д). Указанная анкета, включает 51 вопрос для выявления исследовательской специфики информационных потребностей новаторов на каждом этапе жизненного цикла инноваций от генерации идей до вывода продукта с рынка.

Организационные вопросы.

- 1) Субъект наблюдения – организация, которая проводит исследование.
- 2) Место наблюдения.
- 3) Сроки наблюдения – период в течение которого получают сведения.
- 4) Форма наблюдения.
- 5) Вид наблюдения.
- 6) Способ наблюдения.

В соответствии с указанным перечнем субъектом наблюдения стал Институт информационных и вычислительных технологий МОН РК, являющаяся организацией-исполнителем проекта. Местом наблюдения были дислокации 138 предприятий и организаций, ведущие инновационную деятельность в различных отраслях экономики Республики Казахстан [8].

Наблюдение было проведено в период с января по март 2018 года в соответствии с календарным план-графиком проекта. Интервьюирование экспертов проводилось в письменной (очное) и электронной формах (бесконтактное), способ наблюдения – анкетирование.

3. Статистическая обработка полученной информации для выявления наиболее значимых проблем информационного обеспечения

В Казахстане 36% инновационных предприятий, принявших участие в опросе, – малые предприятия (до 100 человек) и 30% – вузы. Кроме малых предприятий, активное участие в опросе приняли крупные предприятия, научно-исследовательские институты и консалтинговые центры.

Опрос показал, что наибольшее количество организаций оказывает инновационные услуги (57%), на втором месте по количеству ответов – производство инновационной продукции (20%), 11% занимаются доведением продукции до товарного образца, 7% осуществляют прототипирование.

Виды работ, осуществляемых инновационными организациями, иллюстрирует рисунок 1.



Рис. 1 Виды работ, выполняемых инновационными организациями

Немаловажное значение для анализа инновационной деятельности в обеих странах имеет информация об источниках финансирования. В Казахстане в качестве основных источников финансирования инновационной деятельности названы собственные средства организаций (41%), средства бюджетов различных уровней бюджетной системы РК (28%) и средства, привлекаемые из внебюджетных источников (19%).

Результаты анкетирования показали, что дефицит информации возникает в первую очередь на этапе вывода продукта на рынок, проработки идеи инновационного продукта и создании опытного образца продукта (Рисунок 2).

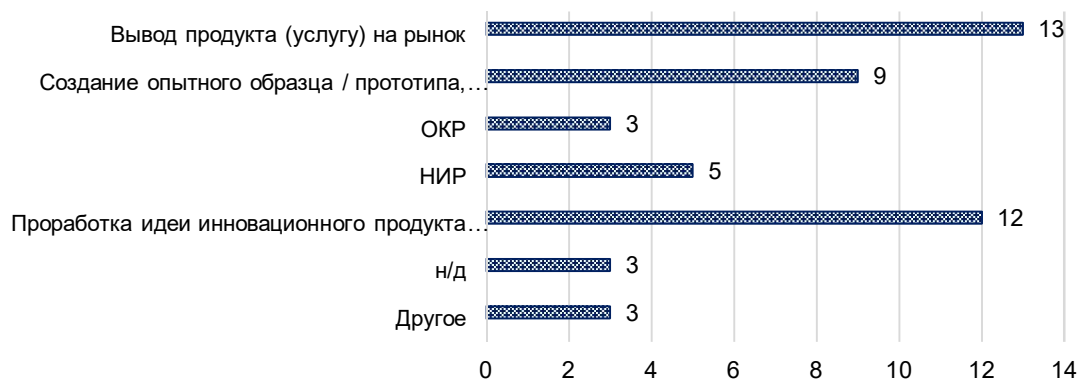


Рис. 2 Ответ на вопрос «На какой стадии Вы ощущаете дефицит информации?»

Согласно полученным данным, недостаток информации в Казахстане ощущается в таких видах деятельности, как получение информации о различных методах, моделях, технологиях; выявлении существующих точек зрения на проблему и сборе фактов (рисунок 3).

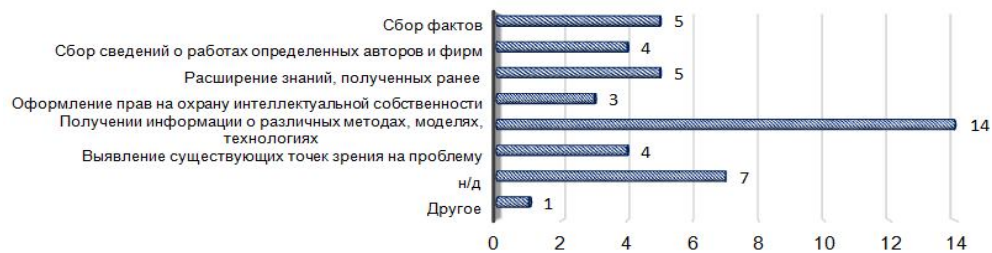


Рис. 3 Результат ответов на вопрос «В каких видах деятельности ощущается, по Вашему мнению, недостаток информации?»

При этом в организациях преимущественно возникает дефицит информации маркетингового характера, информации о новых технологиях и новых инженерно-технических решениях (рисунок 4).

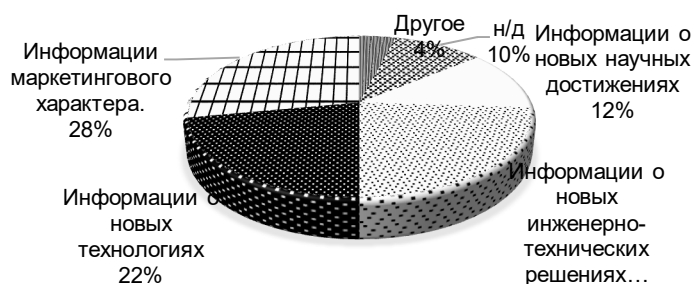


Рис. 4 Ответ на вопрос «Дефицит информации какого характера Вы (сотрудники Вашей организации) испытываете?»

Следует особо отметить, что 82% опрошенных считают информационное обеспечение профессиональной деятельности очень важным (рисунок 5).

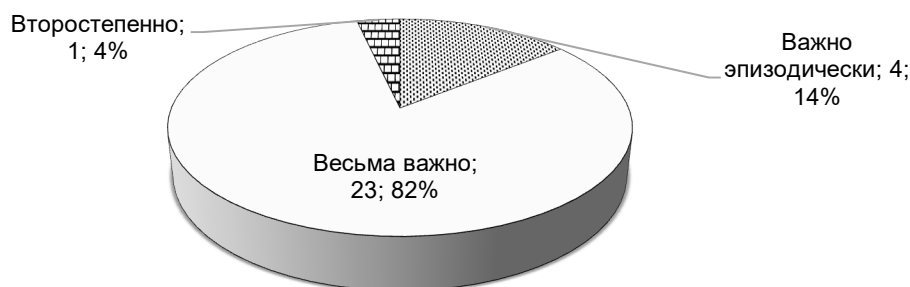


Рис. 5 Ответ на вопрос «Насколько важно в Вашей профессиональной деятельности информационное обеспечение?»

При этом около трети (32%) опрошенных организаций, указавших в качестве источника финансирования инновационной деятельности, в том числе и собственные средства (или только собственные средства), готовы закладывать в бюджет расходы на информационное обеспечение работ. Это свидетельствует о жизнеспособности настоящего проекта по окончании его финансирования.

Полученные в анкетах ответы были обработаны в аналитических программных средствах сервиса «Google Form» и «IBM SPSS statistics 22» с

использованием описательных (рисунок 6), регрессионных (рисунки 7, 8), а также визуальных моделей анализа (рисунок 9).

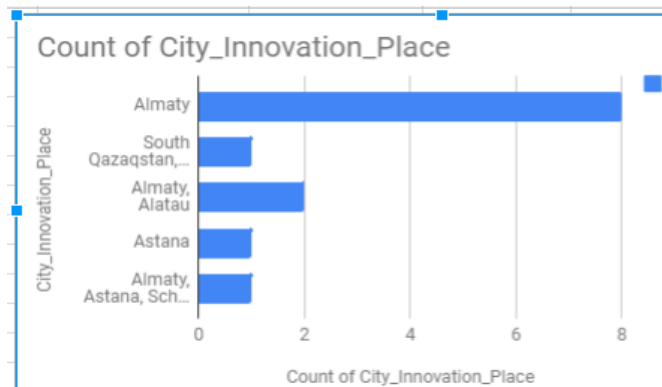


Рис. 6 Частотное географическое распределение организаций

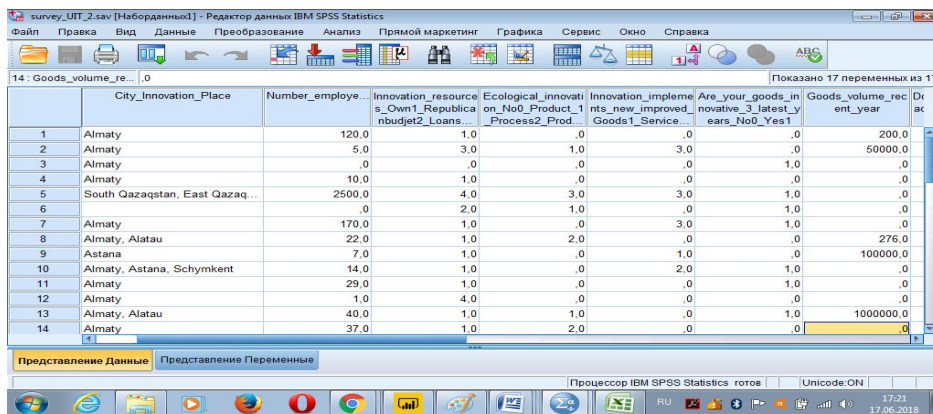


Рис. 7 Данные результатов опроса для обработки в программной среде IBM SPSS statistics 22

Кoeffициенты^a

Модель	Нестандартизованные коэффициенты		Стандартизованные коэффициенты	t	Знач.
	B	Стандартная Ошибка	Бета		
1 (Константа)	-1,457	2,138		-,682	,511
Interested_receiving_new_info_regularly_No-0_Yes_1	-,182	,782	-,066	-,233	,820
Importance_info_provision_No-0_Yes_1_Very-important_2	,953	1,015	,266	,940	,370
Importance_payment_to_other_organizations_for_info_provision_No-0_If-very-needed-1_If-there-are-enough-sources-2_Yes-3	,457	,317	,394	1,441	,180

a. Зависимая переменная: Innovation_implements_new_improved_Goods-1_Services-2_Both-3

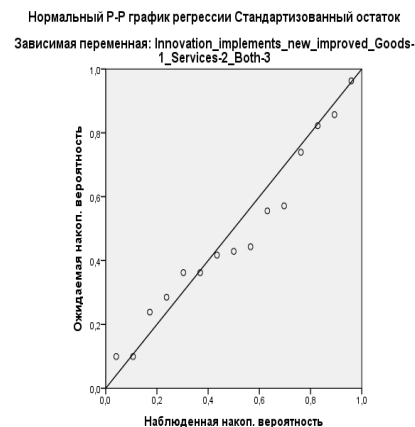


Рис. 8 Результаты регрессионного анализа

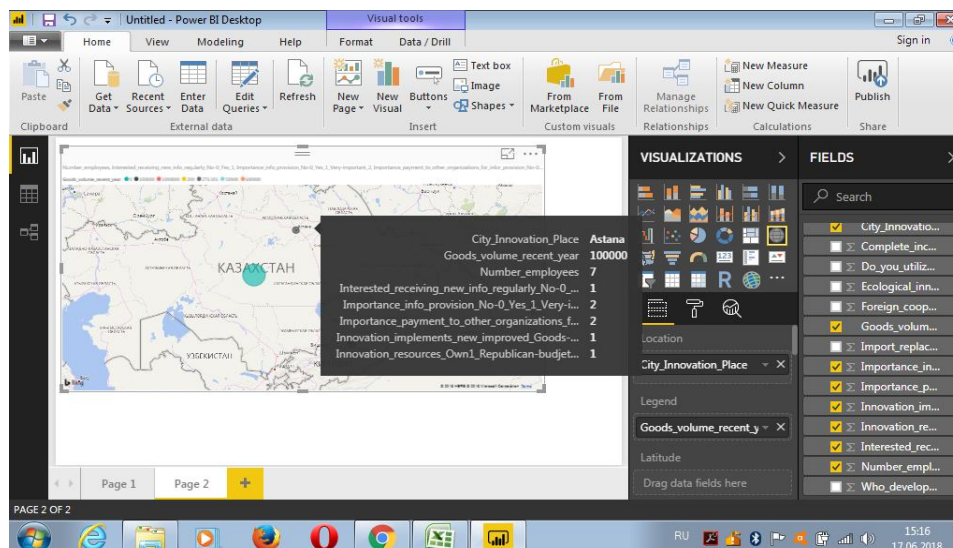


Рис. 9 Визуальное представление суммарной статистики в среде Microsoft Power BI

Из результатов опроса можно сделать предварительные выводы о целесообразности и важности унифицированных источников данных о новейших существующих результатах исследований и разработок для отечественных организаций и производств с учетом местных специфик.

Визуализация географии приложения наглядно демонстрирует, несомненно, более высокую вовлеченность в инновационную деятельность в более густонаселенных областях Республики Казахстан.

Выводы

Разработана модель жизненного цикла инноваций в привязке к внешним информационным ресурсам, необходимым для успешного выполнения работ на его этапах, и особенности информационного обеспечения инновационной деятельности этапов жизненного цикла инноваций.

Проведен многоаспектный анкетный опрос и осуществлен сбор заполненных анкет от 36 (тридцати шести) инновационных и академических организаций и экспертов в РК выявление и сбор данных для исследования проблем информационного обеспечения инновационной деятельности.

Результаты анализа средствами сервиса Google Form и в среде IBM SPSS подтверждают рабочую гипотезу о целесообразности и важности унифицированных источников данных о новейших исследованиях и разработках для отечественных организаций и производств.

Работа поддержана грантом МОН РК (проект № AP05134019 «Разработка научно-методических основ и прикладных аспектов построения распределенной системы информационного обеспечения инновационной деятельности с учетом специфических особенностей каждого из этапов жизненного цикла инноваций»).

Литература

1 Назарбаев, Н. (2017) Третья модернизация Казахстана: глобальная конкурентоспособность: послание Президента Республики Казахстан Н. Назарбаева народу Казахстана. 31 января 2017 г. // http://www.akorda.kz/ru/addresses/addresses_of_president/poslanie-prezidenta-respubliki-kazakhstan-nazarbaeva-narodu-kazahstana-31-yanvarya-2017-g: 29.05.18

2. Государственная программа «Цифровой Казахстан» // <http://mic.gov.kz/ru/pages/gosudarstvennaya-programma-cifrovoy-kazahstan>: 29.05.18.
3. The global innovation index 2014: the human factor in innovation // Soumitra Dutta, Bruno Lanvin, and Sacha Wunsch-Vincent, editors. – Geneva: WIPO, 2014 – XXV: - 400 p.
4. The global innovation index 2015: effective innovation policies for development // Soumitra Dutta, Bruno Lanvin, and Sacha Wunsch-Vincent, editors. – Geneva: WIPO, 2015 – XXXI, 418 p.
5. The global innovation index 2016: winning with global innovation // Soumitra Dutta, Bruno Lanvin, and Sacha Wunsch-Vincent, editors. – Geneva: WIPO, 2016 – 422 p.
6. The global innovation index 2017: innovation Feeding the World // Soumitra Dutta, Bruno Lanvin, and Sacha Wunsch-Vincent, editors. – Geneva: WIPO, 2017 – 463 p.
7. Утепбергенов И.Т., Тойбаева Ш.Д., Буранбаева И.Т., Коржаспаев А.Е., Подход статистическому исследованию проблем информационного обеспечения инновационной деятельности предприятий в Казахстане // Матер. научн. конф. ИИВТ МОН РК «Современные проблемы информатики и вычислительных технологий». Алматы, 2018. – С. 298-306.
8. Лаборатория инновационных SMART технологий // <http://innovationlab.iict.kz/>: 14.08.2018.

ОСНОВНЫЕ ТРЕБОВАНИЯ К ХРАНИЛИЩУ ДАННЫХ ПУТЕЙ ПЕРЕДВИЖЕНИЯ

Утепбергенов И.Т., Яворский В.В., Касымова Д.Т.

e-mail: i.utepbergenov@gmail.com

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** В статье рассматриваются основные требования к хранилищу данных путей передвижения. Проанализированы существующие методы по управлению работой ГТС, информированию пассажиров общественного транспорта, модули интегрированной дорожно-транспортной среды.*

Очевидна важность проблем совершенствования планирования и управления городскими транспортными системами (ГТС), которая определяется тем, что от её качественного решения на соответствующем научно-техническом уровне зависит эффективность функционирования всех городских систем производства и обслуживания. Сложность данной проблемы обуславливается многофункциональным и иерархическим характером транспортных систем, активным и стохастическим поведением транспортных процессов, многообразием функционирования и взаимодействия отдельных элементов транспортных средств.

Совершенствование функционирования такой сложной и многогранной системы, какой является современный город также, как и для других организаций связано с совершенствованием технологий принятия решений. Эти технологии для

организационных систем реализуют процессы управления. Чтобы принимать обоснованные решения обрабатывают колоссальные объемы информации.

Для систем, реализующих процессы обработки информации и подготовки решения, важна реализация трех принципов. Во-первых, собираемая, хранимая и используемая информация должна быть достоверной. Второй принцип связан с актуальностью. Необходимые управленческие воздействия тем эффективнее, чем раньше субъекты управления видят проблемную ситуацию: «проблему надо не упускать, а прогнозировать». Наконец третий принцип связан с процедурой (алгоритмом) формирования обоснованного решения. Очевидно, что при этом должна достигаться максимальная полезность результата с точки зрения всей системы. В условиях максимальной цифровизации процессов в экономике и обществе, которая является стратегической задачей, актуальность указанных принципов еще более возрастает.

Комплекс мероприятий по снижению напряженности организации перевозок в транспортной системе города представляет собой ряд сложных взаимосвязанных задач, ориентированных на приоритетное и эффективное развитие городского общественного транспорта.

Прежде всего, необходимо структурировать сбор и хранение данных. Это требует создания современного хранилища данных для системы управления городским общественным транспортом (ХГОТ), со средствами телекоммуникаций для эффективного использования информации. Для эффективного решения задач управления необходимы мощные средства моделирования процессов передвижения населения в городе; моделирования процессов обслуживания пассажиров и движения транспортных единиц.

Также требуется провести мероприятия по модернизации и совершенствованию инфраструктуры, внедрить автоматизированную систему управления и систему электронных платежей, реализовать комплекс мероприятий по совершенствованию организации движения и формированию эффективного расписания движения.

Очевидно также, что требуется совершенствование управления дорожного движения в крупных городах. Наряду с традиционными мерами регулирования и контроля в современной ситуации для создания интегрированной системы в данном случае также необходимо создание хранилища данных организации дорожного движения (ХОДД), оснащенного мощной коммуникационной системой и средствами моделирования движения транспортных потоков и их рациональной организации. В современных условиях также неизбежны организационные мероприятия, связанные с организацией и управлением парковками. Здесь тоже необходимо создавать систему, которая включает платные парковки, перехватывающие и внеуличные парковки. Управление парковочным пространством города должно быть обеспечено на цифровом уровне.

Весьма эффективным является вывод грузообразующих предприятий за черту города, а также запрет движения грузового транспорта и индивидуального транспорта по определенным направлениям.

Как показывает мировой опыт, наиболее эффективна интеграция ХГОТ, ХОДД, систем автоматизированного управления организацией дорожного движения и общественного транспорта в рамках единой интеллектуальной транспортной системы города (ИТС) с единым хранилищем данных путей передвижения (ХДПП).

Наиболее успешно эти задачи решены в Торонто, Берлине и Турине. В данном случае ИТС интегрирована в органы власти, а основными пользователями выступают (рис. 1) [20]:

- муниципальные органы;
- органы МВД и особенно ГИБДД и комитет по ЧС;
- Скорая медицинская помощь и другие службы экстренной помощи;
- строительные и дорожные подрядные организации;
- всевозможные пользователи транспортных коммуникаций и дорог;
- средства коммуникаций и массовой информации.

