FINANCIAL UNIVERSITY
UNDER THE GOVERNMENT
OF THE RUSSIAN FEDERATION

МТУСИ

TEDS'18

PROCEEDINGS
OF THE INTERNATIONAL CONFERENCE ON
TECHNOLOGY & ENTREPRENEURSHIP
IN DIGITAL SOCIETY

2019

International Conference Technology & Entrepreneurship in Digital Society (TEDS) aims to provide an opportunity for researchers, practitioners, and professionals from both the industry and the academia to share their newest research findings and results on technologies and their business applications in the digital era.

Additional copies may be ordered from:

Real Economy Publishing House
Russia, 190020, Saint Petersburg, Staro-Petergofskiy prospect 43-45 B, of. 4.

Tel.: +7 (812) 346-50-15 (16,17,18)
E-mail: info@jsdrm.ru

Printed in the Russian Federation

Design and Development of the Module to Control the Quadcopter using the Signs shown by Hands

Alex A. Ezhov[1], Vjacheslav I. Voronov[2]
Department of Intelligent Control Systems and Automation
Moscow Technology University of Communication and Informatics
Moscow, Russia
[1]alekcey_ezh@mail.ru, [2]vorvi@mail.ru

***Abstract*. With the development of the industrial Internet of things (IIoT) there is a huge number of tasks that require the use of intelligent methods of information processing. Unmanned aerial vehicles (UAVs) are a class of IIoT devices that require computer vision to solve piloting tasks and provide the necessary functionality. The article describes the development of a module to control the quadcopter using gestures depicted by hands. The solution of this problem is very important to expand the methods of control actions.**

***Key words:* IIoT, tensorflow, quadcopter, robot, neural network..**

I. INTRODUCTION

Currently, the task of creating intelligent machines (robots) is extremely relevant. Robots can perform quite complex work; they are widely used in the industrial Internet of things: in production, in the organization of warehouses, in military Affairs, in medicine. Robots may have a different design, size and level of intelligence.

At the Department of ISUA MTUCI project is being designed for contactless control IIoT with the help of computer vision systems [1–3].

The article describes the development of an automatic mobile robot from the class of unmanned aerial vehicles. Using computer vision technology and machine learning methods. quadcopter will receive and recognize commands in the form of signs, depicted by the hands of the user and perform certain actions.

II. ANALYSIS

When determining the design of the robot, it is required to classify it taking into account the task indicated in the introduction. It is necessary not only select the appropriate housing materials, engines and mechanisms, but also to determine the system of computer vision and technologies of big data processing and machine learning [4–6].

The robot is an automatic device created on the principle of a living organism. Acting on a pre-programmed program and receiving information about the outside world from the sensors, the robot independently carries out production and other operations, usually performed by man (or animals). In this case, the robot can, as well as have a connection with the operator (receive commands from him), and act autonomously.

The basis of the modern robot is a set of actuators and devices with drive systems that allow you to interact with the environment in the performance of a given task. Drive-a set of devices designed to drive machines and mechanisms.

The functionality of the robot, its versatility and compliance with the characteristics is determined primarily by the variety of actions that the robot can perform.

The list of operations that the robot is designed to perform varies, depending on the list of special, specialized and universal robots. Special robots are designed to perform one specific task. Specialized robots can perform several similar operations. Universal robots can perform various basic and auxiliary operations within their technical capabilities. The increase in the degree of versatility of the robot extends the scope of its possible applications, but at the same time inevitably accompanied by underutilization of these opportunities for each specific operation, as well as the rise in the cost of the robot. Optimal in this respect are special robots, but on the other hand it is extremely narrows their market, and, consequently, the volume of production.

Manipulation robot is an automatic machine (stationary or mobile), consisting of an executive device in the form of a manipulator with several degrees of mobility, and a software control device that serves to perform motor and control functions in the production process. Such work is performed in outdoor, overhead and gantry designs. The most widely used in machine-building and instrument-making industries.

A mobile robot is an automatic machine that has a moving chassis with automatically controlled drives.

Mobile robots can be wheeled, walking and tracked (there are also crawling, floating and flying mobile robotic systems) depending on the type and operating conditions [1].

Classification of robots according to the characteristics of their design. These indicators include:

- type of drive;
- carrying;
- the number of manipulators;

- type and parameters of their working area;
- mobility and method of placement;
- intended use.

Drives that are used in manipulators and robot movement systems can be electric, hydraulic and pneumatic. Often they are used in combination.

The load capacity of the robot is the load capacity of its manipulators, and for the transport robot also its chassis.

The number of manipulators in robots in most cases is limited to one. However, depending on the purpose, there are robot designs without manipulators and with 2, 3 and rarely 4 manipulators.

The type and parameters of the working area of the manipulator determine the area of the surrounding space within which it can perform manipulations without moving, i.e. with a fixed base.

The mobility of the robot is determined by the presence or absence of the robot movement system. In the first case, robots are called mobile, and in the second stationary.

According to the method of placement of stationary and mobile robots are floor, suspended and embedded in other equipment.

The use of the robot for its intended purpose depends on the external conditions in which it must operate. There are normal performance, dustproof, heatproof, moisture-proof, explosion-proof, etc.

Automatic classification of robots according to the method of control: 1) Software; 2) Adaptive; 3) Intellectual.

Software control device. Such devices operate according to a predetermined program, mainly designed to solve monotonous problems in a constant environment.

Adaptive control device. In such systems, the control is carried out on the basis of the received information about the current state of the environment and the robot itself, obtained in the process of control from sensor devices (solve typical problems, and adapt to the operating conditions based on the data obtained from the information system).

Intelligent control device. These systems use artificial intelligence techniques to adapt and perform other robot functions.

Control devices can be individual, part of each robot and group, managing multiple robots. Structurally, individual control devices are usually performed separately from the mechanical part of the robot, much less often in the General body. Mobile robots usually have a control device consisting of 2 parts-on-Board and included in the operator panel (or in addition to it).

The vast majority of robots have electronic control devices made on the microcontroller. However, there are also non-electric control devices of robots, often implemented on Pneumatics and intended for use in special explosive and fire hazardous conditions at elevated temperatures.

Since the quadcopter or UAV is used for operation, we will consider its structure in more detail.

UAVs are divided into several classes by weight, in this paper, we consider the AIRCRAFT belonging to the class “mini”, because it has a mass of not more than 4 kg. It consists of body, engine, onboard control system (autopilot), ground control system (NSO) and aerial photography equipment.

The UAV body is made of lightweight but durable materials, such as carbon fiber or Kevlar, to protect cameras and sensors, and the wings are made of extruded polystyrene foam (EPP). This material is light enough and does not break when falling. Manufacturers are trying to make cheaper parts of the body to keep the UAV in working condition was less expensive.

The engine can be either petrol or electric. The gasoline engine has one main advantage: it provides a longer flight, but this engine is less reliable and more complex in design. The electric motor is easy to operate, simple in design and requires less maintenance.

Autopilot is the main control element of the UAV. It has a processor and many sensors that monitor the deviation of the UAV during the flight.

Ground control system (NSO) is a computer; it is installed software to track the flight, which communicates with the UAV. Most often, the task for the UAV is defined as a flight at specified coordinates or points. During the flight, the NSO displays the coordinates of the terrain, as well as the altitude of the UAV flight, so that the operator can redirect, make an emergency landing or adjust the flight.

A quadcopter, also called a helicopter, is a multi-engine helicopter that is lifted and propelled by four rotors. Quadcopters are classified as rotary-wing aircraft, as opposed to aircraft, because their rise is made by means of rotors (vertically oriented propellers).

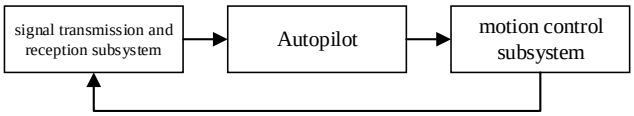


Fig. 1. Block diagram of the quadcopter

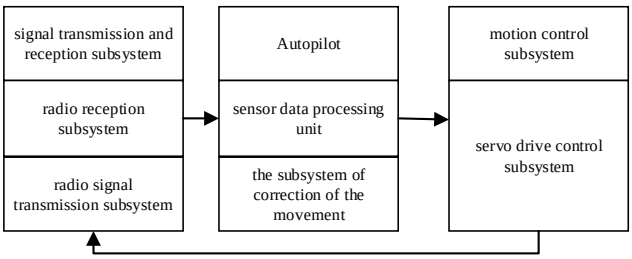


Fig. 2. Structural and functional scheme of the quadcopter

Quadrocopters typically use two pairs of identical fixed propellers; two of which operate clockwise (CW) and two counterclockwise (CCW). For more precise control and control of the quadcopter, an independent change in the speed of each

rotor is used, as well as due to this, you can change the desired total thrust in a large range. The use of changing the speed of individual rotors helps to position the center of thrust in both transverse and longitudinal direction and change the torque or turning force.

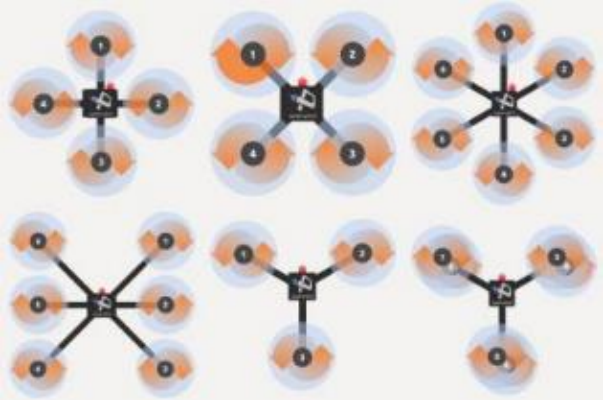


Fig. 3 Aircraft models

Remote control controller controls the quadcopter, by radio waves.

To communicate and respond to the requests of the copter control panel, the UAV is installed electronic systems:

* Flight controller-is an important component of the UAV, as it gives the user the ability to maneuver the device according to his wishes. First, the flight controller receives signals from the operator's transmitter. This allows commands and user instructions to reach the quadcopter. The flight controller is also useful for working together with sensors that are placed on Board for smooth flight. The controller has another function. A quadcopter is multirotor, which comes with four different engines. Each engine has a certain speed. Thus, the flight controller is able to calculate the speed of each of the four engines, and then sends a signal to the electronic speed controllers or ESC. This facilitates a smooth flight. The quadcopter flight controller can perform each of these functions mentioned above many times, resulting in efficient flight.

- An accelerometer is essentially a device that measures the linear acceleration of a quadcopter in a 3-axis system. This sensor gives an output that allows the user to understand the position of the drone. The simple fact that the accelerometer is able to detect and respond to gravity means that it remains stable in the air, ensuring a smooth flight. The correct way to install the accelerometer on the flight controller Board is to align the linear axis of the device with the primary axis of the UAV.
- A gyroscope (like an accelerometer) is useful for measuring the rate of angular change. However, you need to do some iteration to get the angle. The rotational axis of the sensor must coincide with the axis of the UAV.
- Inert-measuring unit (IMU) consists of both gyroscope and accelerometer. Some systems combine the use of a magnetometer and similar devices.
- Magnetometer is an electronic compass placed on the controller Board. This is especially necessary to align the direction of the drone compared to the earth's magnetic field.
- Barometer is a device for measuring pressure. The pressure sensor on the drone can give us the exact altitude value of the device from sea level. The most accurate altitude information is determined by the combination of the GPS pressure and altitude sensor.
- Distance sensor basically improves the height value set by the pressure sensor or barometer. The distance sensor can be ultrasonic, laser or LIDAR technology.

Quadcopters are different from conventional helicopters, which use rotors that dynamically change the height of their blades as they move around the rotor hub.

At first, quadcopters were considered as an alternative solution to some problems of vertical flight. Torque-dependent control problems (as well as efficiency problems arising from the tail rotor, which does not create a useful lift) can be eliminated by counter-rotating, and relatively short blades are much easier to manufacture and adjust.

III. DESIGN

A. Objectives

The Department of ISUA carries out a number of projects on the development of automatic control systems for objects of the Internet of things [7; 8]. One of the subtasks design and build a robot that must perform such actions:

- stay stable in the air and wait for the command from the user,
- find a specified object with the help of the built-in video camera and recognize what you see, taking into account this, perform any actions,
- broadcast over a wireless connection the camera image.

B. General scheme of the architecture of the robot

The General scheme of the robot architecture can be divided into 4 levels: mechanics, electronics, reactive control, tactical control, external control.

The lowest level is the body of the quadcopter. It is mounted on level 1 engine control system and sensors.

These devices are connected to the level 2 microcontroller, in the firmware of which there are simple rules of reactions to the sensor or camera readings, for example, turn to the left. This firmware also contains a list of commands to control the robot's movement (forward, turn, etc.).

The intelligent filling of the robot is contained at level 3, which houses a sufficiently productive minicomputer, to which the camera and the microcontroller of level 2 are connected. The minicomputer analyzes the image and sensor information received from the microcontroller, based on this information

controls the movement. There is also a wireless device (wi-fi) and the corresponding software, which provides a communication channel with level 4.

The highest level (number 4) is an external storage for the accumulation and analysis of large amounts of data and/or an operator that can adjust the behavior of the robot and receive information from it.

TABLE I. The layered architecture of robot

Level 4: external control	• Accumulation of information, • training, • planning, • general management, •operator interface
Level 3: tactical control	• Communication channel and external control protocol, •automatic control software, •camera
Level 2: reactive control	• Simple reaction (stop, forward, right, left), •motor control. •sensor survey
Level 1: electronics	•Motor driver, •simple sensors
Level 0: mechanics	•Body

C. *The hardware*

Set of equipment operation and its layout may be different.

TABLE II. Options kit component parts

<table>
<tr><td>Level 4: external PC control</td><td colspan="3">PC</td></tr>
<tr><td>Level 3: tactical control</td><td>send sensor data directly to the server using a simple minicomputer communication module</td><td>minicomputer</td><td rowspan="2">minicomputer</td></tr>
<tr><td>Level 2: reactive control</td><td colspan="2">microcontroller</td></tr>
<tr><td>Level 1: electronics</td><td colspan="3">sensors and motor driver</td></tr>
<tr><td>Level 0: mechanics</td><td colspan="3">the body of the quadcopter</td></tr>
</table>

Thediagram of the robot hardware is shown in fig. 4.

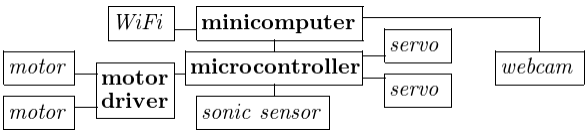


Fig. 4. *Diagram of the robot's hardware*

IV. IMPLEMENTATION

To teach the robot to detect and recognize signs, an open library TensorFlow Object Detection API (Google) was used. It easily allows you to train your network from scratch. It was decided to use the so-called Deep Learning neural networks, because now they are the most advanced. Google provides many options for models. Each differs in speed and accuracy. SSD works fast, but can make mistakes with small objects. In addition, the fastest Faster RUN models work slowly but with greater accuracy. The TensorFlow Object Detection API provides a number of pre-trained models, shown in fig. 5 [10].

Fig. 5. *Pre-trained Tensorflow models*

Model name	Speed (ms)	COCO mAP[^1]	Outputs
ssd_mobilenet_v1_coco	30	21	Boxes
ssd_inception_v2_coco	42	24	Boxes
faster_rcnn_inception_v2_coco	58	28	Boxes
faster_rcnn_resnet50_coco	89	30	Boxes
faster_rcnn_resnet50_lowproposals_coco	64		Boxes
rfcn_resnet101_coco	92	30	Boxes
faster_rcnn_resnet101_coco	106	32	Boxes
faster_rcnn_resnet101_lowproposals_coco	82		Boxes
faster_rcnn_inception_resnet_v2_atrous_coco	620	37	Boxes
faster_rcnn_inception_resnet_v2_atrous_lowproposals_coco	241		Boxes
faster_rcnn_nas	1833	43	Boxes
faster_rcnn_nas_lowproposals_coco	540		Boxes

The faster_run_resnet101_coco model is selected because it allows for maximum detection accuracy.

