

VISHNIAKOU U.A., YIWEI X.

IT DIAGNOSTICS OF PARKINSON'S DISEASE BASED ON VOICE MARKERS AND DECREASED MOTOR ACTIVITY

*Belarusian state University of Informatics and Radioelectronics,
Minsk, Republic of Belarus*

The objectives of the article to propose the method for complex recognition of Parkinson's disease using machine learning, based on markers of voice analysis and changes in patient movements on known data sets. The time-frequency function, (the wavelet function) and the Meyer keprstral coefficient function are used. The KNN algorithm and the algorithm of a two-layer neural network were used for training and testing on publicly available datasets on speech changes and motion retardation in Parkinson's disease. A Bayesian optimizer was also used to improve the hyperparameters of the KNN algorithm. The constructed models achieved an accuracy of 94.7 % and 96.2 % on a data set on speech changes in patients with Parkinson's disease and a data set on slowing down the movement of patients, respectively. The recognition results are close to the world level. The proposed technique is intended for use in the subsystem of IT diagnostics of nervous diseases.

Ключевые слова: Parkinsons disease recognition, machine learning, KNN algorithm, Bayesian neural network

Introduction

Parkinson's disease (PD) is a common neurodegenerative disease of the elderly. The average age of the disease is 60 years, and the number of affected gradually increases with age. Medical studies show that the pathology of Parkinson's disease has two main aspects [1]: depigmentation of the substantia nigra due to degeneration of neurons containing melanin and dopamine; and the formation of Levi bodies in the substantia nigra and other areas of the brain, such as the nucleus accumbens and parts of the cortex.

In [2], the authors divided the disease scheme into six stages, with olfactory and speech disorders occurring in the first two stages. According to the study, 89 % of patients with Parkinson's disease have various degrees of speech impairment, the main symptoms of which are: low volume, slurred speech, hoarseness, monotony, lack of emotional changes, slow speech.

The third and fourth stages of the mild become more pronounced, with the most significant symptoms being tremor at rest, slowing of movement, and in the middle and late stages – balance disorders. In the final stages, the fifth and sixth, all clinical symptoms manifest. Thus, the study of speech and movement disorders as symptoms in patients with Parkinson's disease is the initial information for early diagnosis of the disease, including using IT.

IT diagnostics method

The study was conducted to classify patients with Parkinsonism and healthy people with the allocation of fourteen phonation features and twelve cepstral speech features. The analyzed fourteen phonation signs include five variants of trembling, six variants

of flickering, two signs of harmony and the average autocorrelation of the main frequency characteristic [3].

A method based on recognition and classification using a deep neural network (DNN) in combination with mini-batch gradient descent (MPGS) was proposed to distinguish patients with PD from healthy people by signs of voice change [4]. The authors applied the method of weighted coefficients Mel-cepstrum (WMFCC) to extract objects. The WMFCC approach can solve the problem that the high-order Mel-cepstrum coefficients are small and the features component's ability to represent sound is weak. MPGS reduces the computational load on the loss function and increases the learning rate of the system.

An approach to the analysis of human movement based on the quartile deviation of the normal distribution (QDoND) for the selection of the best features is also proposed [5]. In 2019, [6] proposed an intelligent system for detecting Parkinson's disease based on deep learning methods for analyzing information about a person's gait. The authors used a convolutional neural network (1D-Convnet) to build a classifier based on a deep neural network (DNN).

In [7], the authors analyzed the modern technique of recognizing changes in people's speech using x -vectors, obtaining automatic detection of signs of PD. The growing popularity and success of Transformer networks in natural language processing and image recognition prompted the development of a new method for solving the problem based on the automatic extraction of BP features using Transformers networks. Their use in the 1D signal has not yet become widespread, but the authors of [8] have shown that they are effective in extracting relevant features from 1D signals.

Algorithms for IT-diagnostics

In the article, we will consider the use of machine learning and neural networks as a method for detecting signs of Parkinson's disease. Machine learning algorithms and neural network algorithms can be divided into the following four categories depending on the input data used, the training method and the output data of the studied model.

1. Supervised learning: An algorithm of this type builds a mapping from input data to output based on observations of a series of samples (each of which has a corresponding output signal) and eventually creates a predictive model.

2. Unsupervised learning: This type of algorithm only requires a series of sample points as input and does not require the samples to be pre-labeled with the corresponding output data. The information obtained using the algorithm can be used to build a descriptive model, a classic example of which is a search engine.