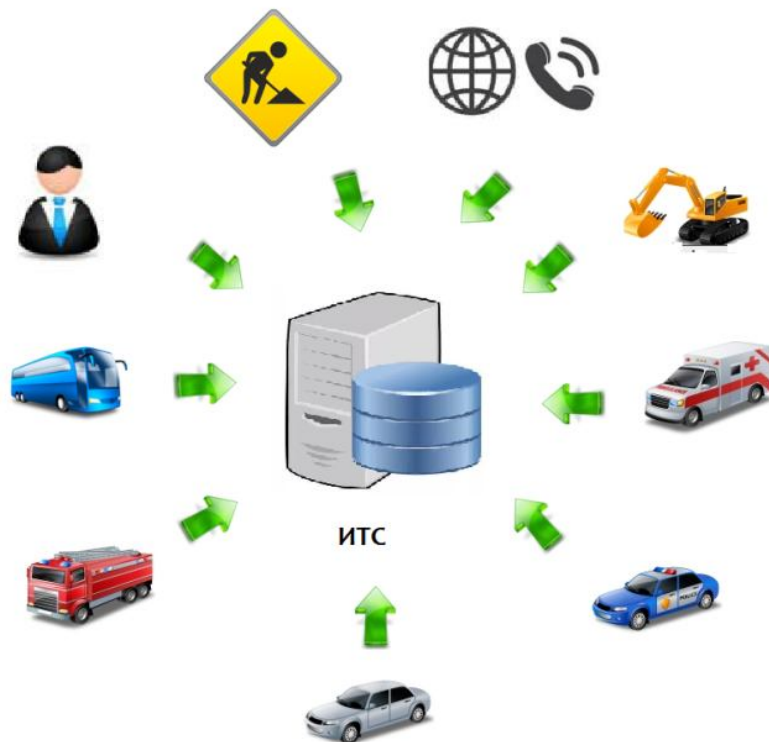


Рис. 1. Основные пользователи ИТС

Основными сервисами ИТС являются следующие (рис.2):

- управление городским транспортом;
- управление движением;
- обеспечение наблюдения и безопасности;
- электронные платежи;
- управление парковками;
- информационное табло и мобильное приложение.



Рис. 2. Основные сервисы ИТС

Концепция ХД предполагает не просто единый логический взгляд на данные, а реализацию единого интегрированного источника данных.

ХДПП строятся на основе многомерной модели данных, которая подразумевает выделение отдельных измерений и фактов, анализируемых по выбранным измерениям. Прежде всего, это данные о маршрутах, в частности координаты остановок и интенсивности потоков входящих и выходящих пассажиров, данные о дорожной сети. Многомерная модель данных физически может быть реализована как в многомерных СУБД, так и в реляционных.

При реализации проектов по построению хранилищ данных возникает ряд общих задач, к числу которых можно отнести:

- проектирование структуры иерархических измерений;
- проектирование структуры медленно меняющихся измерений;
- проектирование и актуализация агрегатных значений.

Основные требования к ХДПП:

- предметная ориентированность - все данные собираются, очищаются, согласовываются, дополняются, агрегируются и представляются в единой, удобной для их использования в анализе форме;

- интегрированность - все данные взаимно согласованы и хранятся в едином ХДПП;

- неизменчивость - исходные (исторические) данные, после того как они были согласованы, верифицированы и внесены в ХДПП, остаются неизменными и используются исключительно в режиме чтения;

- поддержка хронологии - данные хронологически структурированы и отражают историю, за достаточный, для выполнения задач анализа и прогнозирования, период времени.

Математический аппарат формализованного описания транспортной системы подробно рассматривается в ряде таких работ как [7, 10, 11]. В рамках исследований по проекту АР05133699 «Исследование и разработка инновационно-телекоммуникационных технологий с использованием современных кибер-технических средств для интеллектуальной транспортной системы города» института ИВТ КН МОН РК были предложены собственные модели [21, 22]. Транспортная сеть, маршрутная сеть, корреспонденции между пунктами на транспортной сети, потоки на маршрутной сети, пути передвижения между пунктами транспортной системы предлагается представлять гиперсетью, дуги которой являются путями передвижения. Введено точное понятие пути передвижения между корреспондирующими пунктами на городском общественном транспорте (ГОТ) [21]. Описана рекуррентная схема получения путей передвижения между вершинами гиперсети любого порядка.

В мире существует определенный опыт по комплексному внедрению современных технологий управления и электронных приложений в системы управления городским транспортом [12, 20]. В частности, в городах Греции, Великобритании, Германии, Италии, Испании и Швеции реализуется проект Комиссии ЕС QUARTET и QUARTET Plus, который предусматривает разработку приложений по управлению работой ГТС, информированию пассажиров общественного транспорта, модули интегрированной дорожно-транспортной среды, а также приложения по интермодальной маршрутизации, аварийной связи, и экологическому управлению дорожным движением.

Большая программа развития технологий управления дорожным движением, реализованная в городах Берлине и Кельне. Важнейшей составной частью данной программы является внедрение электронных систем управления дорожным движением, управления паркингами, а также обеспечения информирования участников движения с целью решения задач приоритетного использования городского общественного транспорта (ГОТ) (с использованием систем “park&ride”); стимулирования смещения пикового периода на время слабой загрузки улично-дорожной сети (УДС), обеспечения информации о текущей транспортной ситуации в городе и наличии свободных мест для парковки автомобилей, снижению “лишней” интенсивности движения, связанной с поиском места для парковки.

Широко известна система SCOOT динамического светофорного регулирования, которая внедрена в городах Торонто (Канада), Сантьяго (Чили), Бейинг (Китай), Уорчестер (Великобритания), Лондон (Великобритания), Саутгемптон (Великобритания), Ковентри (Великобритания), Сан-Пауло (Бразилия), Наймеген (Нидерланды). Исследование эффективности работы приложений SCOOT различных версий показало, что внедрение системы позволяет сократить продолжительность поездки в среднем на 7,2% и снизить задержки транспорта в среднем на 23,6%. В Лондоне существующая система SCOOT была дополнена функцией дистанционного ограничения доступа транспортных потоков в район перегрузки УДС. Ощутимые результаты от внедрения ИТС в Европе получены в Барселоне и Турине. Внедренные здесь системы организации и управления движением, а также парковочным пространством позволило снизить время в пути пассажиров на 15%.

В целом, как показывает опыт внедрения ИТС основными объектами разработки и внедрения являются следующие:

- телекоммуникационная система формирования и обработки данных;
- хранилище данных и система управления доступом;

- система контроля и обеспечения безопасности;
- система управления городским общественным транспортом;
- система электронных платежей.

Основной эффект достигается за счет интеграции существующих систем управления. При этом ИТС позволяет интегрировать следующие системы и их функции:

- управление городским общественным транспортом прежде всего реализует функции автоматизированного диспетчерского управления и электронные платежи;
- обработка данных о всей ГТС и СОД на базе технологий Big Data и систем искусственного интеллекта;
- обеспечение контроля движения транспорта: выявление опасных ситуаций и заторов; управление полосами и приоритетом; управление движением на съездах/въездах, перекрестках, перегонах УДС; управление парковкой автотранспорта движения;
- обеспечение приема электронных платежей в рамках всей ГТС;
- формирование, обеспечение актуальности и использования хранилищ данных.

В целом схема взаимодействия обеспечивающих подсистем интеллектуальной транспортной системы и подсистем управления ГТС представлена на рисунке 3.

В настоящее время загруженность дорог в городах Казахстана растет очень быстро. Крайние полосы УДС часто используются под парковку. Таким образом, дорога превращается из четырехполосной в трехполосную, из трехполосной – в двухполосную и т.д. В результате даже при неизменном пассажиропотоке на всех видах транспорта имеет место рост загрузки и снижение скорости, как следствие - пробки. При этом надо иметь в виду, что пассажиропотоки в городах постоянно растут.

Первоочередными задачами построения ИТС являются задачи цифрового описания и анализа ГТС:

- 1) создание цифрового хранилища данных о дорожной ситуации;
- 2) создание цифрового хранилища данных о городском общественном транспорте и передвижениях населения;
- 3) разработка информационной цифровой модели транспортной системы города;
- 4) разработка концепции ИТС;
- 5) разработка с использованием информационной цифровой транспортной модели вариантов оптимизации системы управления движением в рамках существующей инфраструктуры.

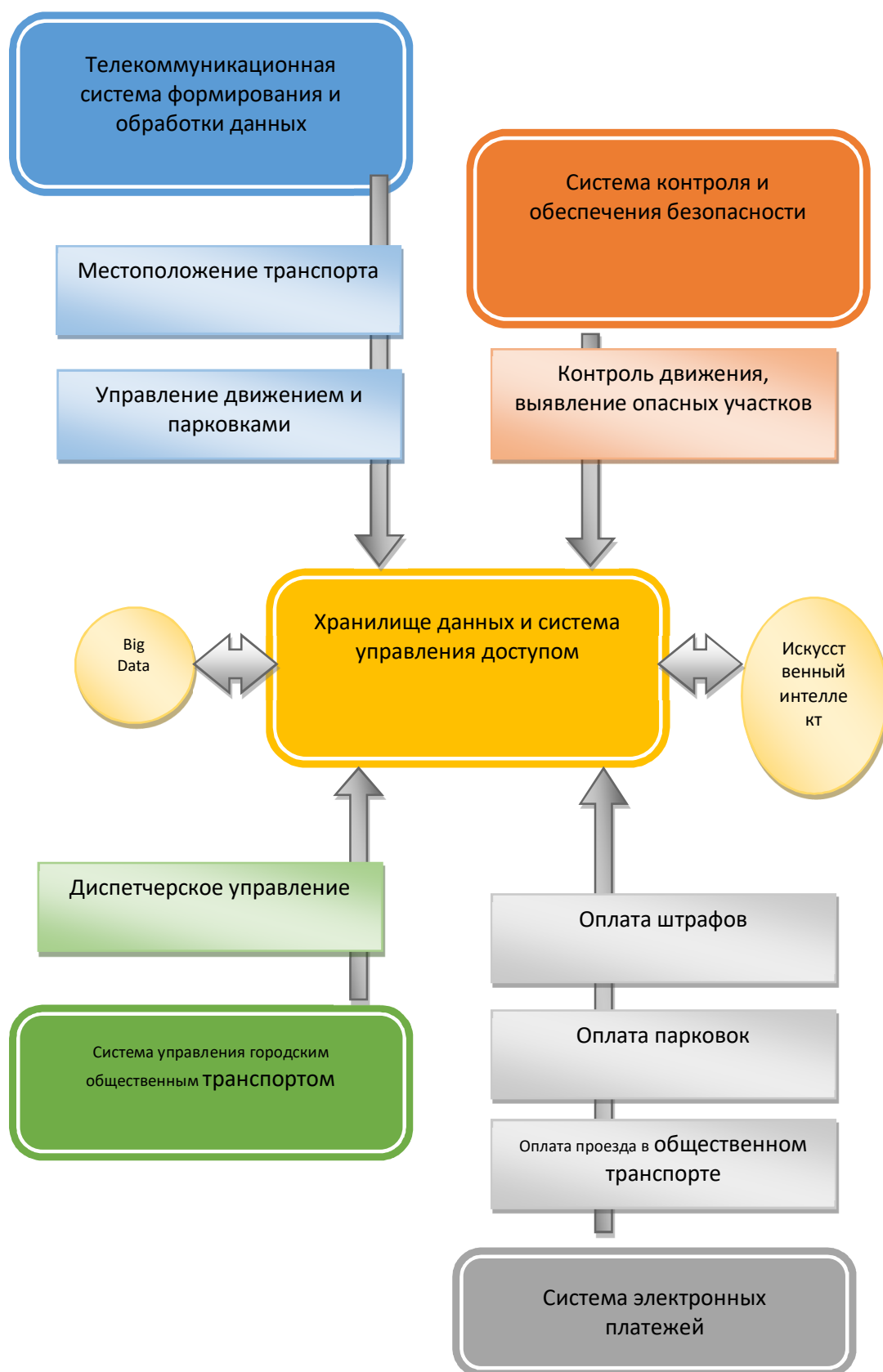


Рис. 3. Схема взаимодействия подсистем ИТС

Можно выделить семь задач решения рассматриваемой проблемы систем организации движения и создания ИТС на предварительном этапе [21].

1. Прежде всего необходимо собрать и оценить имеющуюся информацию о движении транспорта и системе организации движения, способах получения данных, а затем разработать предложения о необходимых обследованиях и системах сбора информации. Определить форматы хранения и принципиальную структуру хранилища данных СОД.

2. После решения первой задачи можно непосредственно приступить к формированию хранилища данных СОД: о движении транспорта, дорожной ситуации, а также запускать в эксплуатацию интерфейсы системы СОД.

3. Третьей задачей является ревизия городского общественного транспорта (ГОТ). Необходимо собрать и оценить имеющуюся информацию о ГОТ, городской территории и пассажирообразующих объектах, пассажиропотоках на маршрутах и корреспонденциях; о способах получения данных и затем, разработать предложения о необходимых обследованиях и системах сбора информации. Определить форматы хранения и принципиальной структуры хранилища данных ГОТ.

4. После нормативного определения способов хранения информации о ГОТ можно приступить к наполнению хранилища данных ГОТ о корреспонденциях и потоках, о движении транспорта, а также запускать в эксплуатацию интерфейсы системы ГОТ.

5. После решения указанных четырех задач необходимо формировать цифровую модель городской транспортной системы города. Разработку информационной цифровой модели города необходимо учитывать специфику передвижения в городах Казахстана и использовать имеющиеся проверенные платформы моделирования компоненты PTV-VISION программы VISUM и VISSIM, а также платформы AUSIAS, TRANUS, SATURN [21]. Запуск в эксплуатацию систем моделирования и оценки систем организации движения является сложной творческой задачей.

6. Получение информации о существующей ГТС и формирование моделей ее развития на новой цифровой основе реально позволит перейти к разработке концепции ИТС, ее структуры, средств и способов реализации, проверке и оценке их работоспособности и эффективности.

7. Собранную информацию, материалы и результаты моделирования необходимо максимально использовать в разработке предложений вариантов оптимизации существующей системы управления ГТС на первом этапе без затратной корректировки улично-дорожной сети.

Система ХДПП создается с целью анализа существующей транспортной системы города, ее корректировки и анализа существующих и синтеза новых транспортных маршрутов. Для анализа требуются данные о транспортной системе города: остановки, дорожная сеть, маршруты, возможные пешие и транспортные пути, а также возможные передвижения по городу с пересадками. Также для получения МПК необходимо иметь данные о пассажиропотоках по городу. Данные о конфигурации транспортной системы города могут быть получены непосредственно с использованием карты города, списка маршрутов и остановок. Возможные пути и передвижения нужно генерировать на основании этих данных. Для получения данных о пассажиропотоках нужно проводить обследование в городе.

В системе все эти данные хранятся в виде многомерной базы данных, созданной в СУБД MS SQL Server 2017. Часть данных непосредственно вводится в

БД с использованием интерфейса разработанного программного обеспечения. Другая часть генерируется подпрограммами системы RouteSearch и также вносится в БД. Система разработана с использованием технологии объектно-ориентированного программирования. Данные представляются в виде объектов, классы которых содержат методы чтения и записи в БД, сортировки и поиска и другие, что делает очень удобной работу с ними.

Литература

1. Яворский В.В., Утепбергенов И.Т. Структурные методы совершенствования управления транспортными системами городов. Караганда: КарГТУ, 2006. – 227с
2. Гасников А.В., Кленов С.Л., Нурминский Е.А., Холодов Я.А., Шамрай Н.Б. Введение в математическое моделирование транспортных потоков: учеб. пособие / Приложения: Бланк М.Л., Гасникова Е.В., Замятин А.А. и Малышев В.А., Колесников А.В., Райгородский А.М; Под ред. А.В. Гасникова. – М.: МФТИ, 2010. – 362 с.
3. Трофименко Ю.В., Якимов М.Р. Транспортное планирование: формирование эффективных транспортных систем крупных городов: монография. – М.: Логос, 2013. – 464 с.
4. Komninos. What makes cities intelligent? // Smart Cities: Governing, Modelling and Analysing the Transition. — Taylor and Francis. — P. 77.
5. Яворский В.В., Утепбергенов И.Т., Ахмедиярова А.Т. Models of analysis of distribution of passenger traffics in routed transport systems // Материалы международной конференции XIII Balkan Conference on Operational Research (BALCOR 2018). — Belgrade, Serbia, 2018.
6. Яворский В.В., Ключева Е.Г., Мутовина Н.В., Касымова Д.Т. Расчет транспортной потребности на основе анализа линейных многообразий // Материалы научной конференции ИИВТ КН МОН РК «Инновационные IT и Smart-технологии», посвященной 70-летию юбилею профессора Утепбергенова И.Т. — Алматы: ИИВТ МОН РК, 2019. — С. 278-286.

КОМПАЬОНЫ КОЛЬЦА ЦЕЛЫХ ЧИСЕЛ

Хисамиев З.Г., Алимжанов С.А.

KhisamievZ@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** В работе изучаются компаньон-модели кольца целых чисел, это исследование является примером изучения классического объекта посредством теории классов Фрэсе разработанной Нуртазиным А.Т. В исследовании описаны структуры компаньонов кольца целых чисел.*

Модели элементарной теории компаньонов кольца целых чисел

Пусть: \mathbb{Z} – кольцо целых чисел; $\mathbb{Z}[\bar{x}]$ – кольцо многочленов от переменных $\bar{x}=(x_1, \dots, x_n)$. Все используемые и неприведенные определения а также обозначения взяты из монографии [1]. Класс компаньонов кольца $\mathbb{Z}[\bar{x}]$ обозначаем $C(\mathbb{Z})$

ТЕОРЕМА 1. *Кольцо целостности $\mathbb{Z}[\bar{x}]$ является компаньоном кольца \mathbb{Z} т.е. $\mathbb{Z}[\bar{x}] \in C(\mathbb{Z})$*

Доказательство. Каждая конечная подмодель кольца \mathbb{Z} является подмоделью кольца $\mathbb{Z}[\bar{x}]$. Обратное пусть имеется конечная подмодель \mathfrak{F} кольца $\mathbb{Z}[\bar{x}]$, \mathfrak{F} задается конечной системой равенств и неравенств (\neq) правые и левые части которых содержат элементы \mathfrak{F} и операции сложения и умножения кольца $\mathbb{Z}[\bar{x}]$. Такая конечная система равенств и неравенств преобразуем в эквивалентную в $\mathbb{Z}[\bar{x}]$ систему $\& f_i(\bar{x}) = 0 \& \& g_j(\bar{x}) \neq 0$. Последняя система эквивалентна в $\mathbb{Z}[\bar{x}]$ системе с одним равенством и одним неравенством $f(\bar{x}) = 0 \& g(\bar{x}) \neq 0$, где $f(\bar{x}) = \sum f_i(\bar{x})$, $g(\bar{x}) = \prod g_j(\bar{x})$. Индукцией по числу переменных n нетрудно доказать существование набора $\bar{a}=(a_1, \dots, a_n) \in \mathbb{Z}$ такого что в \mathbb{Z} выполнено $f(\bar{a}) = 0 \& g(\bar{a}) \neq 0$. Теорема доказана.

Доказанную эквивалентность конечной системы уравнений и неравенств в $\mathbb{Z}[\bar{x}]$ системе из не более чем одного уравнения и не более чем одного неравенства, целесообразно сформулировать в виде

ПРЕДЛОЖЕНИЕ 1. Пусть $g_i(\bar{x}), h_j(\bar{x}) \in \mathbb{Z}[\bar{x}]$. Система $\& g_i(\bar{x}) = 0 \& \& h_j(\bar{x}) \neq 0$, эквивалентна в $\mathbb{Z}[\bar{x}]$ системе из одного уравнения и одного неравенства $g(\bar{x}) = 0 \& h(\bar{x}) \neq 0$.

Пусть $A_f = \{(\bar{a}, a) \mid f(\bar{a}, a) = 0, (\bar{a}, a) \in \mathbb{Z}\}$ – аннулятор f , где $f(\bar{x}, x) \in \mathbb{Z}[\bar{x}, x]$.

ПРЕДЛОЖЕНИЕ 2. Пусть $I(A_f) = \{g \mid g \in \mathbb{Z}[\bar{x}, x], A_f \subseteq A_g\}$, где неприводимый над кольцом \mathbb{Z} многочлен $f \in \mathbb{Z}[\bar{x}, x] \setminus \mathbb{Z}[\bar{x}]$ имеет содержание единицу. Тогда, $I(A_f)$ является главным идеалом в кольце $\mathbb{Z}[\bar{x}, x]$ и $I(A_f) = (f)$.