A dataset was created for training. Consisting of 300 photos. You want to increase the size of the data to increase the recognition accuracy. For the network to understand what it recognizes, you need to create a “label_map” file.txt, which contains the class description. Here is an example of such file:

item {
id: 1
name: ‘Hand’}

Each photo needs to be annotated. This means that for each image you need to create an xml file that contains the name of the photo, its size and 4 coordinates, representing the location of the frame around the object, as well as the class of the desired object. Here is the file:

<annotation>
<filename>182.jpg</filename>
<size>
<width>1280</width>
<height>586</height>
<depth>3</depth>
</size>
<segmented>0</segmented>
<object>
<name>Hand</name>
<bndbox>
<xmin>581</xmin>
<ymin>106</ymin>
<xmax>618</xmax>
<ymax>142</ymax>
</bndbox>
</object>

In addition, the main thing is not to forget to divide the dataset into two samples of 70 and 30 percent. 70 for training and 30 for performance evaluation.

For more or less normal object detection, I recommend training the network for 200 thousand iterations. But for such a long process requires a very powerful iron. The developed network was trained on GPU GeForce 1050 Ti for 25 thousand iterations. The detection accuracy reached 70%. In the future, it is planned to use cloud computing, using, for example, Google VM.

Next, the network was tested on video downloaded from the Internet, respectively; it does not apply to the dataset. The result will result in Fig. 6.

Fig. 6. *The result of testing neural network*

As you can see in the image, the network almost exactly finds both hands, but because I chose a slow model, the number of frames per second fell to 3.

Now you need to decide how to control the quadcopter using a neural network. First, you can show the hands of certain signs that with the camera, the network will recognize. Each such sign will mean some simple action-right, left, up, down. Team of course can be build up over time. Secondly, you can just follow the movement of hands. For example, when flapping up, the quadcopter flies up, when flapping to the right-the quadcopter flies to the right, etc.

The first option has more commands to control, signs hands can come up with a lot. Compared with the second method, which takes into account only the direction of movement of the hands.

The second option is simple to implement, you only need to determine how the coordinates of the movement change and then it will be clear in which direction the hand is moving.

The first method requires training a second network that can classify the characters shown. This is as time-consuming a task as learning how to detect objects. As in the case of detection, the classification requires the creation of a dataset that contains at least 300 photos for each character. Some images are shown in fig. 7 [11].

Fig. 7. Sample images from the dataset

Due to the limited capacity in the creation of the dataset used technology to increase the amount of data by distortion. For example, blurring an image, adding noise, mirroring, and turning.

Now, the network classifies the signs with an accuracy of about 50%. The work shows all the signs for training, because it will have to change and train the network again, so no sense to show it. To improve classification accuracy, you need to change the training dataset and increase its size.

V. CONCLUSION

In today's world, robots are getting a big leap in development. Many technologies make robots “smart” and help people in solving many problems. Therefore, the development of a mobile robot is a very relevant topic and will be such for a long time.

For the purposes of the course project required to design and develop a mobile robot. It was decided to use a model of a quadcopter, which would be controlled by the signs depicted by hands.

The structure and classification of robots are analyzed; the block diagram of the developed robot is presented. Its architecture and hardware are given.

Training of neural networks. Tested on a random video from the Internet, the detection accuracy of about 75%.

The classification of signs requires considerable improvement, but the quadcopter can be controlled if you use the second method given in the work, in which you only need to know the direction of movement of the hands.

VI. REFERENCES

[1] Михаеску С.В., Трунов А.С., Воронова Л.И. Анализ предметной области для разработки системы построения скелетной модели человека на основе массива опорных точек, получаемых совокупностью контроллеров KINECT // Международный студенческий научный вестник. 2015. № 3-4. С. 521–522. [Mihaesku S.V., Trunov A.S., Voronova L.I. Analysis of the subject area of system developing for building a human skeletal model based on an array of reference points obtained from a set of Kinect controllers // International Student Scientific Bulletin, 2015. N 3-4. P. 521–522].

[2] Клешнин Н.Г., Воронова Л.И. Применение нейронных сетей в подсистеме распознавания эмоций для проекта «сурдотелефон» // Телекоммуникации и информационные технологии. 2018. Т. 5. № 1. С. 126–130. [Kleshnin N.G.,

Voronova L.I. Neural networks application in emotion recognition module for “surdotelephone” project // Telecommunications and information technology. 2018. T. 5. № 1. P. 126–130].

[3] Воронов В.И., Воронова Л.И., Генчель К.В. Применение параллельных алгоритмов в нейронной сети для распознавания жестового языка // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2018). VII Международная научно-техническая и научно-методическая конференция: Сб. науч. ст.: В 4 т. / Под ред. С.В. Бачевского. СПб., 2018. С. 207–212. [Voronov V.I., Voronova L.I., Genchel K.V. The use of parallel algorithms in the neural network to recognize the sign language // Actual problems of information and telecommunications in science and education (APINO 2018). VII International scientific tech. and scientific method. conf.: In 4 t. / Ed. S.V. Bachevsky. M., 2018. P. 207–212].

[4] Воронова Л.И., Воронов В.И. Big Data. Методы и средства анализа: Учеб. пос. М., 2016. [Voronova K.I., Voronov V.I. Big Data. Methods and tools for analysis. Moscow, 2016].

[5] Усачев В.А., Воронов В.И. Компетенция «машинное обучение и большие данные» // Приоритетные направления развития науки и образования / Под общ. ред. Г.Ю. Гуляева. Пенза, 2017. С. 97–108. [Usachev V.A., Voronov V.I. The competence “Machine learning and Big Data” // Priorities for development of science and education / Ed. G.Yu. Gulayev. Pensa, 2017. P. 97–108].

[6] Горячев Д.В., Воронов В.И. Большие данные и машинное обучение // Технологии информационного общества: Матер. XII Междунар. отраслевой науч.-техн. конф. М,, 2018. С. 327–328. [Goryachev D.V. Big Data and Machine Learning // Information Society Technologies: Proceedings of 12th the international scientific-practical conference. Moscow, 2018. P. 327–328].

[7] Voronov V.I., Voronova L.I. Features of realization master’s program “Automation of technological processes and manufactures” // International Journal of Applied and Fundamental Research. 2016. № 2. URL: www.science-sd.com/464-25196.

[8] Безумнов Д.Н., Воронова Л.И. О развитии и стандартизации технологии Интернета вещей // Технологии информационного общества: Матер. XII Междунар. отраслевой науч.-техн. конф. М., 2018. С. 293–294. [Bezumnov D.N., Voronova L.I. On the development and standardization of the Internet of Things technology // Information Society Technologies: Proceedings XII International. branch scientific.-tech. conf. Moscow, 2018. P. 293–294].

[9] Blogs about robots. URL: http://enggmechanical.blogspot.com.

[10] Documentation TensorFlow. URL: https://www.tensorflow.org/api_docs/.

[11] De Pontes Oliveira I., Peixoto Medeiros J.L., De Sousa V.F. A Data Augmentation Methodology to improve Age Estimation Using Convolutional Neural Networks. 2016 // 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). 2016. P. 88–95.

Anomaly Detection by Machine Learning Method based on the Gaussian Distribution

Alexander S. Gusev[1], Vladimir N. Repinskiy[2]
Department of Intelligent Control Systems and Automation
Moscow Technical University of Communications and Informatics
Moscow, Russia
[1]gysevalexander@mail.ru, [2]repinski@rambler.ru

Abstract. **The article deals with the problem of using machine learning in detecting anomalies during the operation of IIoT objects, describes method of learning based on the Gaussian distribution, and implements the program that allows finding anomalies in data set.**

Key words: **IIoT, detecting anomalies, the Gaussian distribution, machine learning program.**

I. INDUSTRIAL INTERNET OF THINGS

A number of projects are being developed at the ICSA department of MTUCI using Big Data [1], Data Mining [2], Machine Learning [3] to support the management of IIoT objects.

At the moment, one of the most promising areas of industrial development is the industrial Internet of things. Industrial Internet of Things (IIoT), also known as Internet of things for corporate or industry use, is a system of integrated computer networks and connected industrial facilities with built-in sensors and software for data collection and exchange, with the possibility of remote control and management in an automated mode, without human intervention.

The operating principle of the technology is as follows: sensors, actuators, controllers and man-machine interfaces are initially installed on the key parts of the equipment, then information is collected, this information subsequently allows the company to acquire objective and accurate data on the state of the enterprise. Processed data is delivered to all departments of the enterprise, which helps to establish interaction between employees of different departments and make informed decisions. The obtained information can be used to prevent unplanned outages, equipment failures, reduce unplanned maintenance and failures in supply chain management, which allows the enterprise to work more efficiently.

The main value from using the industrial Internet of things is to achieve the maximum energy efficiency of any production or network. In other words, the calculation is based on the direct cost of optimization through the use of technology. The direction of IIoT is fully focused on the tasks of the industry and its specific branches, such as, for example, municipal lighting systems. First and foremost, IIoT is the interaction of sensors that control the operation of equipment in production, during the processing of raw materials or, for example, in the extraction of oil from the oil derrick. Sensor systems are widely used in industrial production, in the mining of minerals and even in lamps for lighting urban roads, which are switched on and off in a coordinated and automatic manner, depending on their workload. These are all areas for the application of IIoT. At production, the introduction of IIoT implies the structural formation of a digital counterpart of the products produced. As a result of the IIoT application, the total percentage of product rejects is reduced, the main factors influencing the appearance of the rejected products are revealed, the efficiency of the technological chain of production lines is enhanced, it becomes possible to carry out more high-quality process monitoring, it is possible to more clearly trace the product creation chain and optimize the process of the production line.

In order to avoid downtime and to maintain safety at the enterprise, it is necessary to introduce technologies to detect and predict risks, deviations. Continuous proactive monitoring of key indicators makes it possible to identify the problem and take the necessary measures to solve it. For the convenience of operators, modern systems allow you to visualize the conditions of the flow of technological processes and identify factors that affect them. Thanks to such solutions, production data is transformed into useful information that is necessary for the safe and rational management of an enterprise. The introduction of such technologies enables enterprises from different sectors of the economy to gain certain advantages: to increase the efficiency of the use of production assets by reducing the number of unplanned downtime; Reduce maintenance costs by improving procedures for predicting and preventing catastrophic equipment failures and identifying inefficient operations; Increase productivity, increase energy efficiency and reduce operating costs through more efficient use of energy.

Accordingly, it is necessary to create a management system that will automatically monitor, anticipate and prevent possible failures in production. In other words, there is a need of automatic control system for early detection of anomalies in the system.

II. ANOMALY DETECTION TASK

The problem of detecting anomalies today is very relevant. The cases of rejection of any parameters from the norm always require special attention of responsible persons. For example, when analyzing suspicious banking transactions, you can find scammers; revealing deviations from the norm of any parameters in the production, it is possible to determine the defective product; unusual results of medical examination can signal the patient's health problems; in seismology anomalies are a sign of possible cataclysms; non-standard readings of sensors may indicate the occurrence of malfunctions in technical systems; Antiviruses most often find suspicious and undesirable programs for unusual activity.

The automatic control system (ACS) supports or improves the operation of the managed object. In a large number of cases, auxiliary operations for ACS (start, stop, control, commissioning, etc.) can also be automated. ACS operates mainly as part of production or some other complex. Automatic systems used in the automation of production processes, depending on the kind and amount of operations performed by them, can be divided into systems of automatic control, automatic regulation, automatic managing, tracking systems, automatic protection, adaptive, etc. Automatic systems can be combined, i.e. represent a collection of several systems. Automatic systems can also differ in the types of devices used in them, parameters, design solutions, etc.

Technologies do not stand still, they are improving and today there are many tools to diagnose abnormal behavior, for example, the above-mentioned antivirus software, various monitoring and control systems. One of the best methods used in such systems are, in my opinion, the methods of machine learning.

Machine learning is a class of methods of artificial intelligence, the characteristic feature of which is not a direct solution of the problem, but training in the process of applying solutions to a set of similar tasks. When applying these methods, the system “learns” the independent finding of deviations, which reduces the need for human intervention in technological processes.

In order to predict the possible occurrence of a malfunction in the automatic control system, it is necessary to determine the criterion that will be an indicator of this failure. As a rule, any malfunction is accompanied by a change in the operation of the system, and hence its various output values. That is, when anomalous values appear in the system, it can be regarded as a sign of its failure. This conclusion is only half true. Anomalies can also be due to interference effects, voltage surges, and short-term non-critical equipment failures. All these processes are undesirable, but nevertheless, they are not a cause for system repair. The task of predicting possible failures is to determine a certain threshold for the number of anomalies that arise, when exceeding it, it is necessary to give a signal about a failure in the system. The definition of this threshold is individual for each system, and in automatic control systems it must be done using intelligent methods. These methods include the use of neural networks, which allows you to analyze the operation of the system and learn from its values to more accurately determine the required values.

Artificial neural network (INS) is a mathematical model, as well as its software or hardware implementation, built on the principle of organizing and functioning of biological neural networks — nerve cell networks of a living organism. Artificial neural networks are a system of connected and interacting simple processors (artificial neurons). Such processors are usually quite simple (especially in comparison with processors used in personal computers). Each processor of such a network only deals with the signals it periodically receives, and the signals it periodically sends to other processors. And, nevertheless, being connected to a sufficiently large network with controlled interaction, such separately simple processors together are able to perform rather complex tasks. Neural networks are not programmed in the usual sense of the word, they are trained. The possibility of learning is one of the main advantages of neural networks over traditional algorithms. Technically, training is to find the coefficients of connections between neurons. In the process of learning, the neural network is able to identify complex dependencies between input data and output, and perform generalization. This means that in case of successful training, the network will be able to return the correct result based on data that was not available in the training sample, as well as incomplete and / or “noisy”, partially distorted data.

Let us consider the problem of finding anomalies by the method of machine learning [4]. There is an input data set representing the readings of the temperature and load sensors of the system, depending on the operating time of this system.

Table 1. Example of input data

Temperature, °C	13	12	15	18	27	17	...	15
Time, min.	1	2	3	4	5	6	...	300

The average value of this data, for this example, is ~ about 15 °C.. It is necessary to find deviations in the set that are indicators of the malfunction of the system, which in turn can mean the presence of malfunctions, improper operation of personnel, breakdowns, etc.

One of the key points of this task is to determine the conditions of anomaly, that is, what value should be considered as a deviation, and which is not. In this set there are values that differ significantly from the average, but are not anomalies in fact. These values can be considered, for example, 11 and 18. These values appear due to inaccuracies in equipment, the effect of interference, minor voltage surges are not a malfunction signal. Consequently, the problem arises of determining the correct threshold, when exceeding it; the value must be considered an anomaly.

III. PROBLEM SOLVING ALGORITHM

To solve this problem, let’s use a method based on the Gaussian distribution. A normal distribution, also called the Gaussian distribution or Gauss-Laplace distribution, is the probability distribution, which in the one-dimensional case is given by a probability density function that coincides with the Gaussian function. The important value of the normal distribution in many fields of science follows from the central limit theorem of probability theory. If the result of observation is the sum of many random weakly

interdependent quantities, each of which contributes a small contribution to the total sum, then as the number of summands increases, the distribution of the centered and normalized result tends to normal. This law of probability theory has a consequence of the widespread distribution of the normal distribution, which became one of the reasons for its denomination. Therefore, its use to solve the problem of finding anomalies is optimal [5].

The program analyzes the entire input data set and calculates the mathematical expectation μ and variance σ^2 for each value using the following functions:

$$\mu = \frac{1}{m}\sum_{i=1}^{m} x^{(i)}$$

$$\sigma^2 = \frac{1}{m-1}\sum_{i=1}^{m}(x^{(i)} - \mu)^2$$

Then the probability density of the Gaussian distribution for the whole set is determined.

$$p(x) = \prod_{j=1}^{n} p(x_j, \mu_j, \sigma^2{}_j)$$

$$= \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(x_j - \mu_j)^2}{2\sigma^2{}_j}\right)$$

Indications having a lower probability density are more likely an anomaly. On the basis of the received array, the optimal value of threshold is calculated, and then for each value from the input set the condition is checked if $< \varepsilon$ to identify the deviation. Thus, we get a program that learns the independent definition of the optimal anomaly condition.