3. Semi-supervised Learning: This type of algorithm uses a large amount of unlabeled data, and also simultaneously uses labeled data to perform pattern recognition work.

4. Reinforcement learning: This type of algorithm is able to be trained by repeating several times and observing the feedback generated by the environment after each iteration. In fact, the output and actions taken by the model after each iteration will have an impact on the environment, and the environment will provide feedback in response to these impacts. This type of algorithm is mainly used for speech and text recognition.

Parameters are the key to machine learning algorithms and are usually divided into general and hyper parameters. General parameters are variables that can be automatically extracted by the model from the data. For example, the weights of each layer of a neural network, the a priori probability of each class in a simple Bayesian algorithm, etc. Hyper parameters are configurations external to the model, the values of which cannot be estimated from the data and are usually set manually. Hyper parameters are typically empirically defined variables. Different machine learning methods have different hyper parameters. For example, in neural networks, hyperparameters include learning rate, regularization parameter, number of layers, epoch, loss function, weight initialization method, neuron activation function, etc.

Since the data taken is labeled, the experiments given in this article relate to supervised learning. The KNN algorithm is used for machine learning of the BP recognition network by voice and the Bayesian algorithm of the neural network of BP recognition by motion change.

The KNN algorithm. The main idea of the K-nearest neighbor (KNN) algorithm is that if most of the k nearest samples in the object space belong to a category,

then the sample also belongs to this category [9]. The KNN algorithm has two main parameters:

1. Distance calculation methods. Common methods for calculating distances between two points are Manhattan distance, Euclidean distance, and Minkowski distance. Of course, you can add weight to the distance to each point so that the closer point gets more weight. A function is added to the calculated value, such as the inverse function, the Gaussian function, etc.

2. The value of k is usually an integer not exceeding 20, k is too small to be affected by outliers and too large to be affected by sample imbalance.

Input data: k , the number of nearest neighbors; a set of test samples; a set of training samples. Result: a set of test sample labels.

The implementation of the KNN algorithm included the following steps:

1. Input of training and testing data;
2. Definition of parameter k (nearest number of neighbors);

3. Calculating the distance to the object from the given training data (between a point in the dataset of a known category and the current point, between the data point whose class should be predicted and all training data points). Euclidean distance can be used here.

4. Sort the results in ascending order (sequentially from low to high, in ascending order of distance);

5. Selecting k points with the smallest distance from the current point (collecting a category from labels (classifying nearest neighbors based on the value of k));

6. Determining the frequency of occurrence of the category in which k points are located;

7. Return to the category with the highest frequency of occurrence of the first k points as the predicted classification of the current point. When using the nearest neighbor category, which is the most dominant, it is predicted as an object label. Finding the number of classes from the nearest neighbor and setting the data class to be evaluated.

The advantage of the KNN algorithm is that it is computationally simple, and the disadvantage is that it requires a lot of computational effort and requires a lot of memory. It is necessary to choose the appropriate value of k .

The Bayesian neural network algorithm. Bayesian neural networks (BNN) combine probabilistic modeling with neural networks and are able to increase the predicted result. The input layer is used to describe key parameters and as data input for a neural network. The output of a neural network is used to describe a probability or probability distribution. A posteriori distribution is calculated by sampling or variational inference. Bayesian neural networks differ from conventional neural networks in that the weight parameters are random variables, not specific values. In [10], BNN was used to identify the BP by slowing down the movement. Hyperparameters in the Bayesian neural network algorithm include.

1. The number of hidden layers.
2. Regularization parameters. Regularization coefficients affect the ability of the model to generalize
3. The number of neurons per layer.
4. The era of learning.
5. Learning speed. A graph of the learning rate for updating the weight.
6. Hidden layer activation function. Activation functions such as logistic regression function, tanh function, relu function, etc.
7. Mini-packet size for small data packets. The number of samples per training session.
8. Weight optimization solver. 'lbfgs' is an optimizer in the family of quasi-Newtonian methods. "sgd" refers to stochastic gradient descent. "adam" refers to a stochastic gradient-based optimizer.