Доказательство. Действительно, неприводимый многочлен $f(\bar{x}, x)$ в кольце целостности $\mathbb{Z}[\bar{x}, x]$ является также неприводимым многочленом в кольце целостности $\widetilde{\mathbb{Z}[\bar{x}]}[x]$ над полем частных $\widetilde{\mathbb{Z}[\bar{x}]}$. в $\widetilde{\mathbb{Z}[\bar{x}]}[x]$ идеал $I(\widetilde{A}_f)$ будет главным, порождающим элементом которого будет неприводимый многочлен $f(\bar{x}, x) \in I(\widetilde{A}_f)$. Пусть $g \in I(A_f) \subseteq I(\widetilde{A}_f)$, тогда $f \mid g$ (f делит g) в кольце $\widetilde{\mathbb{Z}[\bar{x}]}[x]$, отсюда выполнено равенство $n(\bar{x})g(\bar{x}, x) = m(\bar{x})f(\bar{x}, x)q(\bar{x}, x)$, где $n(\bar{x}), m(\bar{x}) \in \widetilde{\mathbb{Z}[\bar{x}]}$, $q(\bar{x}, x) \in \mathbb{Z}[\bar{x}, x]$. В силу однозначности разложения в кольце целостности на неприводимые множители имеем $f(\bar{x}, x) \mid n(\bar{x})$ или $f(\bar{x}, x) \mid g(\bar{x}, x)$. Очевидно первое не возможно, вследствие выбора $f(\bar{x}, x)$, следовательно $f(\bar{x}, x) \mid g(\bar{x}, x)$. Предложение доказано.

ТЕОРЕМА 2. Пусть $f(\bar{x}, x)$ – неприводимый, с содержанием единица, многочлен над кольцом \mathbb{Z} . Тогда, алгебраическое расширение $\mathbb{Z}[\bar{x}]/f$ кольца $\mathbb{Z}[\bar{x}]$ является компаньоном \mathbb{Z} если и только если аннулятор A_f является бесконечным множеством.

Доказательство. Необходимость. Допустим противное – $A_f = \{(\bar{b}_1, b^1), \dots, (\bar{b}_k, b^k)\}$. Тогда выполнено: $\mathbb{Z}[\bar{x}]/f \models \exists \bar{u} \exists u (f(\bar{u}, u) = 0 \& \& (\bar{u}, u) \neq (\bar{b}_i, b^i))$. Так как \mathbb{Z} и $\mathbb{Z}[\bar{x}, x]/f$ компаньоны, а параметры $(\bar{b}_1, b^1), \dots, (\bar{b}_k, b^k)$ E -выразимы в \mathbb{Z} , тогда в \mathbb{Z} имеется решение $f(\bar{x}, x) = 0$ не принадлежащее A_f . Противоречие. Необходимость доказана.

Достаточность. Пусть A_f бесконечное множество. Достаточно доказать, что в моделях \mathbb{Z} и $\mathbb{Z}[\bar{x}, x]/f$ выполняются одни и те же экзистенциальные предложения. Очевидно, что экзистенциальное предложение истинное в \mathbb{Z} также истинно в $\mathbb{Z}[\bar{x}, x]/f$. Докажем, что каждое экзистенциальное предложение φ выполненное в $\mathbb{Z}[\bar{x}, x]/f$ выполняется в \mathbb{Z} . Как и в теореме 3.1 достаточно доказать, что если $\mathbb{Z}[\bar{x}]/f \models g(\bar{x}, x_f) = 0 \& h(\bar{x}, x_f) \neq 0$ (т.е. $g(\bar{x}, x_f) = 0 \& h(\bar{x}, x_f) \neq 0$ выполнено в $\mathbb{Z}[\bar{x}, x]/f$), то $\mathbb{Z} \models g(\bar{a}, a) = 0 \& h(\bar{a}, a) \neq 0$ для подходящих $(\bar{a}, a) \in \mathbb{Z}$, где $x_f = x + (f)$ – смежный класс в $\mathbb{Z}[\bar{x}, x]/f$, $f, g, h \in \mathbb{Z}[\bar{x}, x]$. Из предложения 1.2 вытекает существование набора $(\bar{a}, a) \in A_f \setminus A_h$. Действительно, если бы $A_f \subseteq A_h$, то $h \in (f)$ и $\mathbb{Z}[\bar{x}]/f \models h(\bar{x}, x_f) = 0$. Противоречие, достаточность а вместе с ней теорема доказаны.

ТЕОРЕМА 3. Пусть $\mathbb{Z}[\bar{x}, y]/f$ и $\mathbb{Z}[\bar{x}, y]/f/g$ (здесь $\mathbb{Z}[\bar{x}, y]/f/g = (\mathbb{Z}[\bar{x}, y]/f)[z]/g$) – простые алгебраические расширения колец $\mathbb{Z}[\bar{x}]/f$ и $\mathbb{Z}[\bar{x}, y]/f$ соответственно, $f(\bar{x}, y), g(\bar{x}, y, z)$ – неприводимые многочлены над кольцом $\mathbb{Z}[\bar{x}]$ и $\mathbb{Z}[\bar{x}, y]/f$ соответственно. Тогда, алгебраическое расширение $\mathbb{Z}[\bar{x}, y]/f/g$ кольца $\mathbb{Z}[\bar{x}, y]/f$ является компаньоном \mathbb{Z} если и только если множество $A_f \times \mathbb{Z} \cap A_g$ бесконечно.

Доказательство. Для $\mathbb{Z}[\bar{x}, y]/f/g$ выполняется следующее предложение

ПРЕДЛОЖЕНИЕ 1⁰. Система $\& g_i(\bar{x}, y, z) = 0 \& \& h_j(\bar{x}, y, z) \neq 0$, где $g_i(\bar{x}, y, z), h_j(\bar{x}, y, z) \in (\mathbb{Z}[\bar{x}, y]/f)[z]$ – кольцо многочленов над $\mathbb{Z}[\bar{x}, y]/f$, эквивалентна в $\mathbb{Z}[\bar{x}, y]/f/g$ системе из одного уравнения и одного неравенства $g'(\bar{x}, y, z) = 0 \& h(\bar{x}, y, z) \neq 0$.

Доказательство. Пусть $d(\bar{x}, y, z) = (g_i(\bar{x}, y, z))$ – наибольший общий делитель многочленов $\{g_i(\bar{x}, y, z)\}$ над полем частных $\overline{\mathbb{Z}[\bar{x}]/f}$ кольца $\mathbb{Z}[\bar{x}]/f$. Пусть $d(\bar{x}, y, z) = \frac{g'(\bar{x}, y, z)}{m(\bar{x}, y)}$, где $g'(\bar{x}, y, z) \in \mathbb{Z}[\bar{x}]/f[z]$, $m(\bar{x}, y) \in \mathbb{Z}[\bar{x}]/f, m(\bar{x}, y) \neq 0$.

Тогда, $f(\bar{x}, y, z)$ является общим делителем $\{g_i(\bar{x}, y, z)\}$ тогда и только тогда когда $f(\bar{x}, y, z)$ делит $g'(\bar{x}, y, z)$ в кольце $\mathbb{Z}[\bar{x}]/f[z]$. Отсюда следует, что система $\& g_i(\bar{x}, y, z) = 0$ и уравнение $g'(\bar{x}, y, z) = 0$ эквивалентны в кольце $(\mathbb{Z}[\bar{x}, y]/f)[z]$.

Пусть, далее $h(\bar{x}, y, z) = \prod_j h_j(\bar{x}, y, z)$. Так как многочлен $f(\bar{x}, y, z)$ неприводим, то имеем эквивалентность $(f(\bar{x}, y, z) | h(\bar{x}, y, z))$ тогда и только тогда, когда для каждого i , $f(\bar{x}, y, z) | h_i(\bar{x}, y, z)$. Предложение доказано.

Перейдем к доказательству теоремы. Необходимость. Опираясь на предложение 1⁰ будем рассматривать бескванторную часть E -предложения φ истинного, как состоящую из одного равенств и одного неравенства $h_1(\bar{u}, v, w) = 0 \& h_2(\bar{u}, v, w) \neq 0$ в $\mathbb{Z}[\bar{x}, y]/f/g$. Допустим противное – $A_f \times Z \cap A_g = \{\bar{b}_1, \dots, \bar{b}_k\}$ – конечно. Тогда выполнено: $\mathbb{Z}[\bar{x}, y]/f/g \models \exists \bar{u} \bar{v} \bar{w} (h_1(\bar{u}, v, w) = 0 \& h_2(\bar{u}, v, w) \neq 0 \& \&_i (\bar{u}, v, w) \neq \bar{b}_i)$. Так как \mathbb{Z} и $\mathbb{Z}[\bar{x}, y]/f/g$ компаньоны, а параметры $\bar{b}_1, \dots, \bar{b}_k$ E -выразимы в \mathbb{Z} , тогда в \mathbb{Z} имеется решение $h_1(\bar{u}, v, w) = 0 \& h_2(\bar{u}, v, w) \neq 0$ не принадлежащее множеству $\{\bar{b}_1, \dots, \bar{b}_k\}$. Противоречие. Необходимость доказана.

Достаточность. Пусть $A_f \times Z \cap A_g$ бесконечное множество. Достаточно доказать, что в моделях \mathbb{Z} и $\mathbb{Z}[\bar{x}, y]/f/g$ выполняются одни и те же экзистенциальные предложения. Очевидно что экзистенциальное предложение истинное в \mathbb{Z} также истинно в $\mathbb{Z}[\bar{x}, y]/f/g$. Докажем, что каждое экзистенциальное предложение φ выполненное в $\mathbb{Z}[\bar{x}, y]/f/g$ выполняется в \mathbb{Z} . Можно считать, что бескванторная формула φ имеет вид $h_1(\bar{u}, v, w) = 0 \& h_2(\bar{u}, v, w) \neq 0$, содержащее решение в кольце целостности $\mathbb{Z}[\bar{x}, y]/f/g$. Докажем, что имеется $(\bar{p}, q, r) \in A_f \times Z \cap A_g$, что выполняется $h_1(\bar{p}, q, r) = 0 \& h_2(\bar{p}, q, r) \neq 0$. В кольце $\mathbb{Z}[\bar{x}, y]/f/g$ рассмотрим идеал $I = \{h \in (\mathbb{Z}[\bar{x}, y]/f)[z] : A_f \times Z \cap A_g \subseteq A_h\}$. Выберем в I многочлен d_1 наименьшей степени по z . Очевидно, что в $(\mathbb{Z}[\bar{x}, y]/f)[z]$ имеет место тождество $e_1(\bar{x}, y)g(\bar{x}, y, z) = e_2(\bar{x}, y)d_1(\bar{x}, y, z)$, для подходящих ненулевых в $\mathbb{Z}[\bar{x}, y]/f$ многочленов $e_1(\bar{x}, y), e_2(\bar{x}, y) \in \mathbb{Z}[\bar{x}, y]$. Отсюда, $d_1(\bar{x}, y, z) | g(\bar{x}, y, z)$, следовательно $d_1(\bar{x}, y, z) = g(\bar{x}, y, z)$ и любой многочлен $h(\bar{x}, y, z) \in I$ имеет вид $e_1(\bar{x}, y)h(\bar{x}, y, z) = e_2(\bar{x}, y)g(\bar{x}, y, z) + h'(\bar{x}, y)$. В свою очередь, любой многочлен $h'(\bar{x}, y) \in I$ имеет вид $a(\bar{x})h'(\bar{x}, y) = b(\bar{x})f(\bar{x}, y)$, для подходящих ненулевых в $\mathbb{Z}[\bar{x}, y]$ многочленов $a(\bar{x}), b(\bar{x}) \in \mathbb{Z}[\bar{x}]$. Таким образом, $e_1(\bar{x}, y)h(\bar{x}, y, z) = e_2(\bar{x}, y)g(\bar{x}, y, z) + a(\bar{x})f(\bar{x}, y)$, для подходящих ненулевых многочленов соответственно в $\mathbb{Z}[\bar{x}]/f$ и $\mathbb{Z}[\bar{x}]$, $e_1(\bar{x}, y), e_2(\bar{x}, y) \in \mathbb{Z}[\bar{x}, y], a(\bar{x}) \in \mathbb{Z}[\bar{x}]$. Тогда, $e_1(\bar{x}, y)h_2(\bar{x}, y, z) = 0$ в $\mathbb{Z}[\bar{x}]/f/g$, где

$e_1(\bar{x}, y) \neq 0$ в $\mathbb{Z}[\bar{x}]/f/g$. Так делителей нуля в $\mathbb{Z}[\bar{x}]/f/g$ нет, то $h_2(\bar{x}, y, z) = 0$.
Полученное противоречие доказывает достаточность, а вместе с ней и теорему.

Литература

1. Нуртазин А.Т. Введение в теорию моделей, элиминация кванторов и экзистенциальная замкнутость. – Алматы: НЦ ГНТЭ, 2017.– 187с.(ISBN 978-601-332-022-9).

АВТОМАТИЧЕСКОЕ УСТАНОВЛЕНИЕ ТОНАЛЬНОСТИ ТЕКСТОВ МЕТОДОМ ИХ КОНЦЕПТУАЛЬНОГО АНАЛИЗА

Хорошилов А.А.^{1,2}, Мусабаев Р.Р.¹, Красовицкий А.М.¹,
Уалиева И.М.¹, Мейрамбеккызы Ж.¹
e-mail: khorosilov@mail.ru

¹Институт информационных и вычислительных технологий
КН МОН РК, Казахстан

²Институт проблем информатики Федерального исследовательского
центра «Информатика и управление» Российской академии наук, Россия

Аннотация. В статье изложено решение задачи создания лингвистического инструментария и методики автоматического определения тональности новостных сообщений, связанных с качеством жизни рядового гражданина. Был определен подход для решения задачи, разработано программное обеспечение и методы автоматизированного создания словарей объектов и словарей оценочных предикатов, а также словарей модификаторов меры оценки. Проведенный эксперимент подтвердил правильность предложенной методики оценки событий, освещенных в новостных сообщениях и работоспособность программного комплекса.

Эта методика, при соответствующем выборе объектов оценки событий, может быть использована при создании тональных портретов конкретных авторов по совокупности их публикаций, а также тональных портретов различных новостных агрегатов по совокупности освещаемых ими событий в конкретный временной интервал.

Введение

На психоэмоциональной настрой современного общества значительное влияние оказывают электронные средства массовой информации (ЭСМИ). Эти средства информируют общество о событиях, происходящих в мире. Но, как правило, большинство событий освещается с разных точек зрения и в значительной части случаев цель такого освещения событий является желание автора публикации сформировать у читателя определенное мнение, созвучное мнению автора. Но самом деле ЭСМИ в ряде случаев навязывают такое мнение читателям, которое заказывается авторам идеологических публикаций. Как правило, чтобы сформировать у читателя определенное мнение по идеологическим вопросам читателю преподносится заранее препарированная информация – подаются

определенные события в позитивном или негативном плане, или некоторые события или вообще не происходили (так называемые фейковые события), или являлись постановочными событиями. Такое освещение событий формирует определенный психоэмоциональный настрой у определенной части общества. Сейчас все такие манипуляции с ЭСМИ называют элементами информационной войны. Для борьбы с такими манипуляциями необходимо выработать автоматизированные средства противодействия, базирующиеся на методах искусственного интеллекта.

Одним из наиболее действенных способов противодействия этим манипуляциям был разработан в США в начале Второй мировой войны, когда Германия объявила войну США. В то время в США были сильны профашистские настроения и функционировало большое число газет, пропагандирующих эти настроения. Поскольку США являлось демократическим правовым государством, то необходимы были формальные основания для закрытия по судебному решению этих профашистских газет. Таким основанием могли бы быть количественные данные по фактам позитивного отношения к фашистской идеологии и одновременно негативного - к устоям демократического общества. Для решения этой задачи были привлечены лингвисты, которые проанализировали большое число газетных публикаций и установили критерии, в соответствии с которыми можно было бы формально отнести газеты к профашистским. При этом была предложена относительно простая методика, заключающаяся в выборе списка значимых фамилий и терминов, относящихся как к фашистской идеологии, так и к демократическим ценностям, с целью определения тональности контента всех случаев упоминания этих терминов и фамилий. Впоследствии такой метод определения тональности текстов назвали контент-анализом.

1. Обзор методов определения тональности текстов

При решении проблемы автоматического установления тональности текстов можно выделить несколько основных подходов, изложенных в работах [1-5]:

- Подходы, базирующиеся на методах машинного обучения и векторном сравнении исходных текстов с размеченными эталонными коллекциями, предполагают предварительное обучение машинного классификатора путем раскладывания каждого документа в виде вектора признаков. При этом каждому документу указывается его тип тональности. При реализации этих методов требуются значительные трудозатраты при разметке текстов, необходимые для обучения алгоритмов обучения. Эти методы достаточно эффективны при анализе однородных текстов слабо подверженным изменениям тематики и их эмотивного настроения.

- Подходы, базирующиеся на выявлении в текстах эмотивной лексики на основе использования тональных словарей. В рамках этого метода часто используются списки паттернов, подставляемых в регулярные выражения, а также правила формирования шаблонов для объединения слов в цепочки. Обычно в качестве шаблонов использовались два типа синтаксических конструкций: «прилагательное + существительное» и «наречие + глагол». При этом тональность текста рассчитывалась как среднее эмоциональных оценок. Оценка для каждой конструкции определяется путем установления совместной встречаемости их со словами тонального словаря.

- С нашей точки зрения наиболее перспективными и универсальными подходами к решению задачи установления тональности текстов являются подходы, ориентированные на глубокий семантический анализ текстов. В рамках этого

подхода возможно обеспечить детальный анализ тональности в пределах каждого предложения и в отношении каждого объекта. Так, например, в работе [1] производилось определение тональности у объектов на основе предикационных отношений в пропозиции. Тональность устанавливалась на основе классификационных тональных словарей.

В работе [5] ставилась задача разработки методов оценки деятельности органов власти на основе набора показателей видов их деятельности и процедурных средств лингвистического анализа текстов. Преимущество таких методов заключается в том, что они обеспечивают получение более объективных количественных показателей и освобождают результаты оценки от субъективного экспертного представления о действительной ситуации.

2. Постановка задачи исследований

Исходя из конкретных задач НИР* в рамках нашего исследования была поставлена задача разработки лингвистического инструментария и методики автоматического определения тональности новостных сообщений, связанных с качеством жизни† рядового гражданина. Как известно, информация, относящаяся к повышению или понижению качества жизни, в значительной степени определяют его эмоциональное отношение к властным структурам. При этом мы исходили из того, что на эмоциональный настрой пользователей ЭСМИ влияет сознательный подбор освещаемых событий и их эмоциональная окраска. Для решения этой задачи необходимо исследовать содержательную и эмоциональную составляющие текстов, составить оценочные словари и разработать программные средства их формализации и выявления в текстах.

Для установления количественных характеристик тональности текстов воспользуемся формулой подсчета значения веса тональности текстов, приведенной в работе [2], но для нормирования числа предложений в документе берем не их абсолютное значение, а логарифм значения:

$$W_k = \frac{\sum_{i=1}^{N_c} W_i}{\log(N)}, \quad (1)$$

где,

W_k – значение веса тональности k -ого документа ;

W_i – значение веса тональности i -ой ситуации;

N – общее число предложений в документе.

Для формирования тонального портрета подмножества документов все тексты T_i помещаются в одномерное эмотивное пространство в соответствии со своим весом W_{T_i} , причем тексты положительной тональности должны занимать в

* Работа проводится в рамках ПЦФ BR05236839 «Разработка информационных технологий и систем для стимулирования устойчивого развития личности как одна из основ развития цифрового Казахстана»

† Совокупность свойств и характеристик жизни человека, относящихся к их способности удовлетворять его существующие и предполагаемые потребности [ГОСТ Р 22.10.01-2001].

этом пространстве местоположение справа, отрицательной тональности – местоположение слева. Тексты нейтральной тональности должны располагаться в центре этого пространства. Среднее значение текстов положительной или отрицательной тональности может быть вычислено по формуле:

$$AW_t^c = \frac{\sum_{i=1}^{N_c} W_{T_i}}{N_c}, \quad T_i \in C \quad (2)$$

где,

AW_t^c – средний вес текстов, относящихся к классу тональности C ;

N_c – число текстов относящихся к классу тональности C ;

После вычисления средних значений текстов положительной или негативной тональности необходимо установить для них границу D – как середину отрезка $[AW_t^- ; AW_t^+]$. На этом отрезке текст T в соответствии со его знаком и значение его веса располагаются справа или слева от границы D . При этом, как уже было отмечено, все тексты справа от границы D относятся к положительному классу, а слева – к отрицательному.

3. Теоретические предпосылки решения задачи

Общеизвестно, что смысловое содержание текстов или их фрагментов выражается через систему единиц смысла и их отношений. Наиболее устойчивыми единицами смысла являются понятия (отдельные слова и устойчивые фразеологические и терминологические словосочетания). Второй по значимости единицей смысла является предложение. Основной чертой предложений является их предикативность – то есть свойство утверждения наличия у объектов определенных признаков и их отношений. В основе смысловой структуры предложений лежит его предикатно-актантная структура, компонентами которой являются понятия-предикаты (признаки и отношения) и понятия-актанты, выступающие в роли описываемых объектов.