IV. ALGORITHM IMPLEMENTATION

To implement this algorithm, let`s use the Matlab software solution.
Load sensor data readings:

```
load('data.mat');,
```

where data is a randomly composed set of data in a certain range, with examples of anomalously large and small values.

Using the Gaussian distribution, we find the optimal value of the threshold ε. For this, apply the functions *estimate Gaussian( ), multivariate Gaussian( )* and *select Threshold ( ).*

The estimate Gaussian ( ) Function, returns the values of μ — mathematical expectation and σ^2 — dispersion necessary for further calculations:

```
[musigma2] = estimateGaussian(X);
```

Where μ is mathematical expectation, σ^2 is dispersion.

The *multivariate Gaussian ( )* Function returns the density of the Gaussian distribution:

```
p = multivariateGaussian(X, mu, sigma2);

pval = multivariateGaussian(Xval, mu, sigma2);
```

The *select Threshold( )* Function, finds the optimal threshold ε using a set of cross checking, taking into account the found values of the Gaussian distribution density.

```
[epsilon F1] = selectThreshold(yval1, pval);
```

Then find deviations in the data set using the built-in Matlabfunction — find. F value that should be considered as a deviation is any value whose probability density is less than the required threshold ε:

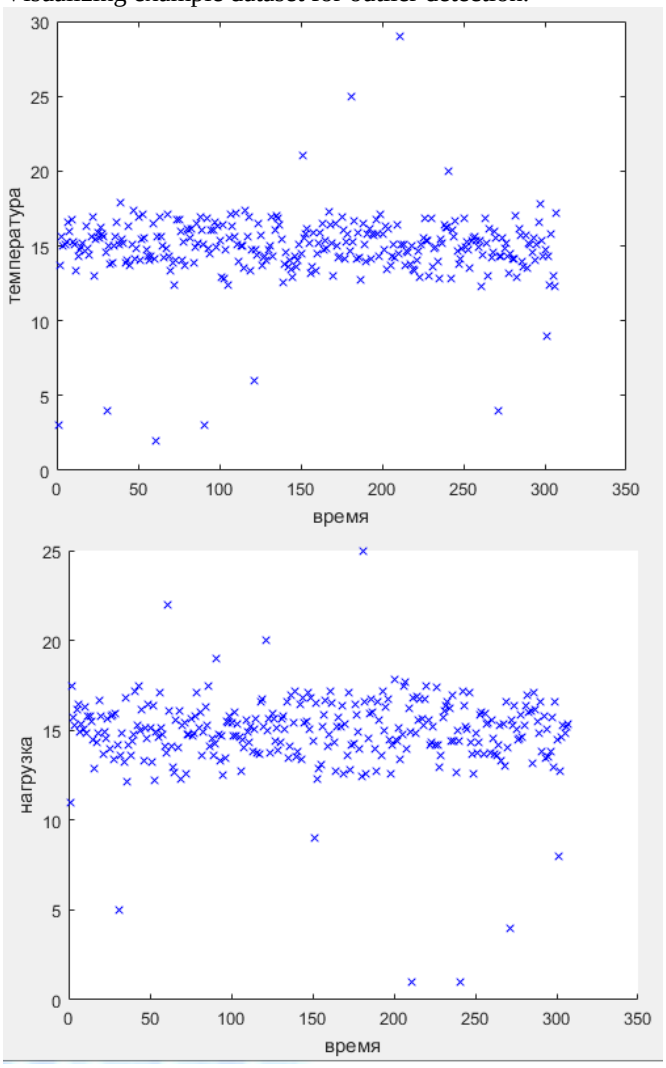
```
outliers = find(p<epsilon);
```

V. RESULTS OF THE PROGRAM WORK

After all necessary calculations, the program displays:

- Initial data set in the form of graphs

Visualizing example dataset for outlier detection.

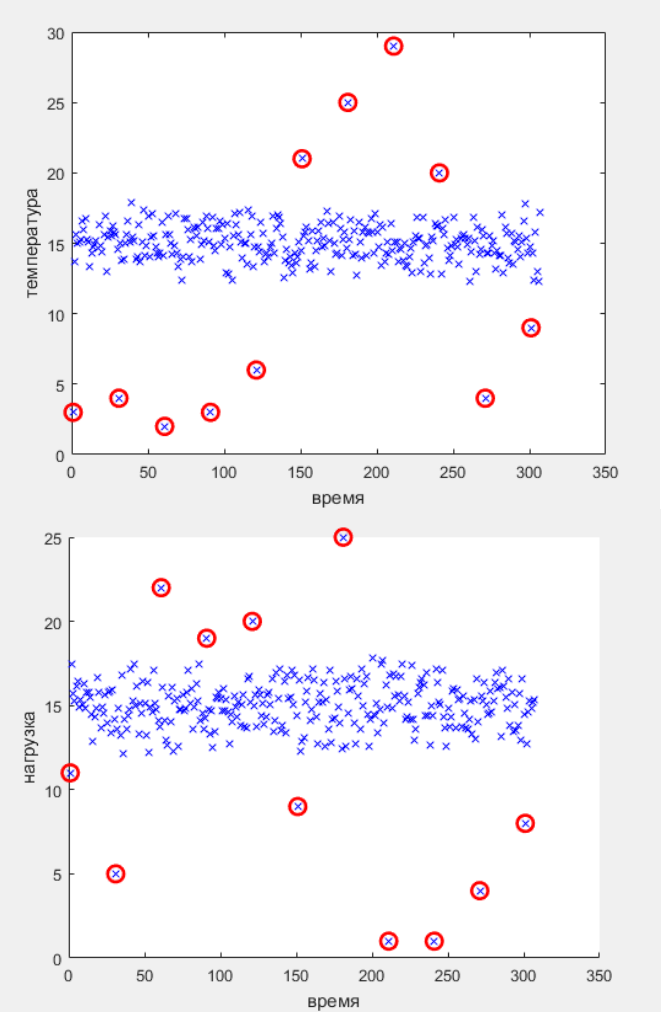


Program paused. Press enter to continue.

- The obtained value of the threshold ε.

Best epsilon found using cross-validation: 6.595492e-05.

- The resulting graphs, in which all the anomalies found are marked with red circles.



VI. ANALYSIS OF THE RESULTS

Analyzing the results of the program, it can be concluded that the problem of detecting anomalously large or small values of the time series is solved. According to the compiled algorithm for solving the problem, written a program that uses the Gaussian distribution to determine the deviations. During the testing of the program it showed good results, because it found all the values that can be considered as abnormal. In addition, the results of the work are displayed on the graphs in a convenient form.

VII. CONCLUSIONS

In this paper, the problem of detecting anomalies by methods of machine learning was analyzed; the basic concepts and theory of this direction are designated, such as: industrial Internet of things, machine learning, neural networks; a teaching method based on the Gaussian distribution is described. The experimental part was the writing of a program to search for anomalies. The resulting program shows good results of finding anomalies, and allows you to visualize the results for convenient work with them.

VIII. REFERENCES

1. Воронова Л.И. Интеллектуальные базы данных: Учеб. пос. М., 2013. [Voronova L.I. Intelligent databases. Moscow, 2013].
2. Воронова Л.И., Воронов В.И. Big Data. Методы и средства анализа: Учеб, пос, / М., 2016. [Voronova K.I., Voronov V.I. Big Data. Methods and tools for analysis. Moscow, 2016].
3. Воронова Л.И., Воронов В.И. Machine Learning: регрессионные методы интеллектуального анализа данных: Учеб. пос. / МТУСИ. М., 2018. [Voronova L. I., Voronov V. I.. Machine Learning: Regression methods data mining: A tutorial / MTUCI. Moscow, 2017].
4. Stanford University Machine Learning // Coursera. URL: https://www.coursera.org/learn/machine-learning.
5. Беляев А.В., Петренко С.А. Обнаружение аномалий в ERP системах // Труды Института системного анализа РАН. [Belyaev A.V., Petrenko S.A. Anomaly detection in ERP systems // Proceedings of the Institute for System Analysis of the Russian Academy of Sciences]. URL: www.isa.ru/proceedings/images/documents/2006-27/130-154.pdf.

Application of Neural Networks for Object Recognition in Video Surveillance Systems of Industrial IoT

Andrew M. Tolkachev[1], Natalia V. Toutova
Department of Intelligent Systems in Management and Automation
Moscow Technical University of Communication and Informatics
Moscow, Russia
[1] dezirtir96@gmail.com, [2]e-natasha@mail.ru

***Abstract*. In this paper the use of neural networks for object recognition in video surveillance systems of the industrial Internet of things is discussed, as well as a comparison of two types of neural networks: a perceptron and a convolutional neural network with regard to the problem of automatic recognition of the queue of people in a supermarket or a bank. The principle of operation of neural networks, their structure and realization are described. Two types of neural network are compared. It is shown that the convolutional neural network surpasses the perceptron in recognition accuracy and it is expedient to use it in the automatic queue recognition system.**

Key words: **neural network, recognition, live queue, perceptron, video surveillance systems.**

I. INTRODUCTION

People recognition is the key component of video surveillance systems, which are widely used for security purposes. Another application is a queue control systems in banks and supermarkets to control and improve quality of service as well as automates goods delivery points.

It is designed using industrial Internet of things (IIoT) concept. Among its components are cameras along with microcomputers, which are running recognition program. They analyze video stream and count the length of the queue in the bank or the supermarket. Then it periodically sends data to the cloud database. The data include time, the queue length and proving pictures. Then the bank or supermarket manager can analyze the data offline and make decisions to improve quality of service.

The recognition is very important part of such systems. Mot recognition algorithms are based on neural networks. Neural networks are models of biological neural networks of the brain, in which neurons are simulated with relatively simple, often identical, elements (artificial neurons) [1; 2].

They are also used in automation of processes of pattern recognition, adaptive control, functional approximation, prediction, the creation of expert systems, the organization of associative memory and many other applications. A wide range of tasks solved by neural networks does not allow to create universal, powerful networks at the present time, forcing to develop specialized networks functioning according to various algorithms.

For years, long queues have become a very serious problem in the life of society. Every person a day at least once meets this problem: in a supermarket, a snack bar, a ticket office. The constant standing in the queues takes away a huge amount of precious time, which is constantly lacking. Recognizing the live queue in the supermarket will help automate the goods delivery points. Automation of points of delivery of goods is a current direction [4].

To develop such a system, it is necessary to choose the structure of a neural network. In this paper, two types of neural networks are compared: a perceptron and a convolutional neural network as applied to the problem of automatic recognition of the queue of people in a supermarket or a bank.

II. MODELS OF THE NEURAL NETWORK

The perceptron is one of the first models of neural networks. The perceptron consists of three types of elements, namely: the signals coming from the sensors are transmitted to the associative elements, and then to the reacting elements. Thus, perceptrons allow us to create a set of “associations” between input stimuli and the necessary reaction at the output. Biologically, it corresponds to the transformation, for example, of visual information into a physiological response from motor neurons. The perceptron is based on a mathematical model of information perception by the brain. Different researchers define it differently. In its most general form, it represents a system of elements of three different types: sensors, associative elements and reactive elements.

The first to work include S-elements. They can be at rest (signal is 0) or in the excitation state (the signal is 1). Then the signals from the S-elements are transmitted to the A-elements along the S-A bonds. These links can have weights that are only –1, 0, or 1.

Then the signals from the sensor elements passed through the S-A bonds fall into the A-elements, which are also called

associative elements. A single element can correspond to several elements. If the signals arriving at the A-element collectively exceed a certain threshold θ, the A-element is excited and gives a signal equal to 1. Otherwise (signal from the S-elements did not exceed the threshold of the A-element), a zero signal is generated.

A-elements are called associative, because A-elements are aggregators of signals from sensory elements. For example, we have a group of sensors, each of which recognizes a piece of the letter "D" on the picture under study. However, only their totality (i.e., when several sensors output a signal equal to 1) can excite the A-element entirely. The A-element does not react to other letters, only to the letter "D". That is, it is associated with the letter "D". Perceptron is used for pattern recognition, weather forecasting, etc.

The neural network is a special architecture of artificial neural networks, proposed by Jan Lekun in 1988 and aimed at effective image recognition, is part of the technology of in-depth training. It uses some features of the visual cortex, in which so-called simple cells that react to straight lines from different angles were opened, and complex cells whose reaction is associated with the activation of a certain set of simple cells. Thus, the idea of convolutional neural networks is to alternate convolutional layers and subsampling layers. The network structure is unidirectional, essentially multilayered. For training, standard methods are used, most often the method of back propagation of the error. The function of activation of neurons is at the choice of the researcher.

Differences:

- Application of the principle of shared weighting coefficients, where the S-type neurons divide the weight coefficients with other neurons of the same layer; this approach allows to significantly reduce the number of free parameters of the neural network and reduce its resource intensity in relation to the required RAM size.
- The multi-layer perceptron architecture contributes to the possibility of reducing the computational complexity of the network by reducing the necessary amount of synaptic connections.
- Perceptron is needed for light tasks (by type to recognize the letter), and the convolutional network allows to optimize the neural network, i.e. first, for example, classifies the object (square), then details the object (road sign), and at the end concretizes the object (the sign "Entry prohibited"). And deep training itself determines how many layers to make, what will be the identifiers. Thus, it determines how to find the road sign and specify it.
- A convolutional layer in a neural network is just a layer that allows you to reduce the dimension of the feature card (features are called features in English literature and lectures). Convolutions are not the opposite of deep neural networks, deep neural networks are simply neural networks with a large number of layers, compared to the perceptron, and that's all.

As a model of a neural network, a perceptron was selected (Fig. 1). In the perceptron there is no feedback and, consequently, memory, all elements of the previous layer are associated with all elements of the next layer using weights calculated using the training set. Hidden layers are involved only in calculations, their number and dimensions affect the accuracy of the neural network and the power required for the calculation [5].

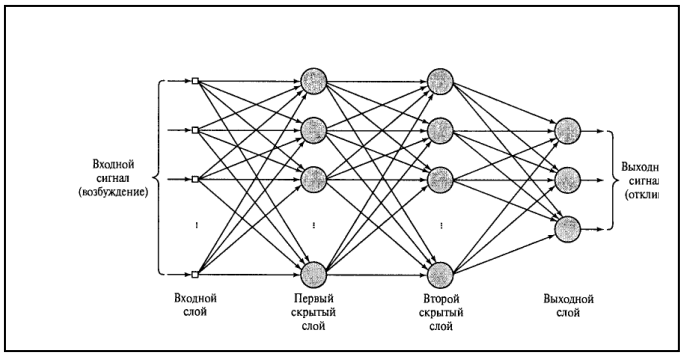


Fig. 1. General view of perceptron

The experiments were carried out in the mathematical package Octave 4.4.1. For realization, a training set consisting of black and white images in the size 20x20 was collected. Each pixel is encoded with one of the shades of gray, taking a value from 0 (black) to 255 (white). An example of an image and a matrix representation of its fragment are shown in Fig. 2.

Fig. 2. An example of a queue image and a matrix representation of its fragment

It is necessary to select extremely carefully images for the training set. We should take into account a number of conditions: different illumination of the room; the camera lens, foreign objects, for example racks or a shopping trolley, may fall; a malfunction of the camera itself, as a result of which noise can be observed in the image (Fig. 3) or "broken pixels".