Neuron network

In [11], speech disorders in Parkinson's disease were studied, and the collected data (<https://archive.ics.uci.edu/ml/datasets/Parkinson%27s+Disease+Classification#>) various algorithms for processing speech signals were applied, taking into account the features of the time frequency, the characteristics of the Mel-cepstrum coefficient, etc. The data used in this study (sound) were collected from 188 patients with PD (107 men and 81 women) aged 33 to 87 years (65.1±10.9) at the Department of Neurology of the Serrahpa Medical Faculty of Istanbul University. The control group consisted of 64 healthy people (23 men and 41 women) aged 41 to 82 years (61.1±8.9). In the process of data collection, the microphone is tuned to a frequency of 44.1 kHz, and after examination by a doctor, a stable phonetics of the vowel "a" was obtained from each subject with three repetitions.

The problem of slowing down movement in Parkinson's disease was investigated in [12]. The Daphnet (action) dataset was recorded in the lab with an emphasis on generating multiple deceleration events (<http://archive.ics.uci.edu/ml/datasets/Daphnet+Freezing+of+Gait>). Users performed three types of tasks: walking in a straight line, walking with multiple turns, and finally a more realistic ADL daily life task in which users entered different rooms, bringing coffee, opening doors, etc. The data set was obtained from three wearable wireless acceleration sensors registering 3D acceleration at a frequency of 64 Hz. The sensors were located on the lower leg, on the hip and on the hip joint.

To recognize the BP by movement, a neural network model has been developed, which uses a 2-layer fully connected network with the first and second layers of size 10, as well as the ReLU function as activation. The ReLU function is better than sigmoidal functions and hyperbolic tangent functions, since it is easier to find

derivatives, and also increases the nonlinearity of the network.

In this article, the above two publicly available datasets (sound, action) were selected as experimental. After data purification and preprocessing, the characteristics of the time-frequency domain, the inverse spectral coefficients of Mel, and the wavelet coefficients in the sound datasets were extracted. In the action data sets, 5 time-frequency characteristics were extracted: the entropy of the wavelet, the energy of the wavelet, the length of the waveform of the wavelet signal, the variance of the wavelet coefficient and the standard deviation of the wavelet coefficient.

Then the set of data features was normalized, divided into a set of training data and a set of test data in a ratio of 9:1. The training data sets were processed and tested using 5-fold cross-validation with 5-fold repetition. A set of test data is used to verify the final results.

Results of IT-diagnostics

Before proceeding with the experiments, we introduce the parameters of the evaluation model, which will help us to have a clear idea of the recognition capacity of the model. The confusion matrix [16] is used to compute various performance measures for multi-class classification problems. Performance measures for a specific class k include sensitivity ($Sens_k$), precision ($Prec_k$), and F1-score ($F1_k$):

$$\begin{aligned}
 Sens_k &= \frac{C_{k,k}}{\sum_{i=1}^K C_{i,k}} ; \\
 Prec_k &= \frac{C_{k,k}}{\sum_{j=1}^K C_{k,j}} ; \\
 F1_k &= 2 \cdot \frac{Prec_k \cdot Sens_k}{Prec_k + Sens_k},
 \end{aligned} \tag{1}$$

where i is the number of CM rows, j is the number of CM columns, k is the number of any class. C is a single element in the confusion matrix.

Since the amount of data in the sound dataset is small, the KNN algorithm was used, which is a simple machine learning classification algorithm. A Bayesian optimizer was also used to improve KNN hyper parameters. The number of iterations of the calculation was 30. The graph of the minimum classification error by speech characteristics using the KNN algorithm is shown in Figure 1.

From Figure 1, it can be seen that the minimum classification error index tends to a stable state when the number of iterations reaches about the 15th iteration, and reaches the minimum value at the 29th iteration, so we take the value of the KNN parameter of the 29th iteration as the penultimate for the experiment.

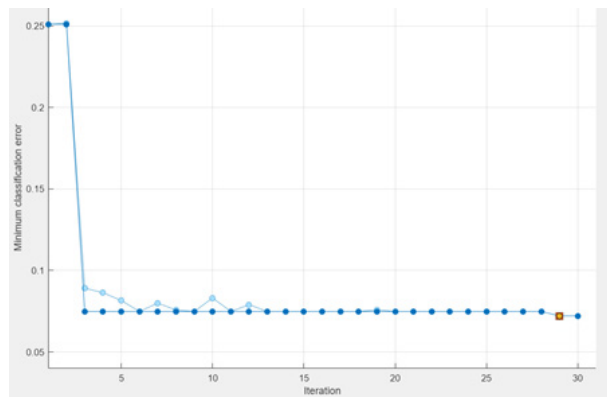


Figure 1. The Minimal classification error plot on speech features using KNN algorithm

The PCA (Principal Component Analysis) method was used to select the 10 functions that best fit the model. The specific parameters of the KNN algorithm were: number of neighbors; distance measure (cosine function); distance weight (inverse distance weighting algorithm). Graphs of the confusion matrix and the ROC curve (Receiver operating characteristics) [12] of the results of a test experiment on the recognition of BP after training the model are shown in Figures 2 and 3.