Для установления психоэмоционального настроения текстов необходимо определить объекты, в отношении которых производится оценка и те компоненты смысловой структуры, которые определяют такую оценку. В качестве объектов обычно выступают существительные, в качестве определителей оценки – глаголы. Для усиления или уменьшения эмоциональной оценки используются прилагательные и наречия. Для более точной оценки тональности текстов были предложены три категории объектов. В качестве таких объектов первой категории фигурируют, все те объекты, явления и ситуации, которые в наибольшей степени влияют на качество жизни гражданина. В качестве предикатов – те процессы, которые улучшают или ухудшают это качество. Примерами таких объектов и их оценок могут служить доступность и стоимость предметов первой необходимости (продуктов питания, медикаментов, одежды и др.), качество здравоохранения, социальной защиты, ЖКХ и борьба с правонарушениями. В качестве объектов второй категории – все те ситуации, которые очень важны для общества в целом, но в меньшей степени влияют на качество жизни конкретного гражданина. К таким явлениям можно отнести борьбу с коррупцией, борьбу с терроризмом, правоохранительную, судебную и налоговую систему. И, наконец, к третьей категории можно отнести все те явления, которые определяют престиж страны в целом. Это международная деятельность, культура, спорт, торговля, отношения с другими странами и др.

4. Методика и технологии создание декларативных средств

При разработке методов оценки текстов необходимо разработать эталонные оценочные словари и методику их применения при анализе текстов. В рамках наших исследований мы пытались решить эту проблему путем автоматизации процесса формирования ситуативных оценочных словарей. Как следует из работ [2] для определения ситуаций и их тональной оценки требуется создать комплекс словарей, состоящий из объектов и связанных с ними предикатов, а также их определителей – наречий (для глаголов) и прилагательных (для существительных), выступающих в роли модификаторов их меры тональности. Обычно решение этой задачи возлагалось на экспертное сообщество. При таком решении затрачивались значительные человеческие ресурсы и длительное время на их создание. Автоматизация этого процесса позволит сократить трудозатраты и время на создание таких словарей.

Основная идея метода автоматизированного составления тональных словарей заключается в том, чтобы на основе уже имеющихся словарных ресурсов обеспечить возможность автоматического выявления в текстах объектов оценки. К таким словарным ресурсам можно отнести различные перечни объектов и текстов, в которых они встречаются. Процесс выявления объектов первой необходимости можно начать, например, с ассортимента потребительской корзины*. Этот ассортимент товаров и услуг необходимо выявить в текстах путем их лингвистического анализа и установить в текстах их смысловую связь с лексическими конструкциями, осуществляющими оценку объектов. В предложении такими конструкциями являются глаголы (предикаты). Эксперту должен для каждой такой пары «объект-предикат» зафиксировать оценочный глагол и назначить этой паре индекс тональности (позитив/негатив). Категория тональности для этих объектов будет устанавливаться автоматически. Далее, на основе установленных предикатов, можно будет определить в коллекции текстов все возможные объекты, встречающиеся совместно с ними. Эксперту должен только установить индекс их тональности. Объекты других категорий, относящиеся к отраслевым видам деятельности, можно будет установить по различным отраслевым документам: глоссариям, тезаурусам, онтологиям, нормативно-справочным документам и др. Здесь также задача эксперта для каждой выявленной пары «объект-предикат» установить индекс тональности. Категория будет устанавливаться автоматически в соответствии с установленной категории для данной отрасли.

Модификаторы степени тональности: наречия (для глаголов) и прилагательные (для существительных) устанавливаются на основе результатов семантико-синтаксического анализа. Задача экспертов также сводится к установлению степени тональности (слабая/средняя/сильная).

После того как выполнен первоначальный этап установления объектов, предикатов и модификаторов, относящихся к оценке тональности объектов первой

* «необходимые для сохранения здоровья человека и обеспечения его жизнедеятельности минимальный набор продуктов питания, а также непродовольственные товары и услуги, стоимость которых определяется в соотношении со стоимостью минимального набора продуктов питания» [Потребительская корзина 2013-2018 год. Информационный портал РФ].

категории, можно будет перейти к реализации второго этапа, целью которого будет выявление объектов второй и третьей категории и их оценочных конструкций. Этот этап по своим задачам гораздо сложнее, поскольку необходимо определить виды деятельности в отраслях и установить наборы их оценочных показателей. Поэтому этот этап разработки должен базироваться в большей степени на технологиях глубинного семантического анализа текстов, позволяющего автоматически выявить в отраслевых текстах компоненты их предикатно-актантной структуры и автоматически назначить им набор тональных признаков. В табл.1 приводятся фрагменты комплекса оценочных тональных словарей.

Таблица 1 - Фрагменты комплекса оценочных тональных словарей

Категория	Компоненты предикатно-актантной структуры	Классифик. индекс	Прям./ Обратн. совместим.	Колич. харак-ка
Основные объекты (виды деятельности в отраслях)				
1	хлебо-булочные изделия	N00001.001	D	1.0
1	вино-водочная продукция	N00001.001	D	1.0
1	куриные яйца	N00001.001	D	1.0
2	налог на имущество	N00254.001	D	0.6
2	подходный налог	N00254.001	D	0.6
3	футбольная команда	N03675.001	D	0.3
Оценочные предикаты (предикативные конструкции)				
0	подешеветь	V00042.001	D	1.0
0	подорожать	V000044.001	N	1.0
0	увеличился	V000048.001	N	1.0
0	понижился	V000049.001	D	1.0
0	выиграть	V000021.001	D	1.0
0	проиграть	V000520.001	D	1.0
Модификаторы тональности - прилагательные				
0	незначительный	A000052.001	D	0.8
0	весомый	A000021.001	D	1
Модификаторы тональности - наречия				
0	очень	Y000042.001	D	1.2
0	незначительно	Y000037.001	D	0.8

В таблице для каждого словарного объекта приводится индекс его категории, классификационный индекс, индекс прямой или обратной совместимости и весовой коэффициент. Классификационный индекс объекта соответствует его родовому индексу в тематической онтологии, а индекс прямой совместимости сочетания «объекта- предикат» соответствует основной тональности предиката, или обратный индекс – изменяет эту тональность на противоположную.

В табл. 2 приведены количественные характеристики словарей текущей версии словарного комплекса.

Таблица 2 – Количественные характеристики словарей

n/n	Наименование словаря	Объем словаря
Словарь объектов (видов деятельности в отраслях)		
1	Словарь «Продуктовая корзина»	897

2	Словарь по ЖКХ	3427
3	Словарь по спорту	2588
4	Налоговый словарь	2764
5	Словарь по международной деятельности	3536
6	Словарь оценочных предикатов	1736
7	Словари модификаторов тональности: прилагательных	237
8	наречий	147
	Общее количество словарных статей	15332

5. Технология автоматического установления тональности текстов

Технология автоматического установления тональности текстов базируется на процедурах их семантико-синтаксического анализа, в результате работы которой смысловое содержание текстов формализуется и структурируется, и выявляются компоненты предикатно-актантной структуры. Эти формализованные компоненты последовательно сопоставляются с комплексом тональных словарей: вначале сопоставляются наименования объектов, после этого, в случае их успешного поиска в словарях объектов, производится сопоставление выявленных в тексте предикатов с элементами словаря оценочных предикатов. При успешном сопоставлении и при наличии в текстах предполагаемых модификаторов меры тональности (наречий и прилагательных) также производится их поиск в соответствующих словарях. Далее после получения всей необходимой информации о событии вычисляется его количественная характеристика по следующей формуле:

Технологию автоматического установления тональности текстов можно представить в виде следующей последовательности лингвистических процедур обработки текста:

1. Выполнение формально-логического контроля текста.
2. Членение на отдельные слова (по пробелам и разделительным знакам между ними).
3. Выполнение морфологического анализа слов.
4. Выполнение членения текста на предложения.
5. Выполнение семантико-синтаксического анализа предложений.
6. Выполнение концептуального анализа предложений и выявление в них компонентов предикатно-актантной структуры.
7. Приведение компонентов предикатно-актантной структуры к их формализованному представлению.
8. Сопоставление компонент предикатно-актантной структуры с элементами комплекса тональных оценочных словарей.
9. Вычисление количественных характеристик тональности событий.
10. Формирование статистического отчета о тональности текста.

6. Программные средства определения тональности текстов

При реализации программно-технологического комплекса определения тональности текстов мы базировались на процедурах программно-лингвистической платформе MetaFraz [6], в рамках которой выполнялась формализация смыслового содержания текстов и выявлялась его формализованная предикатно-актантная структура. Методы и технологии машинной грамматики и семантико-синтаксического анализа достаточно подробно описаны в работах [7,8]. Разработанный в рамках данного исследования программный модуль осуществлял

разбор на отдельные компоненты предикатно-актантную структуру предложений и производил сопоставление по определенной схеме этих компонент с тональными оценочными словарями. Аналогично сопоставлялись с соответствующими словарями модификаторами меры тональности – наречия и прилагательные. Программно-технологическая схема подсистемы автоматического определения тональности текстов приведена на рис. 1.

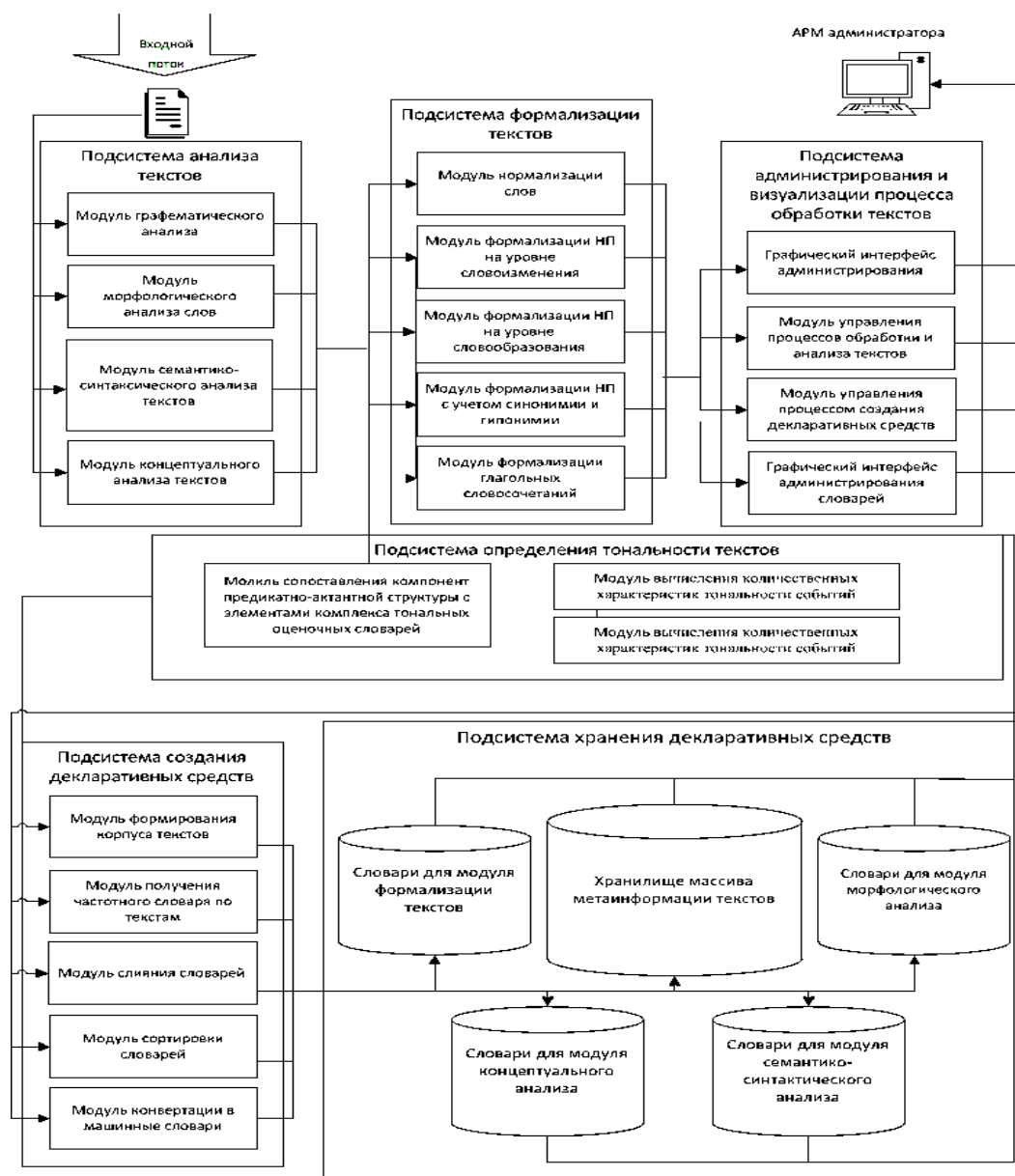


Рис.1 – Программно-технологическая схема подсистемы автоматического определения тональности текстов

7. Методика оценки тональности текстов

Методика оценки тональности текстов должна в какой-то степени моделировать восприятие человеком смыслового содержания новостных сообщения и их эмоциональной окраски. Обычно он воспринимает всю сумму событий, содержащихся в этих сообщениях, и формирует свое собственное мнение об этих событиях. Это мнение во многом зависит, от того насколько эти события

влиять или могут повлиять на его качество жизни и в гораздо меньшей степени на его мнения может повлиять мнение автора публикации. Поэтому мы согласились с авторами статьи [2] утверждавших, что лучшие результаты были получены на основе суммы эмоциональных оценок отдельных сообщений, а не на их частоту употребления, как обычно это применяется в методах машинного обучения.

Количественное значение тональной оценки i -ого события вычислялось по формуле:

$$S_i = C_i \cdot K_{oi} \cdot K_{mi} \cdot K_{ni},$$

где,

S_i – значение тональной оценки i -ого события;

C_i – i -ое событие;

K_{oi} – коэффициент категории i – ого объекта события (диапазон значений коэффициента $[-1, -0.6, -0.3, +0.3, +0.6, \mp 1]$);

K_{mi} ; - коэффициент i – ого модификатора оценочного предиката (диапазон значений коэффициента $[0.1, 1.2]$);

K_{ni} – коэффициент i – ого модификатора объекта события (диапазон значений коэффициента $[0.1, 1.2]$).

8. Эксперимент по автоматическому установлению тональности текстов

Задача эксперимента заключалась в проверке работоспособности предложенных технологических решений и проведения экспертной оценки результатов автоматического установления оценки тональности текстов. Для каждого документа выявлялись все анализируемые ситуации и вычислялась для каждой ситуации его тональная оценка, далее вычислялась абсолютная и средняя тональная оценка документа (W). В табл. 3 приведены результатов автоматической оценки тональности документов, влияющих на качество жизни рядового гражданина.

Таблица 3 – Результаты автоматической оценки тональности документов

Название документа и значение его тональной оценки (Т)	Позитивная оценка			Негативная оценка		
	3-ая кат.	2-ая кат.	1-ая кат.	1-ая ка-я.	2-ая кат.	3-ая кат.
Какие продукты подорожали в Казахстане ($W=-1.49$)			12	14		
Хорошее образование не может быть дешевым ($W=1.34$)		4			1	
Для большинства казахстанцев девальвация – это благо ($W=0.91$)		2	1			
Великая Победа – напоминание всем тем, кто вынашивает планы против нас ($W=0.60$)		3			1	
Проблемы образования касаются каждого ($W=0.83$)		3				
Хорошее образование не может быть дешевым ($W=1.49$)		4				
Чем обернется «дедолларизация» в СНГ, предложенная Владимиром Путиным ($W=0.32$)		3			2	

Празднование 550-летия Казанского ханства должно поднять наш регион на новый уровень (W=1.52)	6	1				
---	---	---	--	--	--	--

9. Перспективы развития метода концептуального анализа текстов

Перспективы развития метода концептуального анализа текстов, ориентированного на автоматическое установление тональности событий, освещаемых в текстах, может развиваться по нескольким направлениям:

1. Совершенствованиям алгоритмов семантико-синтаксического и концептуального анализа текстов связано с повышением точности лексического анализа, разрешением неоднозначной синтаксических и семантических ситуаций и широким применением словарей синтаксических конструкций и системы правил их трансформаций [8].

2. Разработка методов выявления объектов анализа и их оценки должна базироваться на масштабных моделях предметных областей, представленных в виде низшего и среднего уровня тематических онтологии и включающих основной понятийный состав предметных областей и систему смысловых отношений между наименованиями понятий этих предметных областей [5].

3. Совершенствование методов оценки текстовых событий связаны с моделированием восприятия человеком позитивных или негативных событий и формирования у него соответствующего психоэмоционального настроения к причинам, вызывающим эти события [5].

Заключение

В статье изложено решение задачи создания лингвистического инструментария и методики автоматического определения тональности новостных сообщений, связанных с качеством жизни рядового гражданина. Был определен подход для решения задачи, разработано программное обеспечение и методы автоматизированного создания словарей объектов и словарей оценочных предикатов, а также словарей модификаторов меры оценки. Проведенный эксперимент подтвердил правильность предложенной методики оценки событий, освещенных в новостных сообщениях и работоспособность программного комплекса.

Эта методика, при соответствующей выборке объектов оценки событий, может быть использована при создании тональных портретов конкретных авторов по совокупности их публикаций, а также тональных портретов различных новостных агрегатов по совокупности освещаемых ими событий в конкретный временной интервал.

Предлагаемый метод, основанный на методах автоматизированного составления словарей по текстам документов, методах семантико-синтаксического анализа и использования онтологических ресурсов, является универсальным поскольку позволяет проводить детальный семантический анализ коллекций документов по различным многомерным критериям. Комплекс тональных словарей большого объема и средства формализации текстов позволят проводить этот анализ с достаточно высокой точностью.

Работа выполнена в рамках проекта ПЦФ №BR05236839 «Разработка информационных технологий и систем для стимулирования устойчивого развития личности как одна из основ развития цифрового Казахстана»

Литература

1. Пазельская А.Г., Соловьев А.Н. Метод определения эмоций в текстах на русском языке, Dialog 2011. <http://www.dialog-21.ru/media/1475/dialog2011.pdf> (09.06.2019)
2. Клековникова М.В., Котельников Е.В. Метод автоматической классификации текстов по тональности, основанный на словаре эмоциональной лексики. <http://ceur-ws.org/Vol-934/paper15.pdf> (09.06.2019)
3. Посевкин Р.В., Бессмертный И.А. Применение сентимент анализа текстов для оценки общественного мнения. Научно-технический вестник информационных технологий, механики и оптики, №1, том 15, 2015.
4. Ермаков С.А., Ермакова Л.М. Методы оценки эмоциональной окраски текстов. Вестник пермского университета. Вып. (1) 9, 2012.
5. Хорошилов Ал-др. А., Никитин Ю.В., Хорошилов Ал-ей. А., Будско В.И. Автоматическое создание формализованного представления смыслового содержания неструктурированных текстовых сообщений СМИ и социальных сетей // Системы высокой доступности, №3, том.10, 2014.
6. Калинин Ю.П., Хорошилов Ал-др. А., Хорошилов Ал-ей. А. Современные технологии автоматизированной обработки текстовой информации // Системы высокой доступности, № 2, том.11, 2015.
7. Будско В.И., Калинин Ю.П., Козеренко Е.Б., Хорошилов А.А., Хорошилов А.А. Машинная грамматика русского языка // Системы высокой доступности, №3, том.13, 2017.
8. Кан А.В., Ревина В. Д., Руснак В.И., Хорошилов Ал-др А., Хорошилов Ал-сей А. Автоматическое формирование синтаксической модели языка для задач машинного перевода и информационного поиска. Сб. «Научно-техническая информация», Сер. 2, № 12, ВИНТИ, 2018.

РАЗРАБОТКА ИСКУССТВЕННОЙ ИММУННОЙ СЛОЖНОЙ СИСТЕМЫ УПРАВЛЕНИЯ НЕФТЕГАЗОВОЙ ОТРАСЛИ НА ОСНОВЕ АЛГОРИТА КЛОНАЛЬНОЙ СЕЛЕКЦИИ

Ширяева О.И.

e-mail: oshiryayeva@gmail.com

*Институт информационных и вычислительных технологий КН МОН РК,
Республика Казахстан*

***Аннотация.** Представлены результаты синтеза сложного объекта нефтегазовой отрасли – дистилляционной колонны, в которой протекают процессы очистки природного газа от различных примесей. В результате сравнительного анализа, в качестве метода исследования, выбрана такая технология Smart-систем, как канонический алгоритм клональной селекции (CLONALG) искусственных иммунных систем (AIS). С использованием CLONALG, получена процедура синтеза регуляторов изолированных подсистем сложной системы управления процессом очистки газа. Для решения вопроса влияния взаимосвязей на процессы подсистем использовался механизм построения*

развязывающей матрицы, при включении в схему которой многоконтурная система вырождается в набор независимых подсистем.