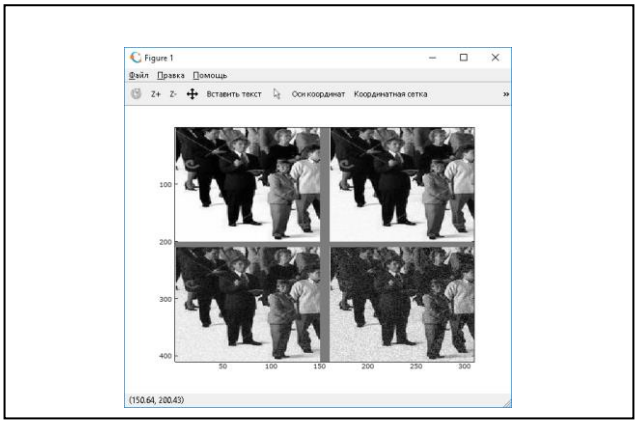


Fig. 3. An example of distortions in the image of a live queue, there is noise

III. TRAINING AND ESTIMATION OF ACCURACY OF NEURON NETWORK

The training set consists of 200 images, 150 of which are distorted copies of the first 50, obtained programmatically, by changing the matrix parameters of the image, and is divided into classes 120 with a live queue, 80 without a queue. Each example includes 400 parameters (the product of width and height), as well as the value of the output parameter that corresponds to the presence of a live queue, "1" is a live queue, "0" is no live queue. To further test the operation of the neural network, the training set is divided into a training set and a test set in 70% to 30% ratios. As a result, a training set of 140 examples was obtained (90 with a live queue, 50 without) and a test set of 60 examples (30 for each class). As a neural network architecture, a perceptron with one hidden layer was chosen. Experimentally, it was decided to use 5 elements of a hidden layer, the results of experiments in the form of a comparative graph are shown in Fig. 4. Training was conducted by the method of back propagation of the error.

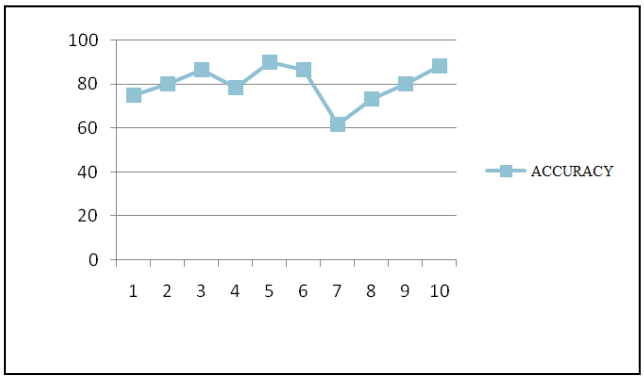
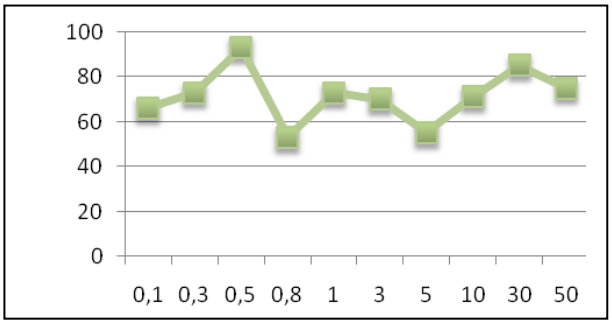


Fig. 4. Changing the accuracy of a neural network with changing the hidden layer

Before learning the neural network, it is also necessary to specify the initial values of weight coefficients that determine the weight of each element of the previous layer with respect to the element of the next layer, and the errors necessary for the back propagation method of the regularization coefficient, which affects the speed and correctness of the algorithm operation, and the termination conditions of the method. Although the weight coefficients are due to the method of back propagation of the error, in order to avoid the symmetry of the coefficients, it is necessary to specify certain initial values that are different from each other. To do this, we perform a "random initialization" assigning weight coefficients random values from the range [–ε; ε]. Thus, the weighting coefficients will take random values: $-\varepsilon \leq \Theta ij\ (l) \leq \varepsilon$. In this problem, ε was chosen to be 0.005. The regularization factor (λ) influences the work of the training method: in its absence, the neural network can be trained to work correctly only with these identical training; If the coefficient is too small, the method will require much more iterations to find the weighting coefficients; if the coefficient is too large, the method may not find the optimal result, since at each iteration it will "step over" through the solution. The regularization factor (λ) can be dynamic or static, at the initial stages of training the neural network is used static, since this simplifies the possibility of manual verification of the results. The initial value of the coefficient, obtained experimentally (Fig. 5), λ = 0.5.

Fig. 5. The accuracy of the neural network from the regularization coefficient

A condition for stopping the work of a learning method can



be a given finite number of iterations or conditions for achieving an error, when as a result of the following iterations there is a slight change in the values. In this problem, as a stopping condition, it was decided to use 100 iterations of the method of back propagation of the error. The accuracy of the neural network was calculated on the number of correctly predicted classes with respect to all classes.

IV. TESTING DEVELOPED NEURAL NETWORK AND COMPARISON WITH ANOTHER MODEL OF NEURON NETWORK

For testing, we chose an image (Fig. 6), which does not apply to either the training set or the test set, and is a queue.

Fig. 6. Image for testing the developed program

The function "imread" was loaded with an image, and also loaded the optimal values of the weights obtained earlier. For classification, the function "predict" was used, which was fed

to the input image matrix and the matrix of the weights. At the output of the neural network we get the probability that the image belongs to this class.

The test result is shown in Fig. 7.

Fig. 7. The result of the developed program

The program recognized this image as a live queue which it follows that the network is working correctly.

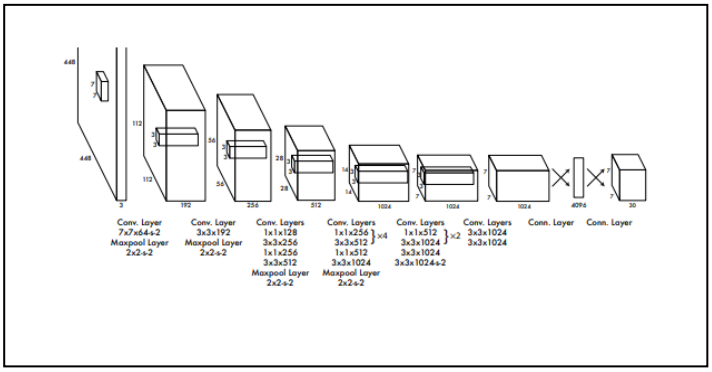
We will conduct such an experiment using the open library of computer vision Yolo, written in C.

Yolo uses a different approach. A single neural network is applied to the full image. This network divides the image into regions and predicts the limiting fields and probabilities for each region. These limiting fields are weighted by the projected probabilities.

The abbreviation YOLO stands for You Only Look Once. This model imposes a grid on the image, dividing it into cells. Each cell attempts to predict the coordinates of the detection zone with a confidence estimate for these fields and the probability of the classes. Then, the confidence estimate for each detection zone is multiplied by the probability of the class to obtain a final rating [6].

Fig. 8. The architecture of the Darknet YOLOv3 neural network

The testing was carried out on a computer with the operating system Ubuntu Linux 16.04 with updated drivers installed



library OpenCV 3.4.3.

In the linux console, we enter the commands for installing the neural network Darknet:

```
git clone https://github.com/pjreddie/darknet
cd darknet
make
```

Next, we load the weights YOLOv3 command:

wget https://pjreddie.com/media/files/yolov3.weights

We load the image with the name WH7GpHNnzpg.jpg into the ./darknet/data/ directory and start the processing process with the command:

```
./darknet detect cfg/yolov3.cfg yolov3.weights
data/WH7GpHNnzpg.jpg
```

After image processing, we get the following result:

Fig. 9. Result of image processing in YOLOv3

V. CONCLUSIONS

A neural network in the form of a perceptron with reasonable accuracy can only give out the fact of having a queue, so this type of neural network is not suitable for use in queue control systems.

The Darknet YOLO convolutional neural network surpasses the perceptron in recognition accuracy and it is advisable to use it in the queue control systems.

VI. REFERENCES

[1] Maas A. L., Hannun A. Y., Ng A. Y. Rectifier Nonlinearities Improve Neural Network Acoustic Models // Proceedings of the 30 th International Conference on Machine Learning. Atlanta, 2013. Vol. 28.

[2] Воронов В.И., Воронова Л.И. О повышении результативности магистерских программ в условиях инновационной экономики // Инновационные подходы в науке и образовании: теория, методология практика / Под общ. ред. Г.Ю. Гуляева. Пенза, 2017. С. 35–44. [Voronov V. I., Voronova L. I. Increasing productivity master's program in the innovation economy. In: Innovative approaches in science and education: theory, methodology, practice / Ed. G. Yu. Gulyaev. Pensa, 2017. P. 35–44].

[3] Воронова Л.И., Воронов В.И.. Machine Learning: регрессионные методы интеллектуального анализа данных: Учеб. пос. / МТУСИ. М., 2018. [Voronova L. I., Voronov V. I.. Machine Learning: Regression methods data mining: A tutorial / MTUCI. Moscow, 2017].

[4] Marmanis H., Babenko D. Algorithms of the Intelligent Web. Greenwich: Manning Publications Co., 2009.

[5] Brink H., Richards J., Fetherolf M. 2016. Real-World Machine Learning. Greenwich: Manning Publications Co., 2016.

[6] Agrawal P., Girshick R., Malik J. Analyzing the performance of multilayer neural networks for object recognition // ECCV. 2014. P. 329–344.

Real-Time Recognition System of Sign Language in the Industrial Internet of Things

Artem A. Goncharenko[1], Lilia I. Voronova[2]
Department of Intelligent Systems in Management and Automation
Moscow Technical University of Communication and Informatics
Moscow, Russia
[1]artem@agsys.ru, [2]voronova.lilia@ya.ru

***Abstract*. Industrial Internet of Things implies the concept of combining all electronic devices into a single network in an industrial enterprise, where each device can communicate with each other. The task of people in such an enterprise is to control automatic systems. In addition, enterprises of increased complexity will always need people. The article describes the process of preparing and partially implementation sign language translation in real time that can be embedded in an enterprise. The choice of architecture, recognition methods and programming languages used to create the system is considered. The system was decomposed into functional subsystems.**

***Key words:* IoT, recognition systems, sign language, real-time systems, API interface, Industrial Internet of Things.**

I. INTRODUCTION

The Internet of things implies the concept of combining all electronic devices into a single network, where each device can communicate with each other. Industrial Internet of things is no exception, but do not forget that people will always be present at any enterprises and industries. Their task is and will be to control automatic systems. In addition, enterprises of increased complexity, where the use of the Internet of things can be limited, for example, steelmaking, the aviation industry or military enterprises will always need people [1; 2].

The world community has taken a direction to develop support for people with disabilities. Due to this, more and more appears on the market of various solutions for the comfortable adaptation of people with various health problems. The development of the industry does not stop at creating physical assistants, for example, a smart home or implantable hearing aids. Various software, automatic and automated systems are being developed that perform various roles to support people.

II. DEVELOPMENT OF THE SIGN LANGUAGE RECOGNITION SYSTEM

The problem of creating an automated sign language recognition system in real time is solved at the MTUCI department of the Intelligent Systems in Management and Automation. The development will allow to expand possibilities of dialogue with surrounding world of people with hearing and / or voice violation without the help of sign language interpreter [3–5].

The purpose of this project is to create an accessible tool in mass use that does not require special additional equipment. Expanding the goal for the industrial Internet of things is the creation of ready-made software that can be installed in enterprises and used in translate a sign language in real time.

The tasks necessary to solve the project are to introduce an application into the mobile device or computer of a normal user, allowing effortlessly translating the displayed gestures by the user to the video camera of the device into text. The received text can be transferred to the interlocutor by sending a message, or reproduced, due to widely used voice assistants.

III. SELECTION OF METHODS OF RECOGNITION AND ARCHITECTURE SYSTEM

To recognize gestures using a mobile device, a 2D image processing method is used. Compared with the 3D method, the 2D method is inferior in quality of recognition, however, for the 3D method to be used, additional equipment is needed, which is contrary to the purpose of the project. The essence of the 2D method is to transform the input image, so that there remains a contour or contour of the hand. This transformation can be achieved by dividing the image into color channels, converting it to monochrome, removing the background and highlighting the outline of the hand. By this principle, all the images collected for the training set of neural networks will be transformed.

The system architecture is chosen the most suitable for expansion and for the distribution of computing resources of the system hardware. Using the API architecture, you can combine different recognition technologies based on neural networks. The most popular and fast programming language for implementing neural networks is Python. Applications installed on users' devices work in various programming languages, such as Java, C #, Swift. The browser solution of the client system can be implemented in the programming languages PHP, Java and JavaScript, or in other languages used in the web. Therefore, to ensure the correct interaction of all components of the sign language recognition system in real time, the API architecture was chosen.

One of the most difficult tasks for creating a system is the implementation of the recognition algorithm. This algorithm is divided into several stages:

- detecting a hand in a frame;
- filtration and selection of hands;
- recognition of the gesture in the frame;
- checking and adding the sequence of gestures from the video stream to a single phrase or sentence.

The step of detecting the hand involves the use of a trained neural network that can select a hand at different angles and on different backgrounds. To train this neural network you need a large amount of training data set. All hands of people are

different, skin color, pigmentation, background depends on the direction of the incident light. To obtain the necessary training set, not one hundred images are required.

To recognize the gesture, the largest training set is required, rather than for detecting the hand. All people who communicate in the language of deaf-mutes do it with the speed of speech of a healthy person. Therefore, it is necessary to take into account the lubrication of the hand movement in the frame, and, consequently, the training set should also provide lubrication.

IV. System Architecture

The architecture of the system involves the interaction of two software packages installed on the user's device and on the data processing server. We suppose that our company has a conference room in which tablets with a video camera are installed, information screens that will display the result of the translation and a server that will provide input and output devices and will recognize the sign language. With this distribution of hardware running on different operating systems, the server will provide data in a standardized format for the system. This data will be converted into a display interface on each of the devices due to the client application. All these actions are possible at the expense of the chosen architecture using the API of the subsystem interaction interface.

The life cycle of the system is shown in Fig. 1. Showing gestures to the device's camera, the user starts processing and transferring video to the server through the installed application. The server performs recognition and response generation and sends it to the application, which displays the transfer information on the screen.

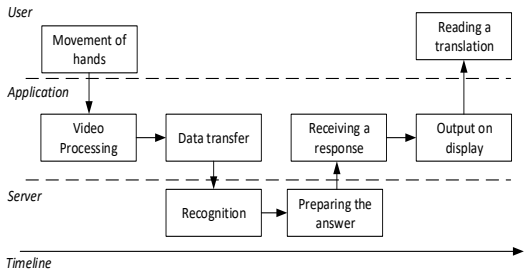


Fig. 1.The life cycle of gesture recognition

To understand the necessary subsystems of the system and their functional features, you need to decompose the system server software and the client application. When decomposing the server software, we have a number of subsystems:

The recognition subsystem is responsible for processing and recognizing the data [6–8]. It includes the following subsystems:

- The image processing and video processing subsystem responds and performs functions for detecting the hand, color processing, framing, bringing the data into a common format for the system, as well as temporarily storing the materials on the server. It has subsystems: compression, detection, storage [9–11].
- The recognition subsystem is responsible for recognizing gestures based on the prepared data, which is provided by the image processing subsystem and video.
- The answer generation subsystem is responsible for combining the received recognition results and converting them into the correct lexical format.

The learning subsystem performs the functions of training the neural networks of the system, as well as collecting data for training. It consists of subsystems:

- The data collection subsystem is responsible for collecting the data of users who have been misinterpreted, as well as collecting data in a special system learning mode, in which users can independently display phrases and write their values in text form.
- The retraining subsystem is responsible for training and retraining neural networks of the system, based on data collected by the data collection subsystem.

The access control subsystem is responsible for limiting and controlling access. Since the system is multi-user and will be used in the live chat mode, this subsystem will ensure the confidentiality of user information. It includes:

- subsystem of registration which makes registration of new users in the system and collecting the necessary information, for example, health data, if you need to urgently call the ambulance service;
- subsystem for authorization which checks the availability of registered user data in the system and the ability to access it;
- authentication Subsystem which performs an access check on every access to the system's actions.