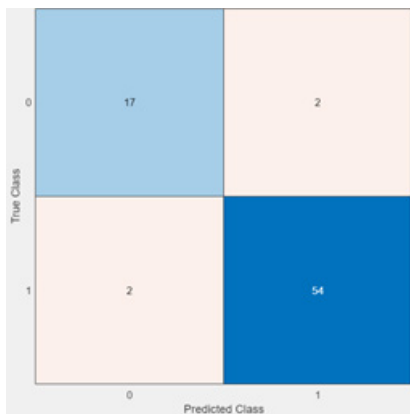


Figure 2. Confusion matrix of test dataset (sound)

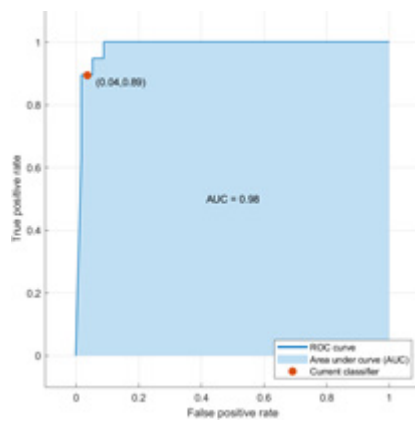


Figure 3. ROC plots of test dataset (sound)

The recognition results for changing the movements of patients for the diagnosis of Parkinson's are as follows. Graphs of the confusion matrix and the ROC curve of the test experiment results for the motion change dataset are shown in Figures 4 and 5.

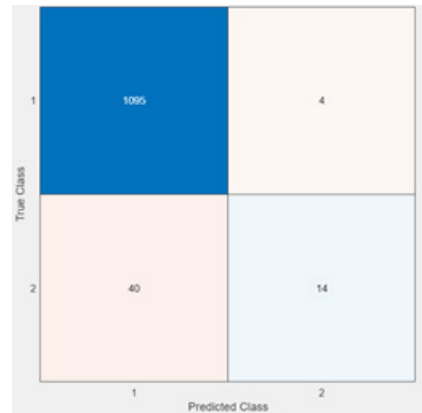


Figure 4. Confusion matrix of test dataset (action)

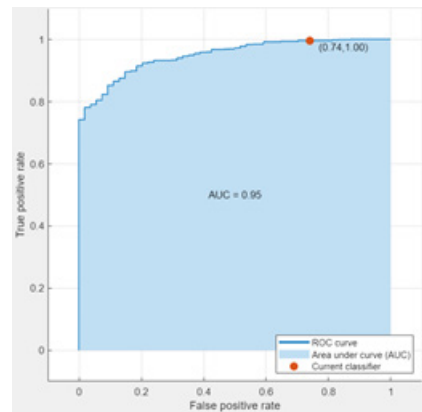


Figure 5. ROC plots of test dataset (action)

Results of Parkinson IT-diagnostics

The data of test experiments for BP on speech modification are given in Table 1.

Table 1

The data of test experiments for speech recognition

| Datasets | Average Precision | Average Sensitivity | Average F1 score | Test Accuracy |
|-----------|-------------------|---------------------|------------------|---------------|
| Pd_speech | 92.95 % | 92.95 % | 92.95 % | 94.7 % |

The model can achieve 94.7 % accuracy in diagnosing Parkinson's disease based on speech data and a high F1 score of up to 92.95 %. The accuracy of the training data set was 92.8 %, and the accuracy of the test data set was 94.7 %. The accuracy of the

test set is 1.9 % higher than that of the training set, which means that the volume of experimental data is too small, which will lead to an uneven slice of the data set, if the model correctly reflects the distribution structure within the data, it is possible that the internal variance of the training set will be larger than the validation set, which will lead to a larger error in training set, which will require reallocation of the dataset to make it evenly distributed. On the same data set [11], one of the best indicators of foreign studies is 95.8 % [13].

The data of test experiments for the recognition of BP by changing movements are shown in Table 2.