Ключевые слова. *Smart-технологии, искусственные иммунные системы (AIS), алгоритм клональной селекции (CLONALG), MIMO-система.*

Введение. Проблема синтеза управления сложными, то есть многомерными многосвязными, технологическими процессами является актуальной задачей в теории управления. Перспективным направлением при проектировании MIMO-систем является применение биоинспирированных алгоритмов, эффективно использующихся при решении оптимизационных задач [1].

В настоящее время большое количество публикаций посвящено синтезу оптимального управления на основе биоинспирированных алгоритмов для одномерных (SISO) систем [2]. Широко используются генетические алгоритмы (GA) [3,4], алгоритмы колонии муравьев (ACO) [5], алгоритмы роя частиц (PSO) [6], искусственные иммунные системы (AIS) [7]. Рассматриваемые алгоритмы имеют следующие отличительные особенности: высокая точность вычислений, быстрое действие, возможность работы алгоритмов при отсутствии полноты информации об объекте управления.

Особый интерес представляют алгоритмы настройки регуляторов на основе AIS. В частности, алгоритма CLONALG, применение которого для решения задач оптимизации и настройки регуляторов даёт более эффективные результаты, чем ACO и GA. В статье [8] результаты моделирования показывают, что алгоритм иммунной системы для настройки ПИД-регулятора имеет высокую точность решения и скорость сходимости.

Применение алгоритма искусственной иммунной системы для задач оптимизации представлено в статье [9], в качестве примера, которого было взято управление скоростью сервопривода постоянного тока. Контроллер ПИД-регулятора был настроен с использованием алгоритма AIS путем минимизации абсолютной ошибки времени. Результаты сравнивались с результатами метода настройки Циглера-Николса, и было очевидно, что AIS даёт лучшие результаты. Алгоритм AIS показал, что он способен находить глобальное оптимальное решение, и он дал ответ лучше, чем ответ традиционных методов настройки с точки зрения времени нарастания, времени регулирования, ошибки установившегося состояния и перерегулирования.

Работы [10, 11] также посвящены настройке ПИД-регуляторов систем с использованием алгоритма AIS. Результаты моделирования показывают, что алгоритм AIS является эффективным методом оптимизации параметров ПИД-регуляторов.

Среди алгоритмов AIS интенсивно используется алгоритм CLONALG. Клональная селекция является теорией, которая используется для объяснения механизмов иммунного ответа при распознавании иммунными клетками образов чужих антигенов антителами [12]. Алгоритм CLONALG, моделирующий основные принципы теории клонального отбора иммунной системы, относится к категории систем вычислительного интеллекта, куда входят системы способные к адаптации своего поведения, при достижении целей, т.е. обладающие способностью обучаться даже в отсутствие человека, в отличие, от систем искусственного интеллекта и так называемых «символьных вычислений», последние из которых неспособны к самообучению вообще.

Алгоритм CLONALG представляет интерес в качестве развития его на класс MIMO-систем управления. Целью данной статьи является получение базовых

концепций формирования алгоритмов синтеза оптимального управления на основе алгоритма клональной селекции CLONALG искусственных иммунных систем для ММО системы управления.

1 Базовые концепции применения алгоритма CLONALG для синтеза оптимального управления сложной системой. В качестве объекта управления нефтегазовой отрасли рассмотрим процесс очистки газа в дистилляционной колонне [13]. Линеаризованная модель дистилляционной колонны может быть представлена системой с двумя входами и двумя выходами [13]:

$$\begin{vmatrix} \chi_D(s) \\ \chi_B(s) \end{vmatrix} = \frac{1}{1+1.9588s} \begin{vmatrix} 0.0042 & -0.0062 \\ -0.0052 & 0.0072 \end{vmatrix} \begin{vmatrix} L(s) \\ V(s) \end{vmatrix}. \quad (1)$$

L – скорость потока жидкости;

V – скорость потока пара.

χ_B – концентрация жидкости в нижней части колонны;

χ_D – концентрация жидкости в дисцилляторе.

В соответствии с (1) введём обозначения для двух управляемых входов, u_1, u_2 , и двух выходных сигналов, y_1, y_2 :

$$\begin{vmatrix} u_1(s) \\ u_2(s) \end{vmatrix} = \begin{vmatrix} L(s) \\ V(s) \end{vmatrix}, \quad \begin{vmatrix} y_1(s) \\ y_2(s) \end{vmatrix} = \begin{vmatrix} \chi_D(s) \\ \chi_B(s) \end{vmatrix}. \quad (2)$$

Пусть $G_{11}(s), G_{12}(s), G_{21}(s), G_{22}(s)$ – передаточные функции объекта, связывающие входы и выходы:

$$\begin{vmatrix} y_1(s) \\ y_2(s) \end{vmatrix} = \begin{vmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{vmatrix} \cdot \begin{vmatrix} u_1(s) \\ u_2(s) \end{vmatrix} = \begin{vmatrix} \frac{0.0042}{1.9588s+1} & \frac{-0.0062}{1.9588s+1} \\ -0.0052 & 0.0072 \\ \frac{1.9588s+1}{1.9588s+1} & \frac{1.9588s+1}{1.9588s+1} \end{vmatrix} \cdot \begin{vmatrix} u_1(s) \\ u_2(s) \end{vmatrix}. \quad (3)$$

Рассмотрим математическую модель изолированных подсистем:

$$\begin{vmatrix} y_1(s) \\ y_2(s) \end{vmatrix} = \begin{vmatrix} G_{p11}(s) & 0 \\ 0 & G_{p22}(s) \end{vmatrix} \cdot \begin{vmatrix} u_1(s) \\ u_2(s) \end{vmatrix} = \begin{vmatrix} \frac{0.0042}{1.9588s+1} & 0 \\ 0 & \frac{0.0072}{1.9588s+1} \end{vmatrix} \cdot \begin{vmatrix} u_1(s) \\ u_2(s) \end{vmatrix} \quad (4)$$

Для изолированных подсистем (4) синтезируем оптимальное управление на основе алгоритма CLONALG.

Постановка задачи. Для изолированных подсистем ММО-системы (4), синтезировать оптимальные регуляторы на основе алгоритма CLONALG, минимизирующие интегральные квадратичные критерии качества:

$$f_i = \int_{t=0}^{\infty} e_i^2(t) dt. \quad (5)$$

где $e(t)$ – ошибка рассогласования между выходным и заданным сигналами;
 i – число подсистем.

Алгоритм CLONALG часто применяется для решения задач оптимизации и распознавания образов, напоминающие параллельный алгоритм с восхождением к вершине и генетический алгоритм без оператора кроссинговера [12]. Алгоритм разработан на основе процессов иммунной системы и её клонально-селективной теории. В контексте данной теории, иммунная система вырабатывает антитела в ответ на внедрение чужеродных антигенов. Антитела размножаются, пропорционально степени распознавания: чем лучше процесс определения антигена, тем больше антител система создает (включается процесс репродукции клеток) [14].

В течение процесса репродукции, отдельные клетки подвергаются процессу мутации, которая позволяет им иметь более высокое соответствие – аффинность к распознаваемому антигену: чем выше аффинность родительской клетки, тем в меньшей степени они подвергаются мутации, и наоборот. В иммунологии, аффинность означает силу связывания между антителом и антигеном. При развитии теории иммунных систем на задачи оптимального управления системами, данная мера определяет, насколько соответствует решение значениям критерия качества и чем меньше расстояние между векторами, тем выше аффинность. Для получения результатов применения алгоритма CLONALG использовалось евклидово расстояние [9].

Основными иммунными механизмами при разработке алгоритма являются обработка определенного множества антител, удаление антител с низкой аффинностью, созревание аффинности и повторный отбор клонов пропорционально их аффинности к антигенам [15].

Концептуальной основой применения алгоритма CLONALG для решения синтеза типовых регуляторов для ММО-систем, является формирование понятий, соответствующих теории иммунных систем и алгоритму клональной селекции. В контексте постановки задачи для системы управления (1), обобщённая форма антител соответствует вектор u аргументов $Ab = (e_1, e_2, u_1, u_2)$, и представляет собой набор решений; в качестве антигенов используется критерий качества, выраженный в виде функций: $Ag = f(e_1, e_2, u_1, u_2)$. Для синтеза типовых регуляторов для ММО-систем используем квадратичный интегральный критерий качества, соответствующий задаче Лагранжа.

Формально алгоритм клональной селекции для решения задач синтеза оптимального управления сложными системами можно представить в виде [14]:

$$\text{CLONALG} = (Ag, Ab, G, S, C, M, f, d, t),$$

где Ag – подмножество антигенов;
 Ab – популяция антител;
 G – представление пространства;
 S – оператор селекции;
 C – оператор клонирования;
 M – оператор мутации;
 f – критерий качества;
 d – количество антител, подлежащих замене новыми;

t – номер поколения.

Для решения поставленной задачи синтеза оптимального управления подсистем данный алгоритм используется в последовательности, представленной на рисунке 1 [16].

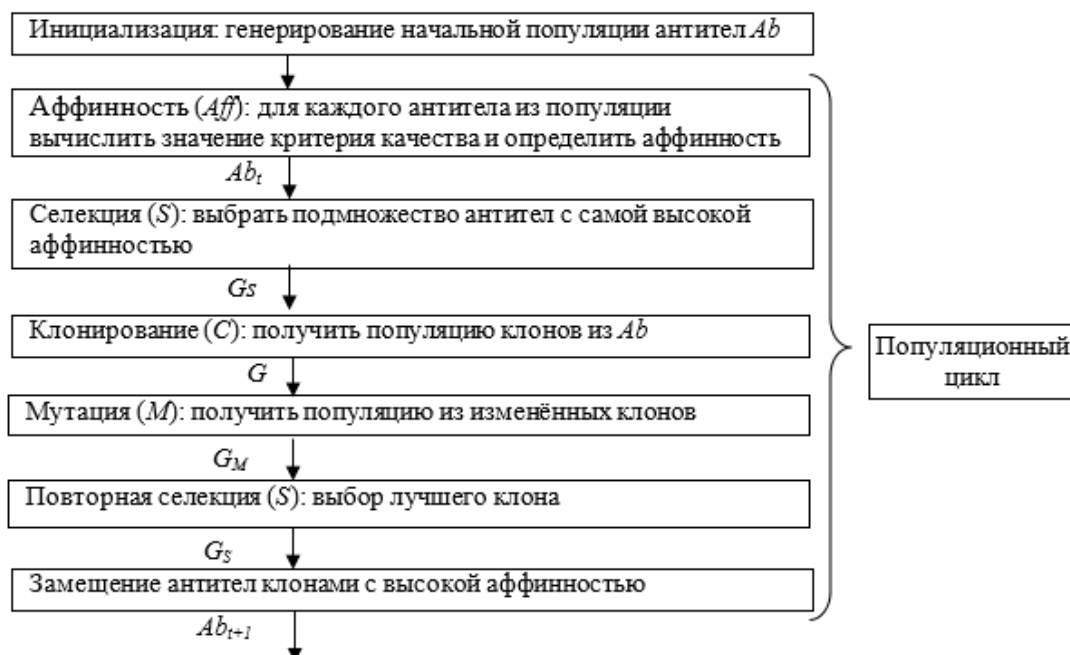


Рис.1. Алгоритм клональной селекции синтеза оптимального регулятора ММО-системы

На этапе инициализации проводится генерирование начального множества решений – коэффициентов типовых регуляторов для двух подсистем, удовлетворяющих критерию (5), случайным образом. Следующим шагом идёт выбор n антител с лучшей аффинностью, после которого осуществляется селекция антител, пригодных к клонированию:

$$N_c = \sum_{i=1}^n \text{round} \left(\frac{\beta N}{i} \right),$$

где β – множитель, контролирующий количество клонов;

N – общее количество антител;

N_c – количество клонов.

В результате клонирования антител с высокой аффинностью, получаем множество, которое подвергается случайной мутации. Если степень аффинности высокая, то частота мутаций низкая:

$$x_{i,m} = x_i + A \cdot (\text{rand}[-1,1]) \cdot (x_{\max} - x_{\min}),$$

где A – является фактором, зависящим от аффинности, и уменьшается с увеличением аффинности.

Заменяются низкоаффинные антитела новыми случайно сгенерированными антителами, представленными выражением:

$$x_i = x_{\min} + (rand[0,1]) \cdot (x_{\max} - x_{\min}).$$

В результате применения алгоритма CLONALG, имеет место множество решений с высокой степенью аффинности. Далее с помощью CLONALG отыскивается глобальный минимум функций (5) координатами, которых являются искомые параметры объекта и регуляторов.

2 Моделирование сложной системы с оптимальными регуляторами. Для того чтобы включить синтезированные регуляторы в схему МИМО-системы с взаимосвязями (3) применим процедуру развязывания. Основная идея данной процедуры заключается в применении дополнительных контроллеров для компенсации влияния взаимосвязей, с целью уменьшения их взаимодействия в контуре управления. В идеале развязывающее управление позволяет влиять только на желаемые регулируемые переменные. Динамически развязанные системы – система, у которой каждый выход зависит от одного и только одного входа. Если все перекрестные функции передачи равны нулю или пренебрежимо малы, многоконтурная система вырождается в набор независимых (развязанных) однопетлевых систем.

Схема моделирования AIS, основанная на CLONALG для МИМО-процесса в среде MATLAB, показана на рисунке 2. Блоки развязки-усиления, «Decoupling-Gain», являются элементами контроллера развязки. При полной развязке многомерная система разделяется таким образом, что в системах управления нет взаимодействий [13].

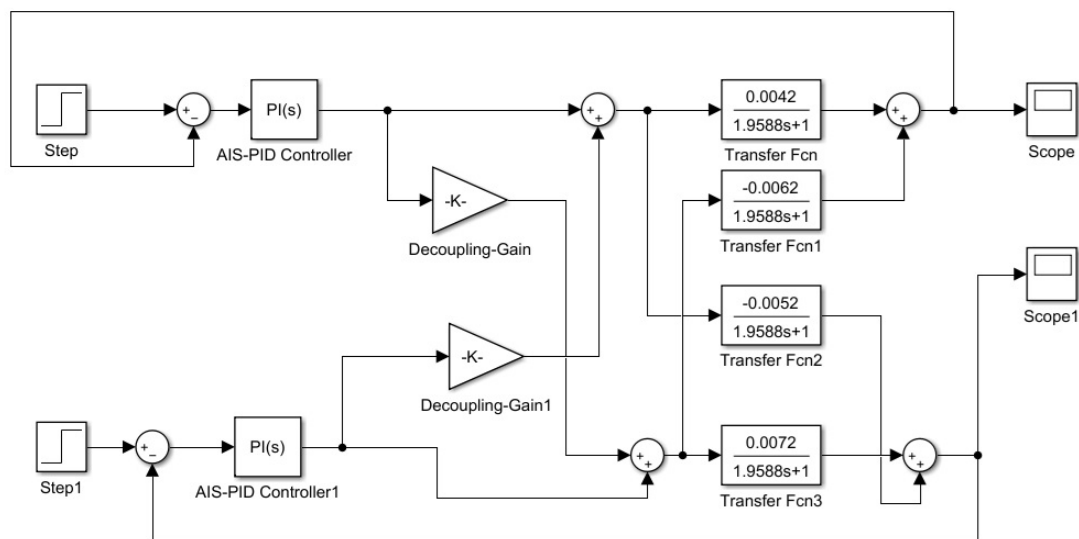


Рис.2 Схема моделирования системы в среде MATLAB

Этот этап развязывания является основной процедурой при применении CLONALG для решения задачи синтеза оптимального управления сложной системой.

3 Сравнительный анализ настройки оптимальных регуляторов на основе Smart-технологий

Сравнительный анализ результатов моделирования изолированных подсистем управления многомерной многосвязной системы управления процесса перегонки газа через дистилляционную колонну с регуляторами, настроенными на основе Smart-технологий:

- генетического алгоритма (GA);
- алгоритм оптимизации методом колонии муравьёв (ACO);
- алгоритма клональной селекции искусственной иммунной системы (AIS).

В ходе сравнения структуры алгоритмов, сделаны следующие выводы (таблица 1).

Таблица 1. Основные показатели алгоритмов Smart-технологий

Компонент	GA	ACO	AIS
Компоненты	Набор хромосом	Набор членов колонии	Набор атрибутов
Расположение компонентов	Динамическое	Динамическое	Динамическое
Структура	Дискретные компоненты	Дискретные компоненты	Дискретные компоненты
Хранение знаний	Набор хромосом	Набор феромонов	Аффинность компонентов
Движущие силы	Эволюционирование	Самоорганизация	Обучение и эволюционирование
Описание движущей силы	Генерация и отбор компонентов	Совокупность компонентов	Генерация и отбор компонентов
Взаимодействие компонентов	Обмен	Обмен	Распознавание
Взаимодействие со средой	Функция приспособленности	Функция распознавания	Функция распознавания

В ходе сравнения результатов применения алгоритмов, сделаны следующие выводы (таблица 2).

Таблица 2. Прямые оценки качества при различных алгоритмах Smart-технологий

	Компонент	GA	ACO	AIS
Результаты моделирования: 1 – лучший результат; 2 – средний результат; 3 – худший результат	Время регулирования	1	2	1
	Перерегулирование	2	1	1
	Колебательность	2	1	1
	Установившаяся ошибка	2	1	3

В соответствии с полученными результатами получены выводы об использовании применения именно искусственных иммунных систем для решения задач оптимального управления сложными системами.

Заключение. В результате проделанной работы получены результаты формирования базовых концепций применения алгоритма CLONALG для решения задач оптимального управления ММО-системой, на примере объекта нефтегазовой отрасли. В перспективе планируется получить результаты

применения алгоритмов искусственных иммунных систем для реализации процедуры развязывания ММО-систем.

Работа выполнена по гранту КН МОН РК по теме: "Разработка когнитивной Smart-технологии для интеллектуальных систем управления сложными объектами на основе подходов искусственного интеллекта" № AP05130018 (2018-2020 гг.).

Литература

1. Требухин, А. В. Методы решения оптимизационных задач с использованием биоинспирированных алгоритмов / А. В. Требухин // ДГТУ: Молодой исследователь Дона. – 2017. – Vol. 9, № 6. – С. 108-111.
2. Бобиков, А. И. Настройка весовых матриц ЗСУР регулятора с помощью биоинспирированных алгоритмов оптимизации / А. И. Бобиков // Вестник РГРТУ. – 2016. – № 55. – С. 77-83.
3. Slavov, T., Roeva, O. Application of Genetic Algorithm to Tuning a PID Controller for Glucose Concentration Control / T. Slavov, O. Roeva // WSEAS Transactions on Systems. – 2012. – Vol. 11, Issue 7. – P. 223-233.
4. Minian, F, Sabouhi, H., Hushmand, J., Hallaj, A., Khaledi, H., Mohammadpour, M. Gas turbine preventive maintenance optimization using genetic algorithm / T. Slavov, O. Roeva // International Journal of System Assurance Engineering and Management. – Springer, 2016. – Vol. 8. – P. 594-601.
5. Chinjiang, L. Optimal design of high-rise building wiring based on ant colony optimization / L. Chinjiang. – Cluster Computing. – 2018. – P. 1-8.
6. Arain, B. A., Shaikh, M. F., Harijan, B. L., Memon, T. D., Kalwar, I. H. Design of PID Controller Based on PSO Algorithm and Its FPGA Synthesization / B.A.Arain, M.F.Shaikh, B.L.Harijan, T.D.Memon, I.H.Kalwar // International Journal of Engineering and Advanced Technology (IJEAT). – 2018. – Vol. 8. – Issue 2.
7. Dasgupta, D., Yu, S., Nino, F. Recent advances in artificial immune systems: models and applications / D. Dasgupta, S.Yu, F.Nino // Applied Soft Computing Journal. – 2011. – Vol. 11, N. 2. – P. 1574–1587
8. Saleh, M., Saad, S. Artificial Immune System based PID Tuning for DC Servo Speed Control / M. Saleh, S. Saad // International Journal of Computer Applications. – 2016. – Vol. 155, No 2. – P.23-26.
9. Castro, L., Zuben, F. The Clonal Selection Algorithm with Engineering Applications / L. Castro, F. Zuben // – Las Vegas, USA: Workshop on Artificial Immune Systems and Their Applications. In Workshop Proceedings of GECCO'00. – 2000. – P. 36-37.
10. Sahraoui, M., Salem, M. Application of artificial immune algorithm-based optimisation in tuning a PID controller for nonlinear systems / M. Sahraoui, M. Salem // International Journal of Automation and Control. – 2015. – Vol. 9, No. 3. – P. 186-200.
11. Wang, M., Feng, S., He, Ch., Li, Zh., Yu, X. An Artificial Immune System Algorithm with Social Learning and Its Application in Industrial PID Controller Design / M. Wang, S. Feng, Ch. He, Zh. Li, X. Yu // Mathematical Problems in Engineering. – Vol. 2017. – 13 p.
12. Литвиненко, В. И., Фефелов, А. А., Горавский, С. П. Объектно-ориентированная реализация алгоритма клональной селекции / В. И. Литвиненко, А. А. Фефелов, С. П. Горавский // Радиоэлектроника. Информатика. Управление. – № 1. – 2003. – С. 81-88.
13. Ширяева О. И., Самигулин Т. И. Разработка SMART-системы управления сложным объектом нефтегазовой отрасли с использованием процедуры развязывания / О. И. Ширяева, Т. И. Самигулин // Вестник КазННТУ. – 2017. – № 5. – С. 50-55.