The subsystem of data collection and analysis solves the problems of collecting data used Internet traffic, the amount of load on the computing power of the server. The entire collection of these data will provide an understanding of the need to distribute computing resources, or upgrade the system. This information is necessary, because the system must allow real-time translation and its delays are significant for users.

The operating system for the system server is selected by the Linux OS family. It is freely distributed, has a large community of developers. With the help of community repositories, you can install all necessary software through the Linux console to implement the server. The software for the MySQL database, the PHP and Python language interpreters were installed separately. Installed software for remote access to the server via SSH.

SSH is a network protocol that allows you to create a secure remote connection and manage the server. All transferred data and files are encrypted when connected via SHH, so using SSH you can create a secure connection in an unprotected environment.

V. WEB SERVER AND API

The interaction of the user through the application of the system with the server of the system will be carried out according to the following algorithm (Fig. 2): the application transmits the request to the web server, which in turn uses the interpreters of the PHP or Python languages to process the data and return the result to the client application.

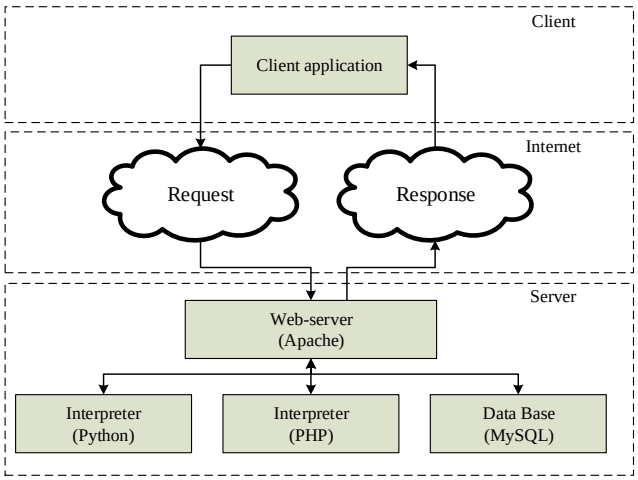


Fig. 2. Process of interaction between the application and the server

Also, for the server to work in the mode of accepting requests from the Internet, you must select and install the web server software. The Apache software is selected as the web server. It provides the use of PHP, Java and Python. Also, this web server is freely distributed, which is convenient for starting development. Python is chosen as the programming language for implementing the execution of neural network algorithms. In this language, the recognition and detection are performed most quickly. Also for Python, there are various libraries that will make it easier to process input images and create neural networks. Interaction with the database and provision of the API work is implemented in the PHP language. In the future, it is planned to translate the development of the interface into the Java programming language. At the moment, the API design of the interface is implemented using the HTTP protocol and looks like this:

https://<domain name>/<API version>/<called method>

To further protect the transmission of data, the HTTPS protocol is used. This protocol is an extension of the HTTP protocol and allows you to use data encryption with the help of special digital certificates. To use the HTTPS protocol, you must obtain a certificate that is issued by specialized certification centers on the Internet. A certificate is issued for each domain name and, depending on the type of certificate, to its subdomains, thereby providing an acknowledgment of access to the domain. There are certificates created independently, they are called “self-signed”, but the responsibility for encrypting the data and security of such a certificate lies with the developer and may be vulnerable. To get a free HTTPS certificate, the global online community offers users a non-profit certification center. Let’s encrypt it. This center issues certificates for a period of 3 months, after which it is necessary to renew the certificate. Its advantage is also that the entire process of obtaining a certificate is automated by installing software on the server to which the domain name belongs.

Recently, all modern browsers prohibit access to the webcam of the device, in the event that the domain name does not have an encryption certificate. Since one of the clients of the system is a web-browser interface, this fact must be taken into account in the development.

In multi-user systems using the API interface, several versions must be considered. With the operation of any program over time, there are errors or the desire for changes, updates from users. In order not to interrupt the system and avoid its inoperability when implementing changes and changes in the current version, it is necessary to separate and conduct version development separately from each other. Therefore, one of the parameters of accessing the interface is its version.

The called method is a pointer to the system, which components and modules need to be connected to perform the requested operation. All methods are described in the structure of the API version of the interface and can be extended or changed when the interface version is updated and changed. In each version of the interface for the sign language recognition system, there is a special router file. This is a kind of router, which describes the interaction and plug-ins, to process the request.

All parameters necessary for processing by subsystems are transmitted to the HTTP server by the method of sending POST. The HTTP protocol supports POST and GET, as well as PUT, DELETE and a number of other methods. However, the most common methods are GET and POST. The difference between them is such that, with the GET method, all the transmitted data is in clear form directly in the URI request. When using the POST method, all transmitted data, including files, is transmitted “hidden” while in the request body.

In response to the request, the API interface connects and calls the necessary methods that ensure the operation of the subsystems and returns a message in the JSON format. The structure of the answer consists of 3 parts. The first part reports on the presence of errors during the execution (error = <0 — no errors, 1 — errors occurred>), the second part transmits a digital number or the number of errors (errorCode = <error description number>), the third parameter transmits the result of the request processing (response = <Processing result>). The received response can be converted for further requests to the system or for performing the output of the results on the device screen by the system client.

VI. INTRODUCTION OF A VIDEO RECORDING AND TRANSMISSION SYSTEM

One of the tasks of implementing the system is to transfer the video stream to the data processing server. The difficulty of solving this problem lies in the fact that if the quality of the Internet connection is poor, the quality of the transmitted video may be degraded or the time delay may occur while data exchange takes place. When solving it, you need to take into account these factors and look for the most acceptable solutions:

- Compress and transfer video. When compressing files, you can achieve a significant difference in the amount of source and compressed files. However, the compression of the images brings quite a few results comparing the output with the input size.
- Conversion to another format is an interesting approach, but it requires large computing resources and speed of the user’s device. It also uses a large amount of time, which in the real-time system is practically nonexistent.
- Cropping video is the most suitable option. To detect the hand and recognize the gesture, a small image size of 80x60 pixels is required. This conversion can be

performed on a user device without significant time delays.

In addition, it is necessary to decide how to transmit the received framed video stream. In the web browser client, there are several ways to implement:

- AJAX technology is built in the JavaScript programming language and allows you to send and receive data without reloading the main page. The downside of this technology lies in the constant multiple access to the server. The results of testing the AJAX technology showed that 7000 calls to the server are created per 1 minute of video transmission, and the server should accept and answer each of them (Fig. 3). This solution is not suitable for a real-time system, because there will be a delay due to frequent connections to the server when transmitting one video stream.

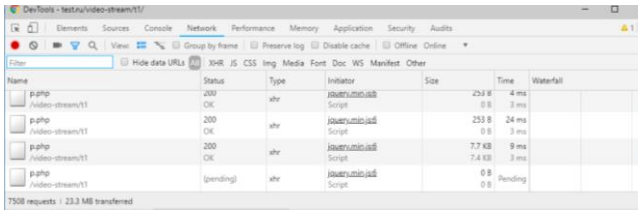


Fig. 3. Test result of AJAX technology

- WebSocket technology allows creating a direct "corridor" between the user and the server and listening to it, which does not create multiple calls to the server (Fig. 4). This technology is widely used in online systems. There are also disadvantages of the technology including being able to connect and "listen" to the server, you need to create a special WebSocket server.

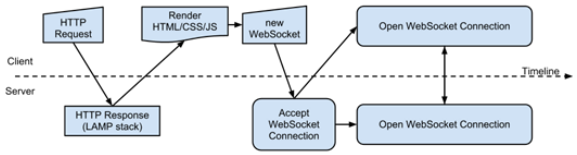


Fig. 4. The scheme of connection of technology WebSocket

VII. CONCLUSION

This article describes a variant of the real-time translation system of sign language that can be implemented in enterprises with industrial Internet of things. When creating this kind of multi-user systems, especially real-time systems using video broadcasting and recognition, a large number of subtasks arise that need to be solved during the implementation of the system and pay attention to various aspects. The article describes the chosen version of the system architecture, decomposition of the subsystems is made, methods for transmitting and processing the video stream are considered.

VIII. REFERENCES

[1] Voronov V.I., Voronova L.I. Features of realization master's program "Automation of technological processes and manufactures" // International Journal of Applied and Fundamental Research. 2016. № 2. URL: www.science-sd.com/464-25196.

[2] Безумнов Д.Н., Воронова Л.И. О развитии и стандартизации технологии интернета вещей // Технологии информационного общества: Материалы XII Международной отраслевой научно-технической конференции. М. 2018. С. 293–294. [Bezumnov D.N., Voronova L.I. On the development and standardization of the Internet of Things technology // Information Society Technologies: Proceedings XII International. branch scientific.-tech. conf. Moscow, 2018. P. 293–294].

[3] Михаеску С.В., Трунов А.С., Воронова Л.И. Анализ предметной области для разработки системы построения скелетной модели человека на основе массива опорных точек, получаемых совокупностью контроллеров Kinect // Международный студенческий научный вестник. 2015. № 3-4. С. 521–522. [Mihaesku S.V., Trunov A.S., Voronova L.I. Analysis of the subject area of system developing for building a human skeletal model based on an array of reference points obtained from a set of Kinect controllers // International Student Scientific Bulletin, 2015. N 3-4. P. 521–522].

[4] Клешнин Н.Г., Воронова Л.И. Применение нейронных сетей в подсистеме распознавания эмоций для проекта "Сурдотелефон" // Телекоммуникации и информационные технологии. 2018. Т. 5. № 1. С. 126–130. [Kleshnin N.G., Voronova L. I. Neural networks application in emotion recognition module for "surdotelephone" project // Telecommunications and information technology. 2018. T. 5. № 1. C. 126–130].

[5] Воронов В.И., Воронова Л.И., Генчель К.В. Применение параллельных алгоритмов в нейронной сети для распознавания жестового языка // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2018). VII Международная научно-техническая и научно-методическая конференция: Сб. науч. ст.: В 4 т. / Под ред. С.В. Бачевского. М., 2018. С. 207–212. [Voronov V.I., Voronova L.I., Genchel K.V. The use of parallel algorithms in the neural network to recognize the sign language // Actual problems of information and telecommunications in science and education (APINO 2018). VII International scientific tech. and scientific method. conf.: In 4 t. / Ed. S.V. Bachevsky. M., 2018. P. 207–212].

[6] Воронова Л.И., Воронов В.И. Big Data. Методы и средства анализа: Учеб. пос. М., 2016. [Voronova K.I., Voronov V.I. Big Data. Methods and tools for analysis. Moscow, 2016].

[7] Усачев В.А., Воронов В.И. Компетенция "машинное обучение и большие данные" // Приоритетные направления развития науки и образования / Под. общ. ред. Г.Ю. Гуляева. Пенза, 2017. С. 97–108. [Usachev V.A., Voronov V.I. The competence "Machine learning and Big Data" // Priorities for development of science and education / Ed. G.Yu. Gulayev. Pensa, 2017. P. 97–108].

[8] Горячев Д.В., Воронов В.И. Большие данные и машинное обучение // Технологии информационного общества: Матер. XII Междунар. отраслевой науч.-техн. конф. 2018. С. 327–328. [Goryachev D.V. Big Data and Machine Learning // Information Society Technologies: Proceedings of 12th the international scientific-practical conference. Moscow, 2018. P. 327–328].

[9] Tsai T.H., Huang C.C., Zhang K.L. Embedded virtual mouse system by using hand gesture recognition // Consumer Electronics-Taiwan (ICCE-TW); 2015 IEEE International Conference on 2015 Jun 6. Taiwan; Taipei; New York: IEEE, 2015. P. 352–353.

[10] Huong T.N., Huu T.V., Le Xuan T. Static hand gesture recognition for vietnamese sign language (VSL) using principle components analysis // Communications, Management and Telecommunications (ComManTel). International Conference on. 2015 Decembee 28. 2015. P. 138–141.

[11] Lee D.H., Hong K.S. A Hand gesture recognition system based on difference image entropy // Advanced Information Management and Service (IMS). 6th International Conference on 2010 Nov 30. Seoul; New York: IEEE, 2010. P. 410–413.

Analysis of Work of YOLO v.3 AND YOLO v.2 Neural Networks

Dmitriy P. Ayrapetov[1], Boris Y. Buyanov[2]
Department of Intelligent Control Systems and Automation
Moscow Technical University of Communication and Informatics
Moscow, Russia
[1]dmitry_a95@live.ru, [2]b.buyanov@gmail.com

***Abstract*. Currently, in the framework of industrial use for solving problems related to the safety of production, widely used computer vision systems, which, as a rule, use neural networks. The article represents an example of settings and operation within the software and hardware complexes of DarkNet Yolo V3 and V2 neural networks, providing the ability to quickly recognize objects. This is critical for making decisions automatically.**

***Key words:* neural networks, object recognition, machine learning, mathematical calculations, industrial Internet of Things.**

I. INTRODUCTION

Currently, in cooperation with the work on the development of artificial intelligence, there is an increased demand for the development of artificial systems based on work algorithms similar to neural networks.

Neural networks are models of biological neural networks of the brain, in which neurons are simulated with relatively simple, often identical, elements (artificial neurons).

Among the areas of neural networks application are automation of image recognition processes, forecasting, adaptive control, expert systems creation, associative memory organization, processing of analog and digital signals, synthesis and identification of electronic circuits and systems [1–3].

The development of modern neural networks is oriented to the processing of natural language, i.e. computer analysis of natural language and its synthesis.

Elimination of objects using learning algorithms solves problems more effectively than human eyesight. Convoluted neural networks are widely used in problems of classification, detection and recognition of images. Gradually, the range of these tasks is expanding, so the development of new architectures, layers of the network and modifications of software platforms is not lost.

Class is a certain group of objects in classification system that combines a certain set of objects by a certain feature or features. Perceiving the outside world, we always classify information, that is, we break them into groups of similar but not identical phenomena. For example, despite a significant difference, one group includes all the letters “A”, written in different handwritings or all sounds corresponding to the same note taken in any octave and on any instrument. To compose the concept of the perception group, it is sufficient to get acquainted with an insignificant number of its representatives [4].

II. THE PROBLEMFORMULATION

The specificity of the task that solves the object recognition system determines a number of requirements for the hardware platform. The platform should be a computer with sufficient processing power for fast image processing, as well as the available ability to connect additional devices, such as a video camera, analog and digital sensors. The platform must also have enough permanent memory to store all the necessary data. In addition, the ability to create a single network for integration with the Industrial Internet of Things, including data transmission, to gain access to the Internet to communicate with the server or use alternative channels for transferring information to a central computer node. If there are several objects of different categories on the image, each of them is recognized [5].

III. INDUSTRIAL INTERNET OF THINGS(IIoT)

Industrial Internet of Things is a system of combined computer networks and connected physical objects (things) with built-in sensors and software for data collection and exchange, with the possibility of remote monitoring and control in an automated mode, without human intervention.

Industrial IoT systems are aimed at minimal user intervention in the management and, accordingly, in independent management of technical processes with computer equipment.

There are features of the industrial Internet of things:

- all devices are controlled automatically from a single command center;
- high-frequency support of user actions;
- automated process of monitoring and managing the life cycle of equipment.

The introduction of computer vision systems into the industrial Internet of things helps to ensure control over compliance with safety requirements at work through automatic control and data analysis.

Based on the above, the task of object recognition is also relevant in the implementation of relevant developments in the IIoT system [6].

IV. IMAGERECOGNIZINGMETHODS

There are various algorithms that allow you to recognize images. The algorithm for learning the machine for pattern recognition, based on the method of secant hyperplanes, consists in approximating the parts of hyperplanes separating the hypersurface and consists of the following main stages:

- training (formation of a separating surface);
- secant planes;
- elimination of excess planes;
- elimination of excess parts of planes;
- recognition of new objects.

When using the method of parallel variants, several devices are trained simultaneously and independently of one another on the same material. When recognizing new objects, each machine will refer these objects to some image, maybe not to the same image. The final decision is made by "voting" machines, the object refers to the image to which it was attributed by a greater number of machines. The way to increase the reliability of recognition is to some improvement in the method of secant planes. It can be assumed that if the secant planes are drawn close to the plane passing through the middle of the straight line connecting the object and the opponent perpendicular to this line, the resultant surface will be closer to the true boundary between the images.