Table 2

The data of test experiments for the identification of Parkinson's movement

| Datasets | Average Precision | Average Sensitivity | Average F1 score | Test Accuracy |
|-----------|-------------------|---------------------|------------------|---------------|
| Pd_action | 62.78 % | 87.13 % | 72.98 % | 96.2 % |

Thus, experiments show that the recognition accuracy when using a two-layer neural network for a set of data on changes in the movement of patients reaches 96.2 %. On the same data set of Daphnet, one of the best indicators of foreign researchers is 98.8 % [14]. However, the confusion matrix shows that the weighted average F1 score of the test dataset is low, only 72.98 %, due to an unbalanced dataset. The motion data in the dataset contains data about various shapes. These data on various forms of movement were mixed, which leads to a decrease in the accuracy of recognition of Parkinson's disease.

Conclusion

1. The time-frequency characteristics (wavelet) used in the experiment included the entropy of the wavelet energy, the wavelet energy, the variance of the wavelet coefficient, the small wavelength and the wavelet coefficient. After using the PCA feature extraction method, 10 wavelet features were obtained. The KNN algorithm was used for speech recognition of patients, while the test accuracy of 94.7 % was achieved in the diagnosis of Parkinson's disease by voice change. The Bayesian neural network algorithm was used to recognize the slowing down of patients' movements, it gave a test accuracy of 96.2 % for the diagnosis of Parkinson's disease by changing the movements of patients.

2. The results of the experiment show that early detection of Parkinson's disease based on data on changes in voice and movement is effective. But it is also clear from the experiment that the data set has a very important influence on the results of the experiment. A more balanced dataset with a larger amount of data can provide better recognition results.

3. IT technology for the recognition of Parkinson's disease has an applied nature and good research significance. Further research work may include: analysis of time-frequency characteristics and investigation of the influence of the number of different wavelets, the levels of decomposition of wavelets and statistical characteristics of wavelets on the recognition accuracy. The sampling rate of the data set can also be increased. Increasing the sampling rate allows you to more fully collect information to obtain better wavelet characteristics. In addition, an experiment with a large data set and balanced categories of data on Parkinson's disease will give more objective and accurate results.

REFERENCES

1. **Davie, C.A.** A review of Parkinson's disease. *Br. Med. Bull.*, Feb. 2008, vol. 86, no. 1, pp. 109-127. doi: 10.1093/bmb/ldn013
2. **Braak, H., Ghebremedhin, E. Rüb U.** Stages in the development of Parkinson's disease-related pathology. *Cell Tissue Res.*, Oct. 2004, vol. 318, no. 1, pp. 121-134. doi: 10.1007/s00441-004-0956-9
3. **Upadhya, S.S., Cheeran A.N.** Discriminating Parkinson and Healthy People Using Phonation and Cepstral Features of Speech. *Procedia Comput. Sci.*, Jan. 2018, vol. 143, pp. 197-202. doi: 10.1016/j.procs.2018.10.376
4. **Xu, Z. [et al.]** Voiceprint recognition of Parkinson patients based on deep learning. *arXiv*, Dec., 2018, pp. 1-10. doi: 10.48550/arXiv.1812.06613
5. **Arshad, H. [et al.]** Multi-level features fusion and selection for human gait recognition: an optimized framework of Bayesian model and binomial distribution. *Int. J. Mach. Learn. and Cybern.*, Dec., 2018, vol. 10, no. 12, pp. 3601-3618. doi: 10.1007/s13042-019-00947-0
6. **Maachi I.E., Bilodeau G.-A., Bouachir W.** Deep 1D-Convnet for accurate Parkinson disease detection and severity prediction from gait. *Expert Syst. Appl.*, May 2020, vol. 143, pp. 1-27. doi: 10.1016/j.eswa.2019.113075
7. **Moro-Velazquez, L. [et al.]** Advances in Parkinson's Disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects. *Biomed. Signal Process. Control*, Apr. 2021, vol. 66, pp. 1-13. doi: 10.1016/j.bspc.2021.102418