14. Кушнир Н. В., Кушнир А. В., Анацкая А. В., Катъшева П. А., Устинов К. Г. Искусственные иммунные системы: обзор и современное состояние / Н. В. Кушнир, А. В. Кушнир, А. В. Анацкая, П. А. Катъшева, К. Г. Устинов // Научные труды КубГТУ. – 2015. – № 12. – 10 с.

15. Бурлаков М. Е. Обзор базового алгоритма искусственной иммунной системы на теории негативной селекции / М. Е. Бурлаков // Уфа: Сборник статей Международной научно-практической конференции, 2014 г. – С. 29-31.

16. Li Zh., Huang H., Tan H., Zhang Y. IA-AIS: An Improved Adaptive Artificial Immune System and Its Application in Tuning of PID Controlled System / Zh. Li, H. Huang, H. Tan, Y. Zhang // Journal of Information & Computational Science. – 2008. – № 5. – P. 2193-2200.

ФОРМИРОВАНИЕ АРХИВА ИСТОРИЧЕСКИХ ДАННЫХ ГЛОБАЛЬНЫХ МОДЕЛЕЙ КЛИМАТА ДЛЯ ПРОВЕДЕНИЯ ИССЛЕДОВАНИЙ В РАМКАХ МЕЖДУНАРОДНОГО ПРОЕКТА CORDEX

**Юничева Н.Р., Нурсейтов Д.Б., Бостанбеков К.А., Алимова А.Н.,
Джамалов Д.К., Есеркепова И.Б.**
naduni@mail.ru, ndb80@mail.ru, k.bostanbekov@norlist.kz, anic2002@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

***Аннотация.** В работе показан сформированный архив данных глобальных моделей климата из списка CMIP5 (Coupled Model Intercomparison Project). В программу проекта CMIP5 включена группа специальных экспериментов по моделированию сценариев, основанных на так называемых исходных траекториях концентраций RCP (Representative Concentration Pathway). Данный архив необходим для проведения исследования в рамках международного проекта «Эксперимент по скоординированному региональному даунскейлингу климата (CORDEX)», которая нацелена на формирование смоделированных данных будущего климата Земли.*

Изучение изменений климата под воздействием антропогенных и природных факторов осуществляется на основе глобальных и региональных гидродинамических моделей. На основе глобальных моделей общей циркуляции атмосферы и океана получены проекции изменений текущего и будущего климата. Но для более детального изучения местности вместо глобальных моделей применяют мезомасштабные гидродинамические модели, которые имеют более высокое пространственное разрешение и где учитываются локальные климатообразующие факторы. Большинство исследований основанные на региональных климатических моделях обособлены и связаны с конкретными интересами, и в настоящее время нет всестороннего представления о региональных перспективных оценках изменения климата. В связи с этим в рамках Всемирной программы исследований климата (ВПИК) была создана Целевая группа по

региональному даунскейлингу климата. Результатом деятельности этой группы стало создание структуры под названием «Эксперимент по скоординированному региональному даунскейлингу климата (CORDEX)» [1]. Цель CORDEX – обеспечить структуру, доступную для широкого научного сообщества при максимальном использовании результатов. Поэтому области CORDEX охватывают большую часть земной поверхности земного шара, включая регион Центральной Азии.

Эксперименты проводятся на базе расчетов по ансамблю моделей общей циркуляции атмосферы и океана нового поколения, участвующих в международном проекте сравнения климатических моделей CMIP5. В сентябре 2008 года состоялась встреча, где приняли участие 20 групп по моделированию климата со всего мира, Рабочая группа по сопряженному моделированию (WGCM) решила продвигать новый набор скоординированных экспериментов по модели климата. Эти эксперименты составляют пятую фазу проекта сравнения климатических моделей (CMIP5) [2]. Модели нового поколения отличаются более высоким пространственным разрешением и усовершенствованными модельными описаниями компонентов климатической системы Земли. Модели климата постепенно перерастают в модели земной системы в целом, которые уже не только рассчитывают динамику атмосферы и океана, но и включают в себя детальные подмодели химии атмосферы, растительности, почвы, химии и биологии моря и других процессов и явлений, влияющих на климат. В программу проекта CMIP5 включена группа специальных экспериментов, направленных на установление причин наблюдаемых климатических изменений и оценку сравнительного значения вызывающих их внешних воздействий. Оценка влияния отдельных внешних воздействий на климат основывается на сравнении наблюдаемых изменений с ожидаемой реакцией на предполагаемые воздействия, полученной с помощью климатических моделей.

CMIP5 поддерживает стандартный набор симуляций моделей для того, чтобы:

- оценить, насколько реалистичны модели при моделировании недалекого прошлого;
- предоставлять прогнозы будущего изменения климата в двух временных масштабах: ближайшем (до 2035 года) и долгосрочном (до 2100 года и далее);
- внести понятие на некоторые факторы, ответственные за различия в модельных проекциях, включая количественную оценку некоторых ключевых обратных связей, которые связаны с облаками и углеродным циклом.

В отличие от сценариев, используемых в четвертом цикле оценок Межправительственной группы экспертов по изменению климата (IPCC), которые основывались на сценариях выбросов парниковых газов, следующее поколение экспериментов по моделированию сценариев основано на так называемых исходных траекториях концентраций (RCP) [3, 4], т.е. заданных траекториях концентраций парниковых газов на протяжении XXI века, соответствующих различным уровням стабилизации радиационного воздействия, которые будут достигнуты к 2100 г. Исходные траектории концентраций (RCP) не являются новыми, полностью интегрированными сценариями (т.е. они не являются полным пакетом социально-экономических, экологических и климатических прогнозов). Они представляют собой согласованные наборы проекций только компонентов радиационного воздействия, которые предназначены для использования в качестве исходных данных для моделирования климата, масштабирования моделей и химического моделирования атмосферы. Как таковые, они дают толчок развитию

сценария среди исследовательских сообществ, из которого можно исследовать неопределенности в отношении социально-экономического, климатического и экологического будущего. Таким образом, они представляют собой только начало параллельного процесса разработки новых сценариев для пятого Оценочного доклада МГЭИК.

RCP - это четыре независимых пути, разработанных четырьмя отдельными группами моделирования. Имеются четыре сценария с уровнями стабилизации 2.5, 4.5, 6.0 и 8.5 ppm (которые соответственно называются RCP2.5, RCP4.5, RCP6.0 и RCP8.5). В рамках CMIP5 при моделировании с использованием глобальных моделей наивысший приоритет отдан уровням RCP4.5 и RCP8.5, которые приблизительно соответствуют сценариям выбросов пятого оценочного доклада МГЭИК.

Глобальная служба предоставляет индикаторы воздействия на климат (СИ) с пространственным и временным разрешением, отличное от современной производственной цепочки, начиная с результатов глобальной климатической модели, за которыми следуют масштабирование и корректировка смещения, до воздействия на моделирование и расчеты индикаторов. Первоначальным источником данных глобальной климатической модели является проект сравнения климатических моделей (CMIP5) МГЭИК, использующий 19 моделей общей циркуляции с разрешением 2 градуса, которые доступны в хранилище климатических данных C3S. Индикаторы воздействия на климат (СИ), связанные с температурой и осадками, представлены с указанным исходным пространственным разрешением (Рис. 1). Связанные с водой данные основаны на моделировании гидрологического воздействия с использованием модели World-Wide-HYPE и предоставляются для водосборов (полигонов) со средним разрешением 1000 км² по всему миру. Временное разрешение колеблется от ежедневных временных рядов до 30-летних средних годовых или месячных значений.

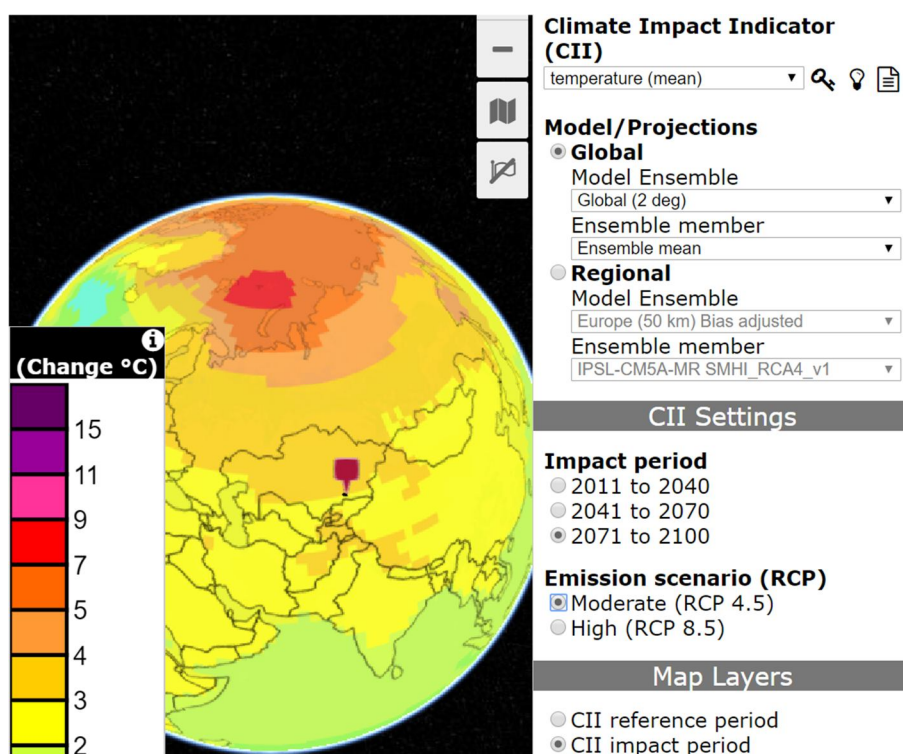


Рисунок 1 - Индикатор воздействия на климат (СИ). Средняя температура по сценарию RCP 4.5 для периода 2071 - 2100 гг.

В оптимальном случае, все региональные модельные эксперименты CORDEX должны охватывать период 1951-2100 гг. для того, чтобы включить последний исторический период и весь XXI век. Однако для многих групп расчеты могут оказаться слишком трудоемкими, чтобы выполнять моделирование CORDEX для всего вышеуказанного периода. Поэтому период 1951-2100 гг. разделен на пять 30-летних временных отрезков, и участвующие группы должны моделировать эти отрезки в следующем порядке приоритетности: 1981-2010 гг., 2041-2070 гг., 2011-2040 гг., 2071-2100 гг. и 1951-1980 гг. Первый период (1981-2010 гг.) является исходным для оценки модели и расчета изменения климата. Второй временной отрезок, охватывающий будущий период, был выбран как компромисс между потребностями сообщества с точки зрения будущего периода времени и потребностями в получении четкого сигнала изменения. Все участвующие группы должны как минимум выполнить эти два временных отрезка, чтобы иметь приемлемый ряд модельных экспериментов для анализа и взаимного сравнения.

Ключевым аспектом программы CORDEX является наличие высокопроизводительных вычислительных средств и управление большим количеством требуемых входных данных моделей, а также выходная продукция моделей и взаимное сравнения. Результаты моделирования необходимо хранить так, чтобы они были легко доступными для конечных пользователей. Также требуется стандартизация форматов (в соответствии с нормативами формата CMIP5).

Данные, рекомендуемые CMIP5 для скачивания включают в себя следующие:

1) исторические атмосферные концентрации, а также концентрации для RCP (2005-2100) и их расширение до 2300 (ECP). В общем, указаны атмосферные концентрации следующих газов: CO₂, CH₄, N₂O, все топливные газы, контролируемые согласно Киотскому протоколу (ГФУ, ПФУ и SF₆), и озоноразрушающие вещества, контролируемые согласно Монреальскому протоколу (ХФУ, ГХФУ, галлоны, CCl₄, CH₃Br, CH₃Cl).

2) исторические данные о выбросах (1850 - 2000), а также выбросы для RCP (2000-2100). В общих выбросах предусмотрены следующие газы: CH₄, SO₂, NO_x, CO, NH₃, а также BC, OC и VOC. Другие дополнительные виды, такие как C₂H₄O (ацетальдегид), C₂H₅OH (этанол), C₂H₆S (диметилсульфид), C₃H₆O (ацетон) и т.д., доступны только для исторических выбросов от сжигания биомассы.

3) исторические данные по аэрозолям (1850-2000) по следующим видам: сульфат (SO₄), нитрат аммония (NH₄NO₃), гидрофобный черный углерод (CB₁), гидрофильный черный углерод (CB₂), гидрофобный органический углерод (CB₁), гидрофильный органический углерод (CB₂), вторичные органические аэрозоли (SOA), пыль (DST01-04, от малых до больших размеров) и морская соль SSLT01-04). Кроме того, температура (T) и поверхностное давление (PS) обеспечиваются для обеспечения возможности перевода единиц (все аэрозоли указаны в кг/кг, сухая масса).

4) исторические и RCP прогнозы землепользования и связанные с ними переходы землепользования [5].

Данные глобальных моделей климата для скачивания доступны на сайте <https://dataserver.nccs.nasa.gov>. Для нашего исследования был сформирован архив данных из 19 глобальных моделей климата по трем сценариям: исторические, RCP 4.5 и RCP 8.5. Для скачивания был выбран домен Центральной Азии, координаты

которой установлены программой CORDEX. Для проведения эксперимента будут использованы три параметра: минимальная температура, максимальная температура и осадки. Сервис для скачивания предоставляет ежесуточные данные по указанным параметрам для следующих периодов: исторические с 1950 по 2005 гг., RCP 4.5 и RCP 8.5 с 2006 по 2100 гг.

Для дальнейшего сравнения и анализа данных с результатами расчета полученные данные были усреднены по месяцам с помощью скрипта, написанный на языке Python. В результате был сформирован архив данных глобальных моделей климата из списка CMIP5 для использования в исследовании в рамках проекта CORDEX.

Данная работа выполнена при поддержке грантового финансирования научных проектов КН МОН РК №AP05135848 «Моделирование будущего климата Центральной Азии в рамках международного проекта CORDEX (COordinated Regional climate Downscaling EXperiment)»

Литература

1. Giorgi F., Jones C. and Asrar G.R. Addressing climate information needs at the regional level: the CORDEX framework // WMO Bulletin. – 2009. – № 58 (3), – PP.175-193
2. CMIP5 – Coupled Model Intercomparison Project Phase 5 – Overview // <https://cmip5.llnl.gov/cmip5/index.html?submenuheader=0>: 25.05.2019
3. Global Climate Impacts // <https://climate.copernicus.eu/global-climate-impacts>: 25.05.2019
4. Wayne G.P. The Beginner’s Guide to Representative Concentration Pathways // 2013. – https://skepticalscience.com/docs/RCP_Guide.pdf: 25.05.2019
5. Wise M.A., Calvin K.V., Thomson A.M., Clarke L.E., Bond-Lamberty B., Sands R.D., Smith S.J., Janetos A.C., Edmonds J.A. Implications of Limiting CO2 Concentrations for Land Use and Energy // Science. – 2009. – №324(5931). – PP.1183-1186.

СОЗДАНИЕ ХРАНИЛИЩ ДАННЫХ ПУТЕЙ ПЕРЕДВИЖЕНИЯ

Яворский В.В., Утепбергенов И.Т., Ахмедиярова А.Т.,

Касымова Д.Т.

e-mail: aat.78@mail.ru

*Институт информационных и вычислительных технологий КН МОН РК,
Казахстан*

Аннотация.

Создана хранилища данных путей передвижения (ХДПП), которая содержит список всех возможных передвижений между каждой парой зон и их основные характеристики. Процесс генерации состоит из трех шагов. Пешие пути получены простым перебором всех пар транспортных районов. Для каждой пары определены расстояние по прямой и, используя заданную скорость пешехода, вычислены время передвижения. Если полученное значение не превышает заданное

максимальное время пешеходного передвижения, то текущий путь заносится в ХД.

Основная задача подпрограмм генерации – создание хранилища данных путей передвижения (ХДПП), которая содержит список всех возможных передвижений между каждой парой зон и их основные характеристики. Процесс генерации состоит из следующих шагов:

- 1) нахождение всех возможных пеших путей между районами;
- 2) нахождение всех возможных транспортных путей между районами;
- 3) собственно, генерация ХДПП.

Пешие пути получают простым перебором всех пар транспортных районов. Для каждой пары определяется расстояние по прямой и, используя заданную скорость пешехода, вычисляется время передвижения. Если полученное значение не превышает заданное максимальное время пешеходного передвижения, то текущий путь заносится в ХД.

Транспортные пути генерируются для каждого маршрута. Каждый реальный маршрут в городе представляется как два однонаправленных маршрута: для прямого и обратного направления. Такое разбиение связано с тем, что часто маршрут в прямом и обратном направлениях идет по-разному. На каждом таком маршруте осуществляется перебор пар остановок: 1-2, 1-3...1-N; 2-3, 2-4...2-N; ... (N-1)-N, где N – количество остановок на маршруте. Далее по графу дорожной сети определяется расстояние между парами остановок и, используя данные о скорости маршрута и интервале движения, вычисляется время передвижения. После чего данные о путях заносятся в БД, где они упорядочиваются по идентификатору первой остановки.

Генерация ХДПП осуществляется алгоритмом поиска в глубину с возвратом, реализованным в виде рекурсивной процедуры. Поиск ведется для каждой пары остановок города, идентификаторы которых и передаются процедуре поиска. Процесс представляет собой движение по путям, найденным на шаге 1 и 2 (назовем их простыми), которое начинается от первой остановки пары и повторяется для каждой следующей, к которой был осуществлен переход. При достижении тупика или установленных ограничений делается шаг назад и продолжается движение по следующему пути. Если больше путей нет, то делается еще один шаг назад и т.д. Каждый шаг движения делается следующим образом:

- 1) выбираем все простые пути, начинающиеся с текущей остановки, и начинаем с первого из них;
- 2) исследуем текущий путь:
 - а) если путь не нарушает ограничений, то исследуем его дальше, т.е. осуществляем переход по нему к следующей остановке и там, если нужно, снова делаем такой шаг (п.1);
 - б) если же ограничения нарушаются, то выбираем следующий простой путь и исследуем его (п.2);
 - в) если больше путей нет, то фиксируем окончание исследования пути (п.3);
- 3) после окончания исследования текущего пути, делаем шаг назад к остановке, с которой он начинался, там выбираем следующий простой путь и исследуем его (п.2).

Поиск заканчивается, когда в п.3 больше нельзя сделать шаг назад (т.е. мы находимся на остановке, с которой мы начинали поиск). В п.2.а после перехода к следующей остановке нужно проверить, не является ли она конечной остановкой, целью поиска. Если это так, то проверяем найденный путь передвижения и заносим

его в ХДПП. Каждый найденный путь передвижения должен удовлетворять следующим условиям:

1) **не должен ни начинаться, ни заканчиваться простым пешеходным путем** (это обуславливается тем, что мы будем искать именно такие корреспонденции, а также тем, что при обследовании фиксируется количество пассажиров, отправляющихся и прибывающих именно на транспорте);

2) **не должен содержать подряд два простых пешеходных пути** (т.к. они образуют один, более длинный простой пешеходный путь);

3) **не должен содержать простых транспортных путей больше, чем задано** (речь идет о пересадках, которых, согласно ранее проводившимся обследованиям, практически не бывает больше двух или трех в зависимости от величины города);

4) **не должен содержать «петель»** (не должен пересекаться сам с собой, не должен проходить два раза через одну и ту же остановку);

5) **не должен содержать передвижения, в которых используется один и тот же маршрут** (это очевидно, так как никто не станет садиться на один и тот же автобус два раза);

б) **время на путь не должно превышать минимальное время передвижения между этими остановками больше, чем задано** (практика показала, что пути, отклоняющиеся от минимального по времени более чем на 15-20%, почти не используются пассажирами).

В процедуре поиска путей передвижения системы RouteSearch для исключения лишних шагов и увеличения скорости проверка этих условий осуществляется не только после нахождения пути передвижения, но и при выполнении шагов поиска (см. п.2.а).