V. OBJECT RECOGNITION STAGES

The task of recognition of objects (images) is determined by the following steps:

- the definition of boundaries is the lowest-level task for which the neural networks are already classically applied;
- the definition of the vector to the normal allows us to reconstruct a three-dimensional image from a two-dimensional image;
- the definition of objects of attention (saliency) is what the person would pay maximum attention to when analyzing the picture;
- semantic segmentation allows you to divide objects into classes according to their structure, not knowing anything about these objects, that is, even before they are recognized;
- the semantic delineation of boundaries is the allocation of boundaries divided into classes;
- the highest-level task is the recognition of objects themselves [7].

VI. YOLO V3 AND V2 CHARACTERISTICS

Yolo v3 (and v2) is a program neural network packet aimed at object recognition. With its help, it is possible to localize and identify an object [8]. All previous detection systems use classifiers or localizers to perform tasks. They apply the model to the image in several places and scales. Areas with high detection density are considered to be detected [9].

Yolo uses a different approach. A single neural network is applied to the full image. This network divides the image into regions and predicts the limiting fields and probabilities for each region. These limiting fields are weighted by the projected probabilities.

The abbreviation YOLO stands for You Only Look Once. This model imposes a grid on the image, dividing it into cells. Each cell attempts to predict the coordinates of the detection zone with a confidence estimate for these fields and the probability of the classes. Then, the confidence estimate for each detection zone is multiplied by the probability of the class to obtain a final rating [10].

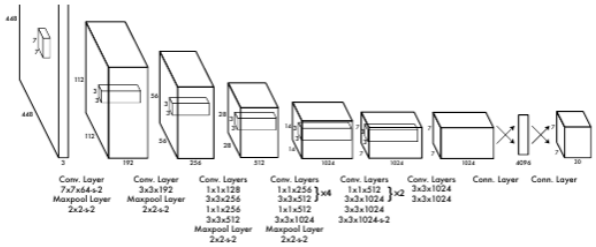


Fig. 1. The DarkNet YOLO v3 neural network architecture

Images that need to be processed can be obtained from different sources: a USB camera or .jpeg files placed in the / darknet / data directory. It is also possible to process video files.

The Yolo model was developed for a neural network based on DarkNet. DarkNet stores the learned coefficients (weights) in a format that can be recognized by different methods on different platforms. DarkNet is written in C and has no other programming interface, so if the platform requirements or your own preferences require you to access another programming language, you will have to work on its integration in addition. Also, it is distributed only in the source code format, and the compilation process on some platforms can be very problematic [11].

The neural network consists of 24 convolution layers, followed by 2 completely connected layers. The rotation of the 1 × 1 convolution layers separates the load on the previous layers. The resolution of the convolution layers in the problem of classifying images is halved (input image 224 × 224), but then the resolution for detection is doubled again [12].

In general, the task of object recognition consists of two parts: learning and recognition. Training is carried out by showing individual objects with indication of their belonging to one or another image. As a result of training, the recognition system should acquire the ability to react with identical reactions to all objects of the same image and other reactions to all objects of distinct images. It is very important that the learning process should be completed only by showing a finite number of objects. Learning objects can be either pictures or other visual images (letters, numbers). It is important that in the learning process only the objects and their belonging to the image are indicated. The training follows the process of recognition of new objects, which characterizes the actions of the already trained system. Automation of these procedures is the problem of learning to recognize patterns. In the case when a person himself unravels or thinks up, and then imposes a classification rule on the machine, the problem of recognition is partially solved, since the main and the main part of the task (training) is taken by the individual.

The input image falls into a network of layers, called filters of different sizes and different complexity of the elements that they recognize. These filters constitute a random index or a set

of characteristics, which then falls into the classifier. Usually it is a multilayer perceptron.

In the image and likeness with a biological neural network, objects of varying complexity are recognized.

VII. PROGRAM TESTING

Configuration of the personal computer (1):
CPU: Intel® Core ™ i7-2670QM CPU @ 2.20 GHz
Memory (RAM) 6.00 GB
Windows 10 Pro x64, Ubuntu 16.04 x64
Graphicscard: ATI Radeon HD6770M
Harddrive: 1.00TB

You need to install the OpenCV package. Installation of CUDA, cuDNN is not required, because on this device there is no video adapter of the corresponding manufacturer (NVidia) [13].

Install the package Darknet Yolo:

We put the file yolov3.weights in the directory darknet;

open the Makefile, set OpenCV and OpenMP = 1.

Next, we execute the assembly commands without leaving the directory.

Run the command for recognition from the web-camera:

./darknet detector demo cfg / coco.data cfg / yolov3.cfg yolov3.weights

The image will be received via USB-web-camera.

It is noticeable that in the terminal the frame processing frequency is counted, 0.2 FPS; it means that one frame leaves 1 / 0.2 = 5 seconds.

Also, the program calculates the percentage of authenticity of object recognition.

Let's estimate the speed of the system: put the image file WP_20160805_009.png in the directory / darknet / data.

Image processing results WP_20160805_009.png:

processing time = 4.85 s.

Repeat the same with another image (WP_20160819_001) .png:

image processing results WP_20160819_001.png:

processing time = 4.94 s.

It turns out that, on average, the image processing takes about 5 seconds.

VIII. PROGRAM TESTING (SECOND EQUIPMENT OPTION)

Let's take a computer of a different configuration (2):

CPU: AMD Athlon 2 x 3 440 @ 3 GHz

Memory (RAM) 4.00 GB

Windows 10 Pro x64, Ubuntu 16.04 x64

Graphics card: NVidia GeForce GTX650Ti 1 GB,

Harddrive: 380 GB

In the Makefile, set the GPU and CUDNN = 1.

For technical reasons, we launch not YOLO v3, but YOLO v2 (the terminal displays a message that there is not enough video memory for the YOLO v3 on the graphics processor, apparently, a higher-performance graphics card is required).

Fig. 2. WP_20160819_001.jpg image processing

We start the recognition from the camera. Immediately note the frequency of 8.9 frames per second. This is 45 times higher than the previous equipment (0.2 frames per second).

We will recognize the same images:

- image WP20160805_009.jpg is detected in 0.096 seconds;
- image of WP20160819_001.jpg was detected in 0.098 seconds.

On average, the processing of one image takes 0.1 s, which is 50 times faster than in the previous configuration (5 s). These examples do not show a backlog in the classification of objects by different versions of the program.

IX. IIoT INTEGRATION

It is considered permissible integration of the software and hardware complex into the IIoT service, namely, the delivery of recognition results to where other services can access them.

To date, a widely distributed solution is python daemon.py, which will run a simple server that displays a video stream from a webcam with forecasts for: http://127.0.0.1:8000/events/.

The task is relevant for monitoring the actions of personnel in production, as well as for monitoring the position of materials on the conveyor.

For the security of the system, data is encrypted. Different solutions are being developed, one of them is the automatic generation of the key by the client device.

Fig. 3. WP_20160805_009.jpg image processing

X. CONCLUSION

The system Yolo v3 and Yolo v2 has been tested. The architecture and process of image processing are considered. The system recognizes objects and defines their names. The neural network is tested on its own equipment.

It is noted that the use of NVidia GTX650Ti for processing graphics allows to process video stream from webcam 45 times faster, and images 50 times faster than on Intel Core i7 2.20 GHz second generation. The system operation is expedient when using the NVidia graphics processor with CUDA cores.

All detected objects are recognized, and the accuracy of the detection is calculated. The algorithms of working with the network are shown.

TABLE 1. PERFORMANCE COMPARISON ON DIFFERENT EQUIPMENT

№	Operation name	Run time (frame rate)		Performance factor (2 relative to 1)
		1	2	
1	Web-camera recognition	0.2 frames/s	8.9 frames/s	44.5 ≈ 45
2	WP_20160819_001 image recognition	4.94 s	0.098 s	50
3	WP_20160805_009 image recognition	4.85 s	0.096 s	50

XI. REFERENCES

[1] Воронов В.И., Усачев В.А. Компетенция «машинное обучение и большие данные» // Приоритетные направления развития науки и образования / Под общ. ред. Г.Ю. Гуляева. Пенза, 2017. С. 97–108. [Usachev V.A., Voronov V.I. The competence “Machine learning and Big Data” // Priorities for development of science and education / Ed. G.Yu. Gulayev. Pensa, 2017. P. 97–108].

[2] Михаеску С.В., Трунов А.С., Воронова Л.И. Анализ предметной области для разработки системы построения скелетной модели человека на основе массива опорных точек, получаемых совокупностью контроллеров Kinect // Международный студенческий научный вестник. 2015. № 3-4. С. 521–522. [Mihaesku S.V., Trunov A.S., Voronova L.I. Analysis of the subject area of system developing for building a human skeletal model based on an array of reference points obtained from a set of Kinect controllers // International Student Scientific Bulletin, 2015. N 3-4. P. 521–522].

[3] Воронова Л.И., Воронов В.И. Machine Learning: Регрессионные методы интеллектуального анализа данных: учеб. пос. / МТУСИ. М., 2017. [Voronova L. I., Voronov V. I.. Machine Learning: Regression methods data mining: A tutorial / MTUCI. Moscow, 2017].

[4] Machine Learning. Seminars on neural networks. URL: http://www.machinelearning.ru/wiki/images/1/1e/Sem07_ann.pdf.

[5] Молодяков С.А., Тышкевич А.И. Принципы выделения параллельных потоков команд обработки видеоизображенгий в smart-видеокамерах // Международный научный журнал. 2016. № 9. С. 76–80. [Molodyakov S.A., Tishkevich A. I. Principles of allocation of parallel streams of commands of processing of video images in smart-video cameras // International science journal. 2016. N 9. P. 76–80].

[6] Consumer Home // Digital Living Network Alliance. URL: http://www.dlna.org.

[7] Reference Model for Service Oriented Architecture // Organization for the Advancement of Structured Information Standards. URL: http://docs.oasis-open.org/soa-rm/v1.0/soa-rm.pdf.

[8] Redmon J., Divvala S., Girshick R. et al. You Only Look Once: Unified, Real-Time Object Detection // The conputer vision Foundation. URL: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf.

[9] Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. 2-е изд., стер. М.: Горячая линия — Телеком, 2002.[Kruglov V. V., Borisov V. V. Artificial neural networks. Theory and practice. 2nd ed., ster. Moscow: Goryachaya Liniya — Telecom, 2002].

[10] Agrawal P., Girshick R., Malik J. Analyzing the performance of multilayer neural networks for object recognition // ECCV. 2014. P. 329–344.

[11] Bengio Y., Courville A.C., Vincent P. Unsupervised feature learning and deep learning: A review and new perspectives // CoRR. 2012. Vol. 1. abs/1206.5538.

[12] Redmon J., Angelova A.. Real-time grasp detection usingconvolutional neural networks // CoRR. 2014. abs/1412.3128.

[13] Буянов Б.Я., Верба В.А. Использование модуля dnn библиотеки OpenCV 3.3 для распознавания объектов // Инновационные технологии в кинематографе и образовании: Сб. IV Междунар. науч.-практ. конф. / ВГИК. М., 2017. С. 47–56. [Buyanov B.Ya., Verba V.A. Using the dnn module of the OpenCV 3.3 library for object recognition // Innovative technologies in cinema and education: Coll. IV International scientific-practical conf. / VGIK. Moscow, 2017. P. 47–56].

Design of a Program Complex Architecture for Equipment Control Using Gestures

Mikhail D. Artemov[1], Lilia I. Voronova[2]
Department of Intelligent Control Systems and Automation
Moscow Technical University of Communication and Informatics
Moscow, Russia
[1]artemov_mikle@mail.ru, [2]voronova.lilia@yandex.ru

***Abstract*. Industrial Internet of Things technologies is widely distributed. Smart cities, factories, shops, house assume using intelligent methods of data processing providing by a sensors. In particular, smart enterprise control is possible by through of gestures and is interesting and perspective task. In the article briefly reviewed methodology of gesture and action recognition by image sequences and proposed neural network architecture for gesture recognition and smart enterprise control.**

***Key words:* gesture recognition, Mask R-CNN, CNN, 3D CNN, LSTM.**

I. INTRODUCTION

Industrial Internet of Things (Internet of Things for industrial/corporate use, IIot) is system of united computer networks and connected industrial facilities having an integral sensors and a software for data collection and exchange, which allow remote control and management in automated mode without a human control.

IIoT has the following dignities:

- Paper documentation is can being rejected.
- Specific knowledge of a specialists can be collected.
- Unscheduled delay, maintenance and equipment breaking are being prevented and fail of delivery sequence can be delayed. Consequently, efficiency of enterprise work can been improved.
- Production stages can be automated and optimized.
- Introduction of network interaction between a machine, an equipment, knowledge and information systems allow effective monitoring and analysis of environment, production process and real-time self-state.
- Risk of emergency can be decreased due to prevent of malfunction and decreasing risk of manufacture delays.

There are the following principles: firstly, sensors, actuating mechanisms, controllers and man-machine interfaces are installed to important equipment parts. Secondly, information is collected and allow to get objective and accurate data. Handled data move to all the enterprise departments and allow to accept a substantiate decisions and to establish interaction between a different departments employees.

For implement of this approach, all necessary information about real resources state within the enterprise and within other enterprises is to be made available to automated control systems of different levels.

Introduction of IoT requires to change approaches of create and use of automated information control systems and common approaches of enterprises control [1]. Obsolete production lines, which cannot be automated using IoT will be changed to new automate and a robotic equipment in future.

For traditional enterprises and them systems, a staff is the base resource, which is necessary for other kinds of resource control. Consequently, voice information exchange and data between employees is main way of information exchange in such systems.

Grow and development of traditional enterprises is related to using IIoT. Adequate interpretation and filtration is high priority task for enterprises, when they are processing huge quantity of unstructured data, and correct information presentation understandable for user is a task acquiring special mean. For those tasks, a modern market is offering an advanced analytical platforms designed to combine, store a analyze data about technical processes and events in real time.

We offer to unify man-machine interface for interaction with equipment; this will allow to simplify handle of huge quantity of data using single information representation and processing.

Unification is meant a control using gestures. Gest is some action or motion of a human body or a him part, which has special meaning, i.e. is sign or symbol [2]. Reject of traditional mechanisms in favor of neural network and a single video camera explicit advantage such control [2; 3].

The video used for control is valuable data source for IoT. So, for example, using machine learning technology will allow to automate processes of algorithms improving, which are being executed by programming to “management cloud”, i.e. this allows to optimize management algorithms, when historical data are being combed, and, consequently, to increase a management efficiency. Collected data can train neural networks to action algorithms, which an operator executes, and in case of unusual system behavior can to notify about that. Moreover, trained system will give remote control opportunity or can independently to launch studied action algorithms [4; 5].

The single interface will allow to unify support such system and to run control of operators efficiency work on a different workstations.

Such system could to provide security of different levels by changing authentication politics and using artificial intelligence, when it analyzes a video frame. Moreover, such interface will be needed single knowledge and skills in the field of cybersecurity, this will allow to simplify them in future. Moreover, single system will ensure data compatibility and give opportunity of secure automated system of patch installing.

New technologies allow enterprises of different industries to achieve significant competitive advantages and take steps to meet recommendations for develop and exploit of IoT.

Such decisions transform production data to useful information, which is required for security and rational enterprise management.

II. RELATED WORK

Existing gesture recognition networks is can categorized by a principle of work with time dimension into different three groups, as proposed in [6]:

- Networks are using 3D filters in the convolutional layer (Fig. 1, *a*). The 3D convolution and 3D pooling in CNN layers allow to capture discriminative features along both spatial and temporal dimensions while maintaining a certain temporal structure.
- Networks are using motion features like 2D dense optical flow maps, which can be precomputed and input to the networks (Fig 1, *b*). Extracted motion features can be fed to the network as additional channels to the appearance ones or input to a secondary network (later combined with
- Networks are using combines a 2D (or 3D) CNN applied at individual (or stacks of) frames with a temporal sequence modeling (Fig. 1, *c*). Recurrent Neural Network (RNN) is one of the most used networks for this task, which can take into account the temporal data using recurrent connections in hidden layers.