8. Nguyen, D.M. [et. al.] Transformers for 1D Signals in Parkinson's Disease Detection from Gait. arXiv, Apr., 2022, pp. 1-7. doi: 10.48550/arXiv.2204.00423
9. Zhang, M.-L., Zhou Z.-H. ML-KNN: A lazy learning approach to multi-label learning, Pattern Recognit., 2007, vol. 40, no. 7, pp. 1-21.
10. Arshad, H. [et. al.] Multi-level features fusion and selection for human gait recognition: an optimized framework of Bayesian model and binomial distribution. Int. J. Mach. Learn. & Cyber, 2019, vol. 10 (12), pp. 3601-3618.
11. Sakar C.O. [et al.] A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. Appl. Soft Comput., Jan. 2019, vol. 74, pp. 255-263. doi: 10.1016/j.asoc.2018.10.022
12. Fawcett, T. An introduction to ROC analysis. Pattern Recognit. Lett. 2006, vol. 27, no. 8, pp. 861-874.
13. Sakar B.E., [et al.] Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. IEEE J Biomed Health Inform, 2013, vol. 17(4), pp. 828-834.
14. Li, B., Yao Z., Wang J., Wang S., Yang X., Sun Y. Improved Deep Learning Technique to Detect Freezing of Gait in Parkinson's Disease Based on Wearable Sensors. Electronics, 2020, no 9 (11), pp. 1-12.

ВИШНЯКОВ В.А., ИВЭЙ С.

ИТ-ДИАГНОСТИКА БОЛЕЗНИ ПАРКИНСОНА НА ОСНОВЕ ГОЛОСОВЫХ МАРКЕРОВ И СНИЖЕНИЯ ДВИГАТЕЛЬНОЙ АКТИВНОСТИ

*Белорусский государственный университет информатики и радиоэлектроники,
Минск, Республика Беларусь*

Цели статьи – предложить метод комплексного распознавания болезни Паркинсона с использованием машинного обучения, основанный на анализе маркеров голоса и изменений в движениях пациента на известных наборах данных. Используются частотно-временная функция (вейвлет-функция) и функция коэффициента Мейера Кепстрала. Алгоритм KNN и алгоритм двухслойной нейронной сети были использованы для обучения и тестирования на общедоступных наборах данных об изменениях речи и замедлении движений при болезни Паркинсона. Байесовский оптимизатор также использовался для улучшения гиперпараметров алгоритма KNN. Построенные модели достигли точности 94,7% и 96,2% для набора данных об изменениях речи у пациентов с болезнью Паркинсона и набора данных о замедлении передвижения пациентов, соответственно. Результаты распознавания близки к мировому уровню. Предлагаемая методика предназначена для использования в подсистеме ИТ-диагностики нервных заболеваний.

Ключевые слова: *распознавание болезни Паркинсона, машинное обучение, алгоритм KNN, байесовская нейронная сеть.*



Вишняков Владимир Анатольевич, д.т.н., профессор, профессор кафедры инфокоммуникационных технологий Белорусского государственного университета информатики и радиоэлектроники. Область научных интересов: информационное управление и безопасность, электронный бизнес и коммерция, интеллектуальные технологии в инфокоммуникационных системах, сети интернет вещей, блокчейн. Член двух докторских Советов по защите диссертаций. Автор более 500 научных работ, в том числе 7 монографий (2 – английском языке), двух учебных пособий с грифом Министерства образования БССР, двух учебных пособий с грифом Министерства образования Республики Беларусь, 28 учебно-методических пособий, включая 8-й томный учебный комплекс «Информационный менеджмент», 190 научных статей, 21 патента.

Тел.: Тел. +375-44-486-71-82

E-mail: vish@bsuir.by

Vishnyakou Uladzimir Anatolyevich, Doctor of Technical Sciences, Professor, Professor of the Department of Information and Communication Technologies of the Belarusian State University of Informatics and Radioelectronics. Research interests: information management and security, electronic business and commerce, intelligent technologies in infocommunication systems, the Internet of Things, blockchain. Member of two doctoral councils for the defense of dissertations. Author of more than 500 scientific papers, including 7 monographs (2 in English), two textbooks with the stamp of the Ministry of Education of the BSSR, two textbooks with the stamp of the Ministry of Education of the Republic of Belarus, 28 teaching aids, including the 8th volume educational complex "Information Management", 190 scientific articles, 21 the patent.

Ся Ивэй, магистр технических наук, аспирант кафедры ИКТ БГУИР. Область научных интересов: сети интернет вещей и интеллектуальные системы ИТ-диагностики.

Xia Yiwei, master of technical science, PhD-student of ICT department of Belarusian State University of Informatics and Radioelectronics. Research interests: Internet of things networks and intelligent IT diagnostics systems.