Так выполнение условия 1 достигается тем, что:

а) на первом шаге рассматриваются только простые транспортные пути;

б) любой выбранный простой пешеходный путь не должен заканчиваться в конечной остановке, цели поиска.

Условие 2 проверяется при выборе следующего простого пути: вводится флаг, разрешающий использование пешеходных путей, который устанавливается после использования пешеходного пути и убирается после использования транспортного.

Для выполнения условия 3 была введена переменная-счетчик, фиксирующая количество посадок в транспорт. И, когда остается только одна посадка до максимального количества, то сразу ищется простой транспортный путь, связывающий текущую остановку с конечной.

Условие 4 проверяется перед каждым шагом. Если этот шаг является простым пешеходным путем, то проверяется, не проходили ли мы раньше мимо конечной остановки этого пути (проверяются не только остановки, на которых были пересадки, но и те остановки, через которые проходят). Для транспортного пути так проверяется не только конечная остановка, но и каждая остановка, через которую проходит этот путь.

Чтобы выполнялось условие 5, перед использованием каждого простого транспортного пути проверяется, использовался ли ранее этот маршрут.

Для того, чтобы найденные пути удовлетворяли требованиям времени (условие б), перед переходом по найденному простому пути, проверяется, не отклоняется ли уже сейчас этот незаконченный путь передвижения по времени от минимального из найденных раньше. Если он уже на этом шаге (не последнем) сильно отклоняется по времени, то его дальнейшее исследование бесполезно.

Также для каждого найденного и удовлетворяющего условиям пути передвижения проверяется, не является ли он минимальным. Если же находится новый минимальный путь, то все найденные ранее пути проверяются уже относительно нового минимального пути передвижения.

После того, как ХДПП создана и все данные внесены в систему, начинается главная процедура анализа – получение МПК. Она осуществляется в два этапа:

- 1) получение первого приближения МПК;
- 2) составление системы уравнений и получение МПК с заданной точностью.

Первое приближение МПК находится методом энтропийного моделирования с использованием данных об обследовании, средних данных по городу и различных коэффициентов, полученных во время предшествующих обследований и приведенных в специальной литературе. Этот метод заключается в распределении пассажиропотоков, отправляющихся из транспортных районов города (остановок) в другие районы в соответствии с функцией тяготения, отражающей трудность передвижения между парой районов. Для реализации была выбрана экспоненциальная модель функции тяготения вида $e^{-l/l_{cp}}$, где l – расстояние между транспортными районами, а l_{cp} – средняя дальность поездки по городу. Но после такого распределения сумма прибывающих корреспонденций по районам не будет соответствовать данным обследования. МПК получается как бы перекошенной, т.е. мы знаем, какими должны быть суммы ее строк и столбцов, но в полученной матрице с этими значениями совпадают только суммы строк (отправления из транспортных районов). Для выравнивания матрицы (достижения совпадения также и сумм столбцов с полученными данными) был применен метод балансировки корреспонденций, в котором расчет ведется в несколько итераций с использованием балансировочных коэффициентов до достижения необходимой точности выравнивания МПК.

Система уравнений, отражающая взаимосвязь пассажиропотоков на маршрутах с корреспонденциями между транспортными районами города, составляется на основании сгенерированных возможных путей передвижения и данных обследования. Так как получающаяся система содержит очень большое количество уравнений и переменных (примерно $4*N$ уравнений и $N*(N-1)$ переменных, где N – количество остановок в городе), то обычные способы решения для этой системы не подходят. А когда система несовместна, получение точного решения и вовсе невозможно. В системе RouteSearch она решается итерационным методом последовательных приближений, в котором в качестве первого приближения берется МПК, полученная на первом этапе. На каждой итерации осуществляется движение точки, которую представляет собой МПК в многомерном пространстве ($N*(N-1)$ измерений), по перпендикуляру последовательно ко всем гиперплоскостям, отображающим в пространстве систему уравнений. В результате получаем точку, более близкую к решению системы или к области гиперпространства, точки которой наиболее полно удовлетворяют всем уравнениям (она является многогранником, образуемым пересечением гиперплоскостей, описываемых уравнениями системы). Для достижения большей точности, шаг немного уменьшается в каждой последующей итерации. После окончания итерации вычисляется отклонение получившейся точки от предыдущей. Процесс заканчивается, когда рассчитанная во время текущей итерации МПК отличается от ранее определенной на величину, не большую заданной точности. Величина уменьшения шага и точность задаются пользователем в окне диалога.

Полученная в результате анализа МПК, может быть экспортирована в виде таблицы для анализа и построения диаграмм. Небольшая дополнительная

подпрограмма RouteSearch позволяет получить укрупненные результаты по зонам города, объединяющих несколько транспортных районов, осуществляя суммирование строк и столбцов матрицы, соответствующих остановкам, входящим в определенный транспортный район. В дальнейшем МПК используется для корректировки и формирования транспортных маршрутов.

Перед началом формирования маршрутов в системе RouteSearch формируется база данных кратчайших путей между всеми парами транспортных районов (БДКП). Поиск ведется по дорожной сети города без учета существующих маршрутов. Алгоритм поиска похож на алгоритм формирования ХДПП (он отличается только требованиями к найденным путям). Для каждой пары районов может быть найдено несколько равноценных кратчайших путей, отклоняющихся от минимального на заданную величину. Формирование нового маршрута происходит так:

1) из полученной МПК выбирается пара наиболее взаимосвязанных транспортных районов, т.е. районов, между которыми величина корреспонденции максимальна;

2) в БДКП находятся пути, проходящие через эти районы;

3) из остановок, лежащих на продолжениях найденных путей, выбирается третья остановка нового маршрута. Критерии выбора этой остановки могут быть различными, но их смысл заключается в следующем: она должна быть наиболее связанной с двумя уже выбранными. Можно искать максимальную сумму корреспонденций, перемещающихся в нее из двух начальных остановок;

4) из всех кратчайших путей, подобранных для двух первых остановок, выбираются те, которые содержат также и третью. Далее таким же образом ищется четвертая остановка (шаг 3) и т.д.

Формирование маршрута заканчивается, когда на продолжения оставшихся путей на шаге 3 больше нет остановок. После этого путем распределения потенциальных корреспонденций между транспортными районами определяются пассажиропотоки на перегонах нового маршрута и другие параметры, на основании которых составляется паспорт маршрута, содержащий следующие данные:

а) длину (определяется длиной кратчайшего пути, на котором был построен маршрут);

б) коэффициент непрямолинейности (отношение длины маршрута к расстоянию по прямой между начальной и конечной остановкой маршрута);

в) объем перевозок (чел/час);

г) количество транспортных единиц;

д) интервал движения.

Полученный паспорт маршрута выдается для оценки экспертом. Если маршрут его удовлетворяет, то пассажиропотоки на нем считаются распределенными и вычитаются из МПК (процедура корректировки) и система приступает к формированию нового маршрута. В противном случае маршрут запрещается и пара остановок, приведшая к его генерированию, больше не рассматривается. Формирование маршрутов заканчивается, когда распределены все потенциальные корреспонденции.

По окончании процедуры синтеза можно посчитать характеристики новой транспортной системы города, такие как средняя дальность поездки, коэффициент пересадочности, среднее время ожидания пассажиром транспорта, загруженность транспорта и другие.

Схема алгоритма работы системы RouteSearch приведена на рисунке 1.

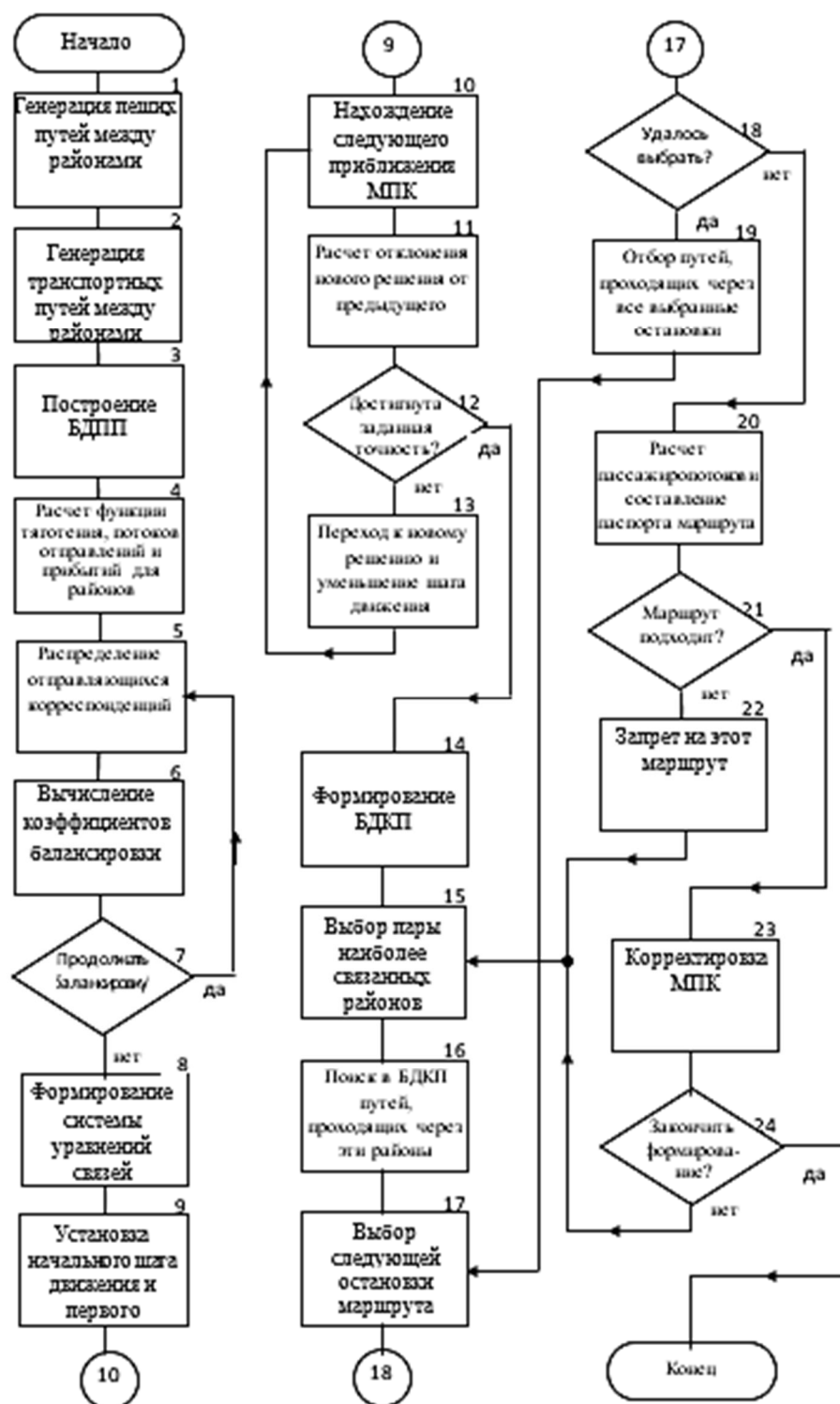


Рис. 1 Укрупненная блок-схема работы системы RouteSearch

Ее можно разбить на следующие части в соответствии с задачами, решаемыми системой, которые были описаны выше:

Блоки 1-3: ввод и генерация данных для анализа транспортной системы города.

Блоки 4-7: Получение МПК методом энтропийного моделирования (шаг 1 процедуры анализа транспортной системы).

Блоки 8-13: Получение МПК с использованием системы уравнений связей пассажиропотоков с корреспонденциями (шаг 2 процедуры анализа).

Блоки 14-24: формирование новых маршрутов по результатам анализа транспортной системы.

Дадим подробное описание системы по блокам:

Блок 1 - производится нахождение всех возможных пеших путей между районами с проверкой ограничений (шаг 1 генерации ХДПП).

Блок 2 - производится нахождение всех возможных транспортных путей между районами с использованием одного маршрута (шаг 2 генерации ХДПП).

Блок 3 - производится поиск всех путей передвижения между транспортными районами по алгоритму, описанному выше, в соответствии с требованиями, предъявляемыми к найденным путям (шаг 3 генерации ХДПП).

Блок 4 - рассчитывается экспоненциальная функция тяготения для всех пар транспортных районов с использованием координат остановок и средней дальности поездки по городу, которая сохраняется в виде матрицы для дальнейшего использования, а также находятся потоки отправления и прибытия для транспортных районов путем суммирования строк и столбцов МПК.

Блок 5 - производится распределений пассажиропотоков отправления из районов в соответствии с вероятностью поездки, определяемой рассчитанной функцией тяготения, с использованием текущих значений коэффициентов балансировки (на начальном шаге они принимаются равными единице).

Блок 6 - вычисляются новые коэффициенты балансировки, показывающие степень выравнивания МПК, полученной распределением на предыдущем шаге.

Блок 7 - вычисляются новые коэффициенты балансировки, показывающие степень выравнивания МПК, полученной распределением на предыдущем шаге

Блок 8 - формирование системы уравнений связи пассажиропотоков с корреспонденциями.

Литература

1. Жуковский О.И. Хранилища данных. – Томск: ТУСУР, 2016. – 145 с.
2. Петрунин, Ю. Ю. Информационные технологии анализа данных. Data Analysis / Ю.Ю. Петрунин. - М.: КДУ, 2010. - 292 с.
3. Ибатов М.К., Яворский В.В., Чванова А.О., Ключева Е.Г. Формирование хранилищ данных транспортной инфраструктуры транспортной системы города // Материалы научной конференции ИИВТ КН МОН РК «Инновационные IT и Smart-технологии», посвященной 70-летию юбилею профессора Утепбергенова И.Т. — Алматы: ИИВТ МОН РК, 2019. — С. 131-140.
4. Яворский В.В., Утепбергенов И.Т., Ахмедиярова А.Т. Models of analysis of distribution of passenger traffics in routed transport systems // Материалы международной конференции XIII Balkan Conference on Operational Research (BALCOR 2018). — Belgrade, Serbia, 2018.

РАЗРАБОТКА РЕКОМЕНДАТЕЛЬНОЙ СИСТЕМЫ ПО УЛУЧШЕНИЮ КАЧЕСТВА ШКОЛЬНОГО ОБРАЗОВАНИЯ НА ОСНОВЕ ИНТЕРПРЕТАТОРОВ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ

Якунин К.^{1*}, Касымжанов Б.¹, Кучин Я.¹, Сымагулов А.^{1,2},
Мурзахметов С.^{1,2}, Булдыбаев Т.³, Мухамедиев Р.^{1,2}
e-mail: yakunin.k@mail.ru

¹ *Институт информационных и вычислительных технологий МОН РК, Алматы, Республика Казахстан*

² *Казахский национальный исследовательский технический университет им. К.И. Сатпаева, Алматы, Республика Казахстан*

³ *АО Информационно-аналитический центр МОН РК, Астана, Республика Казахстан*

Аннотация. В данной работе рассматривается способ применения интерпретаторов моделей машинного обучения для разработки рекомендательной системы по улучшению качества школьного образования. Разработана система показателей Q_{edu} , в основе которой данные о качестве образования. Каждое среднее учебное заведение описывается набором параметров в регрессионной модели. Оценка влиятельности указанных параметров позволяет выработать рекомендации по улучшению позиции отдельного учебного заведения в рамках указанной модели. Кратко проанализированы несколько интерпретаторов *Treeinterpreter*, *LIME*, *SHAP*. Последний применен на практике. Полученные результаты демонстрируют возможность применения данного интерпретатора как основы рекомендательной системы для средних учебных заведений.

Ключевые слова: среднее образование, *SHAP*, качество образования, *machine learning*

1. Введение

В 2015 году на Всемирном форуме по образованию ЮНЕСКО более 180 стран мира, включая Казахстан, приняли новую концепцию развития образования до 2030 года. Суть этой новой концепции отражена в четвертой цели ООН в области устойчивого развития «Качественное образование для всех» - обеспечить инклюзивное и справедливое качественное образование и продвигать возможности обучения на протяжении всей жизни для всех. Для достижения ЦУР-4 необходимо решить 7 задач: всеобщее начальное и среднее образование, развитие детей в раннем возрасте и всеобщее дошкольное образование, равный доступ к техническому / профессиональному и высшему образованию, соответствующие навыки для достойной работы, гендерное равенство и инклюзивность, всеобщая молодежная грамотность, образование для устойчивого развития и глобальное гражданство. Страны всего мира должны приложить все усилия для достижения этой цели в области устойчивого развития к 2030 году. Новая концепция

* Инчхонская декларация и Рамки действий для достижения четвертой цели в области устойчивого развития, ЮНЕСКО, 2015

образования была полностью отражена в стратегических документах по развитию образования в Казахстане.

На данный момент в Казахстане имеется более 7 тысяч школ, в которых обучается 3,2 миллиона учеников и работает 340 тысяч учителей. Для своевременного выявления проблем проводится образовательный мониторинг в Национальной образовательной базе данных (НОБД). НОБД содержит более 10 миллионов записей данных для всех учебных заведений.

В связи с этим появилась возможность применения методов машинного обучения для определения наиболее влиятельных параметров для каждой школы, выявления региональных различий в качестве образования.

Для оценки предметных областей описываемых массивами данных возможно применение методов машинного обучения. В частности, полезно применение методов кластерного анализа, среди которых можно перечислить *isometric mapping* (ISOMAP) [1], *locally linear embedding* (LLE) [2], *t-distributed stochastic neighbor embedding* (t-SNE) [3], *kernel principal component analysis* (KPCA) [4], and *multidimensional scaling* (MDS) [5]. Указанные методы помогают оценить региональный дисбаланс, "завышение" или "занижение" оценок, выявить организации с аномальными значениями параметров.

С другой стороны, методы обучения с учителем, в частности регрессионные модели [6] также позволяют выявить аномальные объекты, но, в дополнение, построенная модель может быть проанализирована с точки зрения выявления наиболее влиятельных параметров. То есть модель желательно проинтерпретировать с целью ответа на вопрос "Почему получился тот или иной результат классификации или регрессии?".

В частности, известны примеры обучения моделей, когда особенности текущего набора данных приводили к правильным результатам классификации на конкретном корпусе при совершенно не правильной модели распознавания [7]. В указанной работе описывается задача распознавания текстов атеистического и религиозного содержания, когда оказалось, что модель правильно отличала первые от вторых, но не по содержанию, а по наличию слова "posting", которое встречалось только 2 раза в текстах религиозного содержания, хотя было весьма частым (21%) в тренировочном наборе.

Другими словами, необходим алгоритм, который поможет ответить на вопрос "Какие переменные и в каких пределах влияют на предсказание?". Это позволяет нам удостовериться, что модель не переобучена и что она формирует результат не случайным образом.

Идеальный вариант интерпретации может быть, например таким, когда общий ответ модели есть сумма значений параметров (X) умноженных на коэффициенты модели (Theta):

$$S = \text{theta}_0 * x_0 + \text{theta}_1 * x_1 + \dots + \text{theta}_i * x_i$$

В этом случае легко понять, как влияют отдельные признаки (параметры) на результат. В плане объяснения результата линейные регрессионные модели являются легко интерпретируемыми.

Однако такой вариант возможен не всегда. В случае сложных нелинейных зависимостей и значительной корреляции свойств, модель представляет собой так называемый "черный ящик", то есть своего рода закрытую систему, которая получает на вход данные, "скрытно" обрабатывает их и выдает результат. Что

делается внутри "черного ящика", как модель связала входные данные с результатом остается непонятным.

Для превращения такого "черного ящика" в "белый", то есть для понимания того как влияют отдельные параметры на предсказание модели в задачах регрессии и классификации полезными инструментами оказывается Treeinterpreter для деревьев решений, LIME (Local Interpretable Model-agnostic), DeepLIFT, похожий на него Layer-Wise Relevance Propagation и наиболее свежий метод, описанный в 2017 году - SHAP (SHapley Additive exPlanations).

Информация о LIME приведена в [8,9]. Имеется проект на Github [10]. Приложения метода описаны в [11, 12]. LIME не зависит от модели машинного обучения и может применяться к любой модели, которая поддерживает метод predict_proba из библиотеки sklearn, или чьи результаты могут быть представлены в виде вероятности (например, softmax). Вкратце суть метода заключается в том, что модель предполагает линейную зависимость вывода классификатора от изменения свойств объекта в небольшом диапазоне. Создается интерпретируемая модель, например, линейная, которая обучается на небольших возмущениях параметров оцениваемого объекта (добавляется "шум"), добиваясь хорошей аппроксимации исходной модели в этом небольшом диапазоне. "Шум" создается по-разному для разных видов задач, например, путем добавления, собственно шума к непрерывным функциям, удаления слов или скрытия части изображений. Однако эта относительная простота интерпретатора LIME является причиной его основного недостатка. Для сложных, трудно интерпретируемых моделей линейные аппроксимации могут быть неадекватными, даже если анализируется небольшая область вокруг значений параметров. В таких случаях LIME может давать ошибочные рекомендации.