3D filters in the convolutional layer are used [7; 8]. A space-temporal feature learning approach is offered [7], which use a deep 3D convolutional networks, which are trained on large number controlled video datasets. Their findings are three-fold:

- they are more suitable for spatiotemporal feature learning compared to 2D ConvNets;
- a homogeneous architecture with small convolution kernels in all layers is among the best performing architectures for 3D ConvNets;
- a simple linear classifier outperforms state-of-the-art methods and achieving 52.8% accuracy on UCF101 dataset with only 10 dimensions and is also very efficient to compute due to the fast inference of ConvNets.

This network conceptually very simple and easy to train and use.

It is proposed hand gesture recognition system that interleaves depth and intensity channels to build normalized spatio-temporal volumes, and train two separate subnetworks with these volumes [8]. The VIVA challenge's hand gesture dataset contains 885 intensity and depth video sequences of 19 different dynamic hand gestures performed by 8 subjects inside a vehicle. Both channels were recorded with the Microsoft Kinect device and have a resolution of 115×250 pixels.

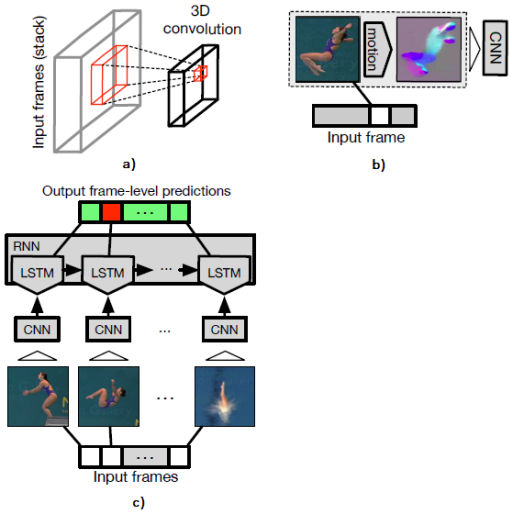


Fig 1. Groups of neural networks:
a) Networks are using 3D filters in the convolutional layer; b) Networks are using motion features like 2D dense optical flow maps; c) Networks are using combines a 2D (or 3D) CNN applied at individual (or stacks of) frames with a temporal sequence modeling [6]

An approach [9] is related to the second group propose semi-supervised approach using a deep neural network, by combining an autoencoder with a classification loss function, and training both of them in parallel. Other work [10] is related to same group propose three simple, compact yet effective representations of depth sequences, referred to respectively as Dynamic Depth Images (DDI), Dynamic Depth Normal Images (DDNI) and Dynamic Depth Motion Normal Images (DDMNI).

Other approach [11] is related to third group, which use a very large, annotated video dataset of the dynamic hand gestures and neural networks are trained on this data. This solution require a single camera and worked on various of platform. To train system, used a large dataset of short, densely labeled video clips that was crowd-acted by our community of crowd workers. The dataset contains ~150,000 videos across 25 different classes of human hand gestures, split in the ratio of 8:1:1 for train/dev/test. It also includes two "no gesture" classes to help the network distinguish between specific gestures and unknown hand movements.

Long Short-Term Memory Recurrent Neural Network (CNNLSTM) is proposed for the problem of dynamic gesture recognition[12]. That model consists of two consecutive convolutional layers, a flattening layer, a Long Short-Term Memory recurrent layer and a softmax output layer.

III. MASKCNNLSTM

In our project, we use approach based on using of a 3D CNN, which is individually applied to set of frames, which are modeling a temporal sequence.

Gesture control requires user's gesture recognition in real-rime. Therefore, the system should divide the video sequence into frames at a speed of 15–20 frames per second.

Task target neural network is human gestures recognition. Therefore, on introductory stage, we should detach a human and

his hands. This problem is solved by segmentation using by a Mask R-CNN (Fig. 2).

Apart from a good results of instance segmentation and object detection,Mask R-CNN is suitable for determining human pose estimation in photographs. For this, the keypoints select is important. That is a left shoulder, a right elbow, a right knee, etc. Using such points allows you to draw human pose skeleton. A neural network is being trained for select of the keypoints, and after that it should receive masks, in which just a one pixel have 1, and all other have 0. At the same time, the network is being trained to output the several single-pixel masks one at each keypoint [13].

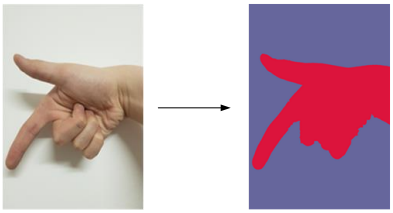


Fig. 2. The input image and the Mask R-CNN result

Use of segmentation will allow to simplify training of neural network, which is being making gesture recognition. Moreover, this lets to increase a classification accuracy and to detail human pose estimation with the aim of improve gesture definition, when key points are being added.

For recognition we propose to use MaskCNNLSTM architecture, which shows in Fig. 3.

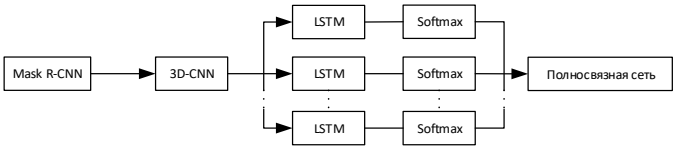


Fig. 3. Proposed gesture recognition architecture (MaskCNNLSTM)

After segmentation the frames sequence, on first stage, input to the 3D-CNN. That network handles the frames as seen in Fig. 4.

1) An input is being got sequence of the L frames.

2) Convolutional layers have filters of size [*k*,*k*,*d*], there *k* is dimension of reception field; *d* is frames count, which will be handled by filter. In our case, *d* = *L*, i.e the filter have [*k*,*k*,*L*] size. Our network will have the small height and weight of a convolution kernel (for example, 2×2 or 3×3) with the same filter size on all convolutional layers, since this will allow to receive best performing architecture [7].

3) In the output of the network we get a 3D feature map of size [*f*,*h*,*n*], which can be considered as sequence of *n* map with size [*w*,*h*].

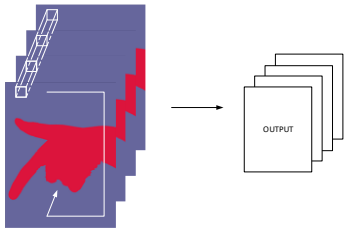


Fig. 4. Handling of the image sequence using the 3D-CNN

Features, which got from the 3D-CNN and related with action, can exist only at special position or time periods. Consequently, it is necessary to introduce a focusing mechanism for improve classification efficiency. Therefore, the feature maps received from the 3D-CNN are entering to inputs of LSTM networks, which can naturally introduce a temporal memory using previously received context.

The LSTM has been successfully combined with 2D CNN to incorporate visual attention and locate region of interest in the video sequence.

Usually, the LSTM incorporates all the input as a vector, which could not preserve the spatial correlation within an image. For solution of that problem, we will use a model [14]. The model is based on the LSTM, which combines the temporal and spatial attention mechanisms into a single scheme.

The LSTM model is being described by following equations:

$$i_t = \sigma\left(x_t W_{xi} + h_{t-1} W_{hi} + c_{t-1} W_{ci} + b_i\right) \quad (1)$$

$$f_t = \sigma\left(x_t W_{xf} + h_{t-1} W_{hf} + c_{t-1} W_{cf} + b_f\right) \quad (2)$$

$$o_t = \sigma\left(x_t W_{xo} + h_{t-1} W_{ho} + c_{t-1} W_{co} + b_o\right) \quad (3)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tanh\left(x_t W_{xc} + h_{t-1} W_{hc} + b_c\right) \quad (4)$$

$$h_t = o_t \circ \tanh\left(c_t\right), \quad (5)$$

where x_t is the input to the LSTM block; i_t, f_t, o_t, c_t, h_t are the input gate, the forget gate, the output gate, the cell state and the output of the LSTM block respectively at the current time step *t*. *L* is the weights between the input layer and the input gate, the forget gate and the output gate respectively. W_{hf}, W_{hi}, W_{ho} are the weights between the hidden recurrent layer and the forget gate, the input gate and the output gate of the memory block respectively. W_{ci}, W_{cf}, W_{co} are the weights between the cell state and the input gate, the forget gate and the output gate respectively and finally; b_i, b_f, b_o are the additive biases of the input gate, the forget gate and the output gate respectively. The set of activation functions consists of the sigmoid function $\sigma()$, the element-wise multiplication $\circ()$ and the hyperbolic activation function $\tanh()$.

Feature map was weighted by a spatial attention weight matrix, this allows to safe spatial attention.

$$\alpha_{t,i} = \frac{\exp\left(W_i e_{t,i}\right)}{\sum_{j=1}^{w \times h} \exp\left(W_i e_{t,i}\right)}, \quad (6)$$

where $X_{t,i}$ is feature map [*m*,*h*] received from 3D-CNN; *i* is cell of $X_{t,i}$; W_i is the weight mapping to element *i* of the spatial weight matrix $X_{t,i}$; $e_{t,i}$ is multilayer perceptron conditioned on the current input $X_{t,i}$ and the previous hidden state h_{t-1}:

$$e_{t,i} = \tanh\left(W_{\alpha} X_{t,i} + W_{\alpha h} h_{t-1} + b_i\right). \quad (7)$$

The $\alpha_{t,i}$ evaluates the importance of the i region to the frame per point t.

$$x_t = \sum_{i=1}^{w \times h} \alpha_{t,i} X_{t,i} \ . \quad (8)$$

Thus, at each time step, the LSTM will predict the weight matrix for the next time step, and output of the LSTM will includes two parts, which are the activity labels and the weight matrix.

Every a softmax layer makes a predict for each the LSTM output. Finally, a fully connected network determines class of the source videosequence.

IV. CONCLUSION

Recently, interest in recognizing actions and gestures has much increased. In the this work, we briefly review a methodology of gesture recognition by an image sequences and propose a neural network architecture MaskCNNLSTM, which include a Mask R-CNN, a 3D CNN and a LSTM networks, which will be to used for gesture recognition.

A proposed system will could to provide security of different levels by changing authentication politics and using artificial intelligence, when analyze a video frame. Moreover, such interface will be requiring single knowledge and skills in the field of cybersecurity, this will allow to simplify them in future. Moreover, single system will ensure data compatibility and give opportunity of secure automated system of patch installing.

V. REFERENCES

[1] Воронов В.И., Усачев В.А. Компетенция «машинное обучение и большие данные» // Приоритетные направления развития науки и образования. Под общ. ред Г.Ю. Гуляева. Пенза, 2017. С. 97–108. [Usachev V.A., Voronov V.I. The competence “Machine learning and Big Data” // Priorities for development of science and education / Ed. G.Yu. Gulayev. Pensa, 2017. P. 97–108].

[1] Михаеску С.В., Трунов А.С., Воронова Л.И. Анализ предметной области для разработки системы построения скелетной модели человека на основе массива опорных точек, получаемых совокупностью контроллеров Kinect // Международный студенческий научный вестник. 2015. № 3-4. С. 521–522. [Mihaesku S.V., Trunov A.S., Voronova L.I. Analysis of the subject area of system developing for building a human skeletal model based on an array of reference points obtained from a set of Kinect controllers // International Student Scientific Bulletin, 2015. N 3-4. P. 521–522].

[2] Воронов В.И., Воронова Л.И., Генчель К.В. Применение параллельных алгоритмов в нейронной сети для распознавания жестового языка // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2018). VII Международная научно-техническая и научно-методическая конференция: Сб. науч. ст.: В 4 т. / Под ред. С.В. Бачевского. М., 2018. С. 207–212. [Voronov V.I., Voronova L.I., Genchel K.V. The use of parallel algorithms in the neural network to recognize the sign language // Actual problems of information and telecommunications in science and education (APINO 2018). VII International scientific tech. and scientific method. Conf.: In 4 t. / Ed. S.V. Bachevsky. M., 2018. P. 207–212].

[3] Горячев Д.В., Воронов В.И. Большие данные и машинное обучение // Технологии информационного общества: Матер. XII Междунар. отраслевой науч.-техн конф. М.,2018. С. 327–328. [Goryachev D.V. Big Data and Machine Learning // Information Society Technologies: Proceedings of 12th the international scientific-practical conference. Moscow, 2018. P. 327–328].

[4] Voronov V.I., Voronova L.I. Features of realization master’s program “Automation of technological processes and manufactures” // International Journal of Applied and Fundamental Research. 2016. № 2. URL: www.science-sd.com/464-25196.

[5] Asadi-Aghbolaghi M., Clapes A., Bellantonio M. et al. A Survey on Deep Learning Based Approaches for Action and Gesture Recognition in Image Sequences // 12th IEEE Conference on Automatic Face and Gesture Recognition. 2017. P. 476–483.

[6] Tran D., Bourdev L., Fergus R. et al. Learning Spatiotemporal Features with 3D Convolutional Networks // IEEE International Conference on Computer Vision. 2015.

[7] Molchanov P., Gupta Sh., Kihwan Kim K. et al. Hand Gesture Recognition with 3D Convolutional Neural Networks // IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2015.

[8] Gupta O., Raviv D., RaskarR. Multi-velocity neural networks for gesture recognition in videos. URL: https://arxiv.org/abs/1603.06829.

[9] Wang P., Li W., Liu S. et al. Large-scale Isolated Gesture Recognition Using Convolutional Neural Networks // 23rd International Conference on Pattern Recognition (ICPR). Cancun, 2017.

[10] Gesture recognition using end-to-end learning from a large video database. 2017. URL: https://medium.com/twentybn/gesture-recognition-using-end-to-end-learning-from-a-large-video-database-2ecbfb4659ff.

[11] Tsironi E., Barros P., Wermter S. Gesture Recognition with a Convolutional Long Short-Term Memory Recurrent Neural Network // Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN). 2016. P. 213–218.

[12] MaskR-CNN: Architecture of modern neural network for object segmentation by image. 2018. URL: https://habrahabr.info/development/image-processing/4391-mask-r-cnn-architecture-of-a-modern-neural-network-for-segmenting-objects-on-images.html.

[13] Lu N., Wu Y., Feng L. et al. Deep Learning for Fall Detection: Three-Dimensional CNN Combined With LSTM on Video Kinematic Data // IEEE Journal of Biomedical and Health Informatics. 2019. Vol. 23, N 1. P. 314–323.

Real-Time Object Classification on Aircraft Board using Raspberry Pi 3

Nikita V. Belov[1], Boris Y. Buyanov[2], Vera A. Verba[3]
Department of Intelligent Control Systems and Automation
Moscow Technical University of Communications and Informatics
Moscow, Russia
[1]djnikone2777@yandex.ru, [2]b.buyanov@gmail.com, [3]verba@list.ru

***Abstract*. The approach to solving the problem of classification of objects in real time and its application on an unmanned aerial vehicle are investigated. When implementing this system, a Raspberry pi 3 microcomputer with a video camera connected to it, as well as an artificial neural network of the Fast Yolo architecture is used.**

***Key words:* unmanned aerial vehicle; hexacopter, artificial neural network; Fast YOLO v2; Raspberry pi 3.**

I. INTRODUCTION

This article uses INS architecture Fast YOLO V2 (You Only Look Once), which is a faster development of the YOLO v2 architecture, and allows you to detect objects on the video stream on embedded devices in real time. Firstly, it is used the evolutionary structure of deep intelligence for the of YOLO v2 network architecture and create an optimized architecture at that has 2.8 times less parameters. To further reduce power consumption on embedded devices while maintaining performance, the adaptive movement method is introduced into the proposed fast YOLO fast structure to reduce the frequency of the outputs using Fast YOLO v2 based on the time characteristics of the motion. Fast YOLO architecture reduces the number of pins by an average of 38.13% to detect objects in the video stream compared to the original YOLO v2.