В свою очередь SHAP [13, 14] качественно отличается от простого поиска корреляций, в виду того, что он использует модель для получения знаний о нелинейных и немонотонных взаимозависимостях параметров при влиянии на конечный результат. Метод предназначен для работы в случае существенной зависимости между параметрами. В общем случае, метод требует перетренировки модели на всех подмножествах $S \subseteq n$, где n множество всех признаков. Метод присваивает значение важности каждому свойству, которое отражает влияние на прогнозирование модели при включении этого свойства. Чтобы вычислить этот эффект, модель $f(S \cup \{i\})$ обучается с этим свойством, а другая модель $f(S)$ обучается с исключенным свойством. Затем прогнозы этих двух моделей сравниваются на текущем входном сигнале $f(S \cup \{i\}(x_S \cup \{i\})) - f_S(x_S)$, где x_S представляет значения входных свойств в наборе S . Поскольку эффект исключения признака зависит от других признаков в модели, то указанная разность вычисляется для всех возможных подмножеств $S \subseteq n \setminus \{i\}$. Затем вычисляется средневзвешенная величина всех возможных различий:

$$\varphi_i = \sum_{S \subseteq \{1,2,\dots,n\} \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} (f(S \cup \{i\}) - f(S)), \quad (1)$$

Это и есть оценка важности (влияния) свойств (параметров) на оценку модели. В целом, это подход, основанный на теории игр, который по выводам авторов алгоритма обеспечивает общность интерпретации и пригодность для широкого круга методов машинного обучения.

Последний метод применен в разработке пилот проекта рекомендательной системы по улучшению качества школьного образования.

2. Данные

Источником выбора показателей общего среднего образования послужили административные данные МОН РК. В Казахстане сбор административных данных в области образования осуществляется в рамках Национальной базы данных образования (НОБД). НОБД содержит данные о более чем 20 тысячах образовательных организаций. Источником данных об окончательных результатах обучения и распределении стипендий (грантов) является Национальный центр тестирования (НЦТ). Сбор данных проводился за 2018-2019 учебный год, данные о результатах обучения модели за 2017-2018 учебный год.

НОБД содержит следующие данные для каждой школы:

- О местоположении, типе, внутреннем распорядке
- О материальном обеспечении (библиотека, питание, отопление, ремонт, площадь помещений и пр.)
- О кадровом составе (количество, звания, опыт и пр.)
- О контингенте (язык обучения, дети с особыми потребностями, половой и национальный состав и пр.)
- доступность школьного образования (клубы, кружки, транспорт, питание и т. д.)
- результаты обучения (успеваемость).

Принимая во внимание доступные переменные в базах данных, было идентифицировано и собрано около 200 переменных общего среднего образования, в том числе около 10, полученных в результате feature engineering.

На основании полученных данных были спрогнозированы и рассчитаны 4 показателя качества образования Q-Edu:

Q-Edu-1 - Процент выпускников, поступивших на грант на лучшие специальности в лучшие университеты

Q-Edu-2 - Количество элитных достижений (AltynBelgi, республиканские и международные олимпиады)

Q-Edu-3 - Доля выпускников, которые набрали проходной балл на ЕНТ (единое национальное тестирование)

Q-Edu-4 - Процент выпускников, продолживших обучение

Q-Edu-1,2,3 были рассчитаны на основе имеющихся данных, и для формирования Q-Edu-1 (доля выпускников, поступивших на грант на лучшие специальности в лучшие университеты) были использованы следующие дополнительные данные:

1) Национальный рейтинг университетов - <http://egov.kz/cms/ru/articles/2Fbestuniinkz30>

2) Проходные баллы на грант по специальностям за 2017 год - <https://univision.kz/prohodnye-bally-na-grant-2018.html>

В результате было отобрано 9 университетов, которые занимают высшие позиции в национальных рейтингах:

- КазГУ
- КБТУ
- МУИТ
- СДУ
- КазНМУ им. С.Д. Асфендиярова
- Евразийский национальный университет имени Л.Н. Гумилева
- КазНИТУ им. К.И. Сатпаева

- Карагандинский экономический университет Казпотребсоюза
- КазНПУ им. Абая

3. Результаты

На основе имеющихся данных и рассчитанных показателей качества Q-edu-1-4 были построены и обучены модели машинного обучения для прогнозирования показателей качества, на основе имеющихся предварительно обработанных данных. В результате экспериментов была применена модель, основанная на деревьях решений GradientBoosting, реализованная в пакете scikit-learn. Построенные модели состояли из 100 деревьев, каждое с максимальной глубиной 5.

В результате были получены модели с коэффициентами детерминации, показанными ниже. Коэффициент детерминации - это показатель качества, который сравнивает обученную модель с моделью, которая всегда прогнозирует математическое ожидание, для которого значение 0 соответствует модели, прогнозирующей математическое ожидание, а значение 1 соответствует идеальной модели.

- Q-Edu-1 - 0.98
- Q-Edu-2 - 0.84
- Q-Edu-3 - 0.77
- Q-Edu-4 - 0.94

Затем алгоритм SHAP, описанный выше, использовался для объяснения полученных моделей и ранжирования значимости факторов. В результате результаты были получены в виде, показанном на рисунке 1.

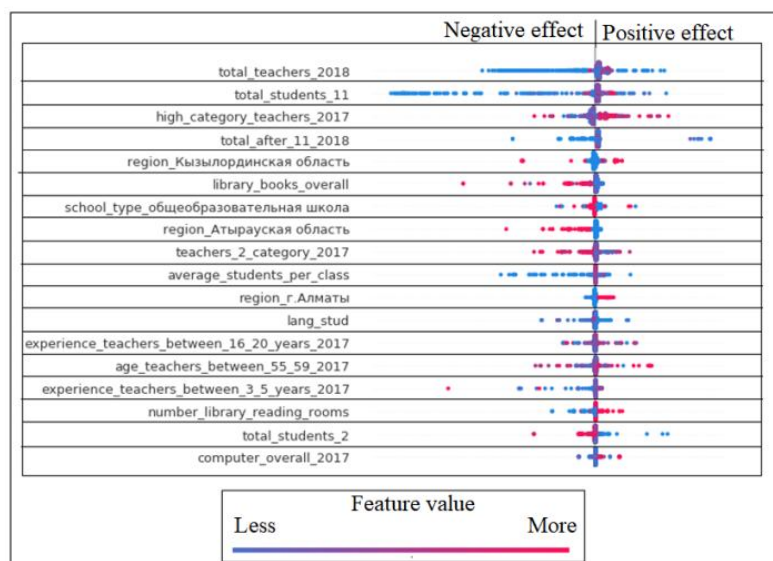


Рисунок 1. Влияние факторов на показатель Q-edu 3 (Доля выпускников, которые набрали проходной балл на ЕНТ)

Давайте рассмотрим рисунок 1 и дадим его интерпретацию. Каждая строка на рисунке отражает определенный фактор, факторы отсортированы по значимости в порядке убывания. Каждая точка представляет отдельную школу, и ее положение вдоль горизонтальной оси показывает, насколько отрицательно (слева) или положительно (справа) этот фактор повлиял (в данном случае) на показатель Q_edu

3 этой школы. В то же время цвет точки указывает на значение фактора - красные точки имеют высокое значение этого фактора (выше, чем среднее значение по выборке), а синий цвет соответственно - низкое значение. Для категориальных / бинарных факторов цвет точек означает соответственно принадлежность к категории или наличие / отсутствие фактора.

Рассмотрим пример фактора «Region_Almaly» - принадлежность к этому региону оказывает положительное влияние на показатель Q-edu 3 в обученной модели, и, соответственно, отсутствие принадлежности к этому региону в некоторых случаях приводит к меньшему отрицательному значению в абсолютной величине.

Следует отметить, что высокая степень влияния не означает автоматически наличие причинно-следственных связей, а только указывает на то, что модель использует тот или иной фактор как наиболее подходящий для статистического разделения по целевым показателям или, другими словами, как фактор, имеющий наибольшую информационную энтропию для задачи.

4. Выводы

Была предложена многомерная оценка качества Q_edu_school.

Предлагаемая система рекомендаций определяет ключевые параметры, влияющие на ключевые показатели качества, включая региональные дисбалансы и региональные особенности. Есть возможность проанализировать наиболее важные параметры для всей выборки, отдельных школ, а также провести анализ взаимозависимости различных факторов. Количество показателей Q-edu может быть увеличено, а их содержание должно быть согласовано со специалистами в области среднего образования. Чтобы более точно сопоставить показатели с реалиями школьного образования, предлагается ввести экспертную оценку веса параметров. Расширение набора данных (по годовым срезам) позволит оценить динамику изменения показателей Q-edu.

Работа финансировалась за счет гранта №BR05236839 Министерства образования и Науки Республики Казахстан «Разработка информационных технологий и систем для стимулирования устойчивого развития личности как одна из основ развития цифрового Казахстана».

Литература

1. Tenenbaum J B, De Silva V, Langford J J S 2000 A Global Geometric Framework for Nonlinear Dimensionality Reduction *Science* vol 290 no 5500 pp 2319-2323
2. Roweis S T, Saul L K J S 2000 Nonlinear Dimensionality Reduction by Locally Linear Embedding *Science* vol 290 no 5500 pp 2323-2326
3. Der Maaten L V, Hinton G E 2008 Visualizing High-Dimensional Data Using t-SNE *Journal of Machine Learning Research* pp 2579-2605
4. Scholkopf B, Smola A J, Muller K J N C 1998 Nonlinear component analysis as a kernel eigenvalue problem *Neural Computation* vol 10 no 5 pp 1299-1319
5. Borg I, Groenen P J F J J o E M 2003 Modern Multidimensional Scaling: Theory and Applications *Journal of Educational Measurement* vol 40 no 3 pp 277-280
6. Muhamedyev R 2015 Machine learning methods: An overview *CMNT* 19(6) pp 14-29

7. Ribeiro M T, Guestrin S S C 2016 Local Interpretable Model-Agnostic Explanations (LIME): An Introduction A technique to explain the predictions of any machine learning classifier
8. Ribeiro M T, Singh S, Guestrin C 2016 Why should i trust you?: Explaining the predictions of any classifier *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* C 1135-1144
9. Hulstaert L, 2018 Understanding Model Predictions with LIME. Learn about Lime and how it works along with the potential pitfalls that come with using it - <https://www.datacamp.com/community/tutorials/understanding-model-predictions-lime>
10. Lime - <https://github.com/marcotcr/lime> 24 mar 2019
11. Local Interpretable Model-Agnostic Explanations (LIME): An Introduction - <https://www.oreilly.com/learning/introduction-to-local-interpretable-model-agnostic-explanations-lime> 24 mar 2019
12. Koehrsen W A Complete Machine Learning Walk-Through in Python: Part Three Interpreting a machine learning model and presenting results- <https://towardsdatascience.com/a-complete-machine-learning-walk-through-in-python-part-three-388834e8804b> 24 mar 2019
13. Lundberg, Scott M, Lee S I 2017 A unified approach to interpreting model predictions *In Advances in Neural Information Processing Systems* pp 4765-4774
14. Shap - <https://github.com/slundberg/shap> 24 mar 2019

СОДЕРЖАНИЕ

Aitimov A., Turan C., Dusebekov Zh., Amirgaliyev Ye.	Reliable gesture recognition methods	4
Amirgaliyev Ye., Jantayev R.	Methods for accurate gesture recognition	9
Mladenovic N., Krassovitskiy A., Mussabayev R.	Decomposition approach for clustering problem	15
Nurumov K., Baimakhanbetov M., Buldybayev T., Akoyeva I., Ospanova U.	An approach for detecting sentiments in mass media publications using content specific informative features	21
Toleu A., Tolegen G., Mussabayev R.	Initial exploration for Topic modeling with word sense	31
Калимолдаев М.Н., Абдилдаева А.А., Жукабаева Т.К., Галиева Ф.М.	Қазақстанның электр энергетикалық жүйесі үшін математикалық модель құру	45
Қалимолдаев М.Н., Мамырбаев О.Ж., Мекебаев Н.О., Тұрдалыұлы М.	Машиналық оқуды қолдануда дауыстың гендерлік жіктелінуі	51
Капалова Н.А., Самрат С.М.	Ақпараттық қауіпсіздік қауіп-қатерлерінің жіктелуі	57
Капалова Н.А., Хаумен А., Сулейменов О.Т.	Зиянды бағдарламалардың жұмысына талдау жасау	61
Сейтжаппар М.А., Рахимова Д.Р.	Қазақ тілінің автоматтандырылған маркерлік корпусын әзірлеу	66
Сейтқали Ғ.Т.	Автоматтандырылған жүйелердегі ақпаратты қорғау әдістеріне шолу	74

Wojcik Waldemar, Калижанова А.У., Кашаганова Г.Б., Оразбеков Ж.Т., Арсланова С., Искакова М.Т.	Разработка интеррогаторов для датчиков на основе волоконной решетки Брэгга	77
Алиаскар М.С., Айпанов Ш.А., Тусупова С.А., Карымсакова Н.Т., Амирханов Б.С.	Биометрическая идентификация человека по отпечаткам пальцев	83
Амирханова Г.А., Дуйсенбаева А.Ж.	О применении одного способа кластеризации для поиска минимальной энергии плоского кристалла	88
Ахмеджанов А.Х., Ахметов Н.Д., Караданов Т.К.	Изменения содержаний малых газов в атмосфере Казахстана по данным дистанционного зондирования	94
Ахметжанов М.А., Алпысбай Н.Э., Даулетбек Е.Т.	Регуляризация движения робота по заданной траектории	99
Ахметжанов М.	Движение рабочего инструмента для робота 3RPR по заданной траектории с минимальной суммарной длиной штанг	105
Бадаев С.А., Калмурзаев Б.С.	О специальных классах позитивных предпорядков	118
Бакирова Э.А., Асанова А.Т.	Об аппроксимации задачи управления для интегро-дифференциальных уравнений параболического типа	121
Барахнин В., Дуйсенбаева А., Кожемякина О., Кучин Я., Якунин К., Мухамедиев Р.	Изменение публикационной активности в области обработки естественного языка	130
Бегалиева К.Б., Аршидинова М.Т., Кудайкулов А., Ташев А.А.	Автоматизация исследования термо- напряженного состояния стержня ограниченной длины при воздействии локального бокового теплового потока	135

Бияшев Р.Г., Бегимбаева Е.Е., Рог О.А.	Разработка автоматизированной системы защиты информации в процессе трансграничного обмена	148
Варенников А.В.	Краткий обзор основных принципов разработки систем управления криптографическими ключами	154
Дауылбаев М.К., Атахан Н.	Краевые задачи для сингулярно возмущенных интегро-дифференциальных уравнений	160
Еримбетова А.С., Мурзин Ф.А.	Разработка лингвистического и алгоритмического обеспечения технологий информационного поиска	167
Зиятбекова Г.З., Мазаков Т.Ж., Кисала П.	Обзор литературы по проблеме численного моделирования речных потоков	174
Калимолдаев М.Н., Тынымбаев С.Т., Магзом М.М., Тананова Д.Д.	Программа расчета параметров непозиционной системы счисления	179
Капалова Н., Хаумен А.	Модель управления ключами алгоритмов шифрования данных	186
Кубеков Б.С., Утегенова А.У., Науменко В.В., Аленова Р.А.	Компетентностная модель представления знаний и концепция Smart-контракта в проектировании образовательной программы	190
Кудайкулов А.К., Ташев А.А., Аршидинова М.Т., Бегалиева К.Б.	Разработка численного метода учета наличия локальных поверхностных теплообменов в стержнях переменного сечения	201
Кунелбаев М.М. Сундетов Т.Р., Айткулов Ж., Черикбаева Л.Ш., Орманов Т.А.	Разработка алгоритма управления и программного обеспечения солнечного котроллера	211
Литвиненко Н.Г., Литвиненко А.Г., Мамырбаев О.Ж., Шаяхметова А.С.	К вопросу о поколениях в байесовских сетях	217

Макаренко Н.Г., Пак А.А., Панина А., Каримова Л.М., Пазовский Н.С.	Обнаружение онкологических паттернов методами дискретной дифференциальной геометрии	225
Мейрамбеккызы Ж., Аманбай А., Мусабаев Р.	Автоматическое формирование тематических словарей социально-значимых понятий	233
Мусабаев Р., Мухамедиев Р., Кучин Я., Сымагулов А., Якунин К., Мурзахметов С.	Метод мультимодальной оценки средств массовой информации на основе тематической модели корпуса текстов	239
Мусабаев Р.Р., Сейткали Д.	Реферирование текстового документа с помощью Word Mover's Distance и извлеченных ключевых слов документа	247
Мустафин С.А., Арсланов М.З., Шангитбаев Ж.К.	Определение класса прочности закладки, основанное на методах распознавания образов	253
Нургазин А.Т.	Реализация и опускание типов в форсинг- структурах	259
Сақан Қ., Алғазы К.	Программная реализация алгоритма симметричного блочного шифрования «QAMAL» для исследования его криптографических свойств	264
Самбетбаева М.А., Батырханов А.Г.	Технология создания тематических интеллектуальных ресурсов для поддержки научно-образовательной деятельности на основе онтологического проектирования	270
Самигулина Г.А., Самигулина З.И.	Модифицированный алгоритм оптимизации серых волков для Smart-технологии прогнозирования и управления сложными объектами	280
Самигулина Г.А., Масимканова Ж.А.	Разработка модифицированного кооперативного алгоритма роя частиц с весом инерции для многофункциональной искусственной иммунной системы	286

Суйменбаев Б.Т., Трушляков В.И., Ермолдина Г.Т., Суйменбаева Ж.Б., Батышев А.М.	Разработка информационно-аналитической системы районов падений космодрома Байконур на основе ГИС-технологий для выбора оптимальных координат точек приземления при управляемом спуске отработавших ступеней ракет-носителей	291
Сулейменов И.Э., Бакиров А.С., Матрасулова Д.К., Тасбулатова З.С., Гришина А.В., Косцова М.В., Мун Г.А.	Новые подходы к использованию систем искусственного интеллекта в учебном процессе	306
Терехов А.Г., Егембердиева З.М., Джаксылыкова А.Б.	Обзор нейросетевых методов разрешения анафорических связей	315
Терехов А.Г., Пак А.А.	Спутниковый мониторинг водных объектов Синьцзян-Уйгурского автономного района КНР	319
Тлеубергенова М.А., Жаманшин А.У.	Хаос Пуанкаре в нейронных сетях	327
Уалиева И.М., Мусабаев Р.Р.	Исследование признаков для автоматической идентификации жанров публикаций в СМИ	331
Усатова О.А., Нысанбаева С.Е.	Обеспечение защиты информационной системы с помощью двухфакторной аутентификации	337
Утепбергенов И.Т., Бобров Л.К., Буранбаева А.И., Тойбаева Ш.Д., Коржаспаев А.Е.	Исследование проблем обеспечения информационной поддержки этапов жизненного цикла инноваций	343
Утепбергенов И.Т., Яворский В.В., Касымова Д.Т.	Основные требования к хранилищу данных путей передвижения	353
Хисамиев З.Г., Алимжанов С.А.	Компаньоны кольца целых чисел	361
Хорошилов А.А., Мусабаев Р.Р., Красовицкий А.М., Уалиева И.М., Мейрамбеккызы Ж.	Автоматическое установление тональности текстов методом их концептуального анализа	365

Ширяева О.И.	Разработка искусственной иммунной сложной системы управления нефтегазовой отрасли на основе алгоритма клональной селекции	375
Юничева Н.Р., Нурсеитов Д.Б., Бостанбеков К.А., Алимова А.Н., Джамалов Д.К., Есеркепова И.Б.	Формирование архива исторических данных глобальных моделей климата для проведения исследований в рамках международного проекта CORDEX	383
Яворский В.В., Утепбергенов И.Т., Ахмедиярова А.Т., Касымова Д.Т.	Создание хранилищ данных путей передвижения	387
Якунин К., Касымжанов Б., Кучин Я., Сымагулов А., Мурзахметов С., Булдыбаев Т., Мухамедиев Р.	Разработка рекомендательной системы по улучшению качества школьного образования на основе интерпретаторов моделей машинного обучения	394
Содержание		401

МАТЕРИАЛЫ
научной конференции
ИИВТ МОН РК
«Современные проблемы информатики и
вычислительных технологий»
1 - 4 июля 2019 года

(Сборник статей)

Подписано в печать 18.06.2019 г. Формат А4.
Печать цифровая. Бумага офсетная. Усл. печ. л. 25,12.
Тираж 300 экз. Заказ № 01306.
Отпечатано в ИИВТ МОН РК.
Алматы, ул. Пушкина, 125