II. DETECTION OF OBJECTS AND THEIR LOCALIZATION

In the industrial Internet of things, a hexacopter with an installed object recognition system can be used for visual inspection of ships and determining breakages, tracking the spread of weeds, analyzing the number of people at enterprises and so on. By transmitting a video signal and communicating with any device with a wi-fi network on board, it can act as a sensor and help in the automation of production.

Detection of objects [1; 5] is one of the most difficult tasks in the field of technical computer vision. The purpose of object detection is to localize various objects in the scene and assign labels to object restrictions. The most common approach [1; 2] to solving this problem is to reuse existing trained classifiers to assign labels to the bounding rectangles in the scene. For example, you can use the standard slip approach [1], where the classifier defines the existence of an object and the associated labels for all possible windows in the scene. However, this approach has significant limitations in terms of not only high computational complexity, but also a high rate of detection errors.

Deep neural networks (DNN) have shown excellent performance in various applications [3; 4], and object discovery is one of the key areas where DNN has far exceeded all previously known approaches. In particular, convolutional neural networks (CNN) have demonstrated the best characteristics when solving the object detection problem. For example, the CNN approach uses an architecture to create bounding rectangle suggestions in an image instead of a moving window approach, and thus, the classifier divides into classes by bounding rectangles. Although the modernized architecture of R-CNN (Region-based Convolutional Network) is capable to obtain maximum accuracy, the entire procedure is slow and cannot be optimized, since each component must be trained individually.

In 2016, the “Detecting an object in one pass (YOLO)” [6] approach was proposed, which mitigated the complexity of computations associated with R-CNN, posing the problem of object detection as a problem with a single regression, where the coordinates of the bounding framework and the probability of classes are computed parallel. Although it was demonstrated that YOLO has significant speed advantages over R-CNN (for example, 45 frames per second on the Nvidia Titan-X graphics processor), it has also been shown that the localization error of YOLO is significantly higher than in architectures such as Faster R-CNN [7].

J. Redmon and A. Farhadi [8] proposed an improved YOLO method (called YOLO v2), where the so-called “anchor boxes” are used to predict bounding rectangles. In addition, compared to YOLO, YOLO v2 does not have fully connected layers in its network architecture. To make the network faster, YOLO v2 uses the new architecture of CNN (Darknet-19), unlike other frameworks that use VGG-16. Darknet-19 requires 8.52 billion floating-point operations, which is significantly lower than VGG-16, for which requires 30.69 billion floating-point operations on each pass. Experimental results for this approach have shown that YOLO v2 can perform object detection with 67 FPS on the Nvidia Titan-X graphics processor, achieving the fastest detection characteristics.

III. PRINCIPLE OF WORK YOLO

The principle of YOLO is the following: at each step you run the classifier to get a prediction of what sort of object is inside the current window. Using a sliding window gives several hundred or thousand predictions for that image, but we will keep only the ones, in which the probability of finding an object is more than 30%.

This approach shows good results, but obviously, it will be very slow, since it requires multiple launch of the classifier. A more effective approach is to predict which parts of the image contain interesting information (the so-called regional proposals), and then we run the classifier only in these areas.

YOLO V2 uses a different approach. This is not a traditional classifier, which is an object detector, a YOLO V2 divides the image into cells 13x13 pixels, as shown in Fig. 1.

Fig. 1. Separation of the image into separate cells [10]

Each of these cells is responsible for the prediction of 5 bounding rectangles. The bounding box describes the rectangle that surrounds the object.

YOLO also calculates a confidence score that shows how reliably it is determined that the predicted bounding box actually covers an object.

The predicted bounding rectangles look as shown in Fig. 2 (the higher the confidence indicator, the greater the thickness of the lines that limit the field).

Fig. 2. Example of forecasting [10]

For each bounding block of cells, the class also predicts. This works just like a classifier: it gives the probability distribution over all possible classes. In the version of YOLO, training was conducted on a PASCAL VOC data set [9], which can detect 20 different classes, such as bicycles, boats, cars, cats, dogs, people, etc.

The confidence estimate for the bounding framework and class prediction is combined into one final score, which indicates the probability that this bounding box contains a certain type of object. For example, a big fat yellow line on the left indicates a dog with a probability of 85% (Fig. 3).

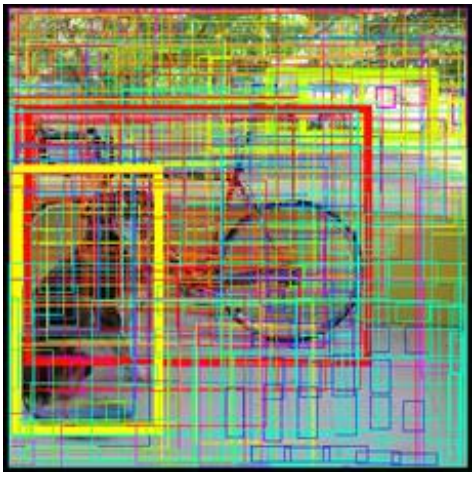


Fig. 3. Selecting objects in the picture [10]

Since there are 13 × 13 = 169 grid cells, and each cell predicts 5 bounding rectangles, it eventually results in 845 bounding rectangles. Most of these cells will have very low confidence ratings, so only those rectangles are saved, for which the final result is at least 30%.

In our case, the final prediction is shown in Fig. 4.

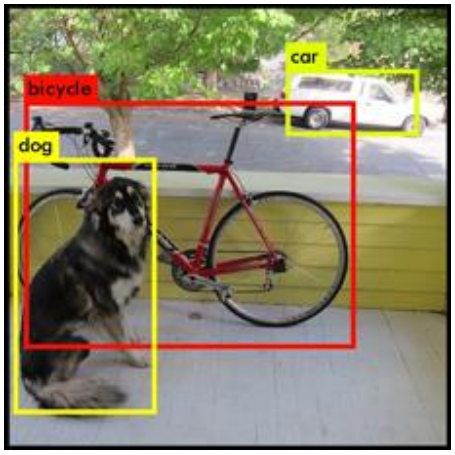


Fig. 4. Final prediction [10]

Of the 845 common restrictive boxes, only these three are preserved, because they give the best results. But even if there were 845 individual predictions, they were all done simultaneously - the neural network worked in one pass.

IV. HEXACOPTER CONSTRUCTION

To test Fast YOLO V2, a UAV was chosen onboard, a hexacopter is capable to obtain high-resolition aerial photographs and stabilized high-definition video.

The following components were used to implement the hexacopter:

- DJI F550 ARF kit;
- 3300mAh 3S 35C LiPo battery with T-connectors;

- DJI DT7 and DR16 transmitter and receiver;
- Eken H9 camera with microSD card;
- FeiYu Tech Mini 3D 3-axis gimbal;
- D-Link DWA-160;
- Raspberry pi 3.

Before the beginning of the project it is necessary to determine the purpose of each part of the hexacopter. Below is the specification of hexacopter elements:

- Frame. The frame provides a convenient way to fasten engines and electronics. DJI Flame Wheel 550 (F550) is used in this project:.
- Brushless motors and propellers. The engines rotate the propellers to create traction for lifting the hexakopter. Engines DJI E305 960KV and propellers DJI 9450 are in this project.
- Electric speed controllers (ESC). ESCs nourish brushless motors and provide a PWM interface that allows the flight controller to monitor the speed and thrust of each engine. DJI controllers E305 960KV are used in this project.
- Flight Controller. The flight controller controls how the hexacopter flies. The controller is a small computer with an inertial measuring block or IMU (which includes a gyroscope and an accelerometer) to support and stabilize a hexacopter: the flight controller also has a barometer, GPS, a magnetometer that allows the flight controller to know at what altitude it is, where he is on land and in what direction it flies, respectively. NAZA-M Lite with GPS kit is used in this project.
- Voltage regulator, flight controller status LED / USB interface. The voltage regulator provides a constant 5V for the flight controller. The status LED transmits information about the state of the hexacopter to the pilot during the flight. The indicators include a battery and a GPS. The USB interface also allows you to configure the flight controller from the PC. Used voltage regulator and status LED are included in the delivery NAZA-M Lite.
- Battery. Lithium-polymer (LiPo) batteries are the preferred method for providing the energy of a hexacopter due to the high energy-to-weight ratio and high maximum discharge currents.
- Transmitter + Receiver. The transmitter receives the pilot commands with a few joysticks and switches and transmits them over the wireless network to the receiver on a hexakopter. The receiver decodes the pilot's commands and sends them to the flight controller. DJI DT7 transmitter with DR16 receiver is used in this project.
- Gimbal + Camera. The camera and the gimbal turn the hexacopter into an aerial photography platform, capable of capturing professional quality frames. The camera is attached to the gimbal, which, in effect, performs opposite movements of the hexacopter, making sure that the camera is always horizontal. Eken H9, FeiYu Tech Mini-3D Gimbal are used in this project.
- Raspberry pi 3. The camera is connected to Raspberry Pi 3 with a 5GHz WiFi transmitter. This allows Raspberry Pi to broadcast live stream from the camera to any device connected via WiFi. The Raspberry Pi 3 has a Ubuntu Mate system, which has a wide range of applications. Model Raspberry Pi 3, D-Link DWA-160 are used in this project.

Video from Raspberry Pi 3 is sent via the Wi-Fi access point to the receiving device. There are probably three ways that you can place a Wi-Fi access point:

- The receiving device and the aircraft are connected to the same network. This approach allows you to use a powerful WiFi-router and increase the transmission range of the signal. The disadvantage is that the router is another object that must be placed on the UAV. The router also requires a separate power source.
- A Wi-Fi hotspot located on the Raspberry Pi 3, to which the receiving device is connected. This solution will work with any device, regardless of whether it can host a Wi-Fi hotspot or not.
- The Wi-Fi hotspot is located on the receiving device. Hosting the Wi-Fi access point on the device eliminates the need to move the Wi-Fi router and additional power source and makes it easy to place a WiFi access point.

In this system, we need to add signal protection functions, such as a password for accessing the wi-fi network wherein the devices are located, and also protecting the radio signal to prevent interception of device control. For this, the transmission frequency changes every 2 ms (i.e., 500 times per second), the transmitted data is mixed with a pseudo-random sequence, for more efficient transmission, and for noise immunity.

Figure 5 shows the final assembled aircraft.

Fig. 5. Collected UAVs

Fast YOLO V2 provides high performance enough to solve the task in real time on a powerful graphics processor, but it is still very difficult to use this approach to detect objects in real time in a video stream on embedded computing devices with limited processing power and limited memory. For example, in various real-world applications, such as real-time output to smart phones or embedded video surveillance, the available computing resources are limited to a combination of embedded graphics processors with low power. Therefore, the detection of objects in real time in the video stream on embedded devices remains a big problem.

After the UAV assembly, Fast Yolo v2 is configured on the Raspberry Pi 3 microcomputer.

V. IMPLEMENTING YOLO ON RASPBERRY PI 3

The Yolo models used on the Nvidia Titan-X GPU:

- yolo.cfg is based on the extraction of the network. It processes images at a rate of 45 frames per second;
- yolo-small.cfg has smaller fully-connected layers, so it uses much less memory. It processes images at a rate of 50 frames per second.

As you can see, Fast Yolo v2 produces more frames per second on the Nvidia Titan-X graphics processor, compared to Yolo v2, which is why it was decided to use it on Raspberry pi 3.

The original YOLO model uses a lot of GPU memory but because is used Raspberry Pi 3, you need to use a smaller version of the YOLO model yolo-small.cfg.

The small version of YOLO uses only 1.1 GB of GPU memory, so it is suitable for Raspberry pi 3.

For implementation, a pre-trained model was used. Coefficients of the network are available on the official website of YOLO [10]. The next task was to import the DarkNet weights in TensorFlow, for this you need to install DarkNet. After you need to run Fast Yolo v2 in test mode and because DarkNet is used on Raspberry Pi 3, 6-12 images should be processed per second.

We present the classification of objects in real time with Raspberry pi 3.

Running Fast YOLO v2 on the test data is only needed for testing the system, in a real system, the video stream from the video camera is used.

An example of how Fast YOLO v2 works on a video stream from a video camera is shown in Fig. 6.

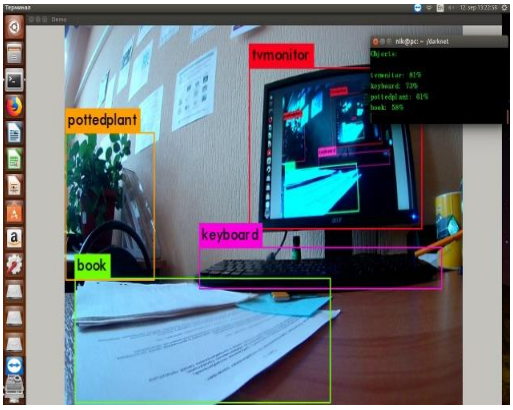


Fig. 6. Example of the work of Yolo

Fast YOLO v2 displays the current FPS and predicted classes, as well as the actual image with the bounding rectangles drawn on top of it.

As you can see in Fig. 7, the FPS (frames per second) is 9.8, which is not enough to use Fast Yolo V2 on Raspberry pi 3 in real time, because to use the FPS real-time system, it must be at least 24.

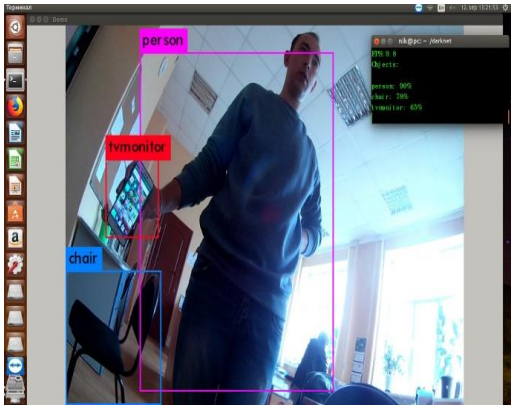


Fig. 7. Example of Yolo algorithm work

VI. CONCLUSIONS

The article discusses the INS architecture of Fast YOLO V2, this is the new INS deep learning architecture for solving real-time object detection problems on a video stream. Although YOLO v2 is considered the most modern architecture and allows you to solve the task in real time on powerful graphics processors, its direct use on embedded devices in real time is not yet possible.

In the work, the hexacopter was assembled and tuned, components were selected and a real-time video stream was transmitted, which was preprocessed on board the hexacopter with the Raspberry Pi 3 microcomputer with the preinstalled Ubuntu Mate 16.04.1 operating system.

Fast Yolo v2 on Raspberry pi 3 allows you to speed up the process of recognizing objects in real time by 33%, relative to using Yolo v2. Fast YOLO v2 can achieve an average run time of 3.3 times faster than the original YOLO v2, but this is not enough to use it in real time on Raspberry pi 3.

REFERENCES

[1] Viola P., Jones M. Rapid object detection using a boosted cascade of simple features // Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001. Kauai, 2001.

[2] Lienhart R., Maydt J. An extended set of haar-like features for rapid object detection // Proceedings. International Conference on Image Processing. Rochester, 2002.

[3] Krizhevsky A., Sutskever I., Hinton G. Imagenet classification with deep convolutional neural networks // NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems. 2012. Vol. 1. P. 1097–1105.

[4] Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition // Proc. Int. Conf. Learn. Represent. 2015. https://arxiv.org/abs/1409.1556.

[5] Girshick R., Donahue J., Darrell T. et al. Rich feature hierarchies for accurate object detection and semantic segmentation // Computer Vision and Pattern Recognition. 2014.

[6] Redmon J., Divvala S., Girshick R. et al. You only look once: Unified, real-time object detection // Computer Vision and Pattern Recognition. 2016.

[7] Ren S., He K., Girshick R. et al. Faster RCNN: Towards real-time object detection with region proposal networks // Advances in neural information processing systems. 2015.

[8] Redmon J., Farhadi A. YOLO9000: better, faster, stronger // Computer Vision and Pattern Recognition. 2017.

[9] Pascal. URL: http://host.robots.ox.ac.uk/pascal/VOC/.

[10] Yolo: Real-Time Object Detection. URL: https://pjreddie.com.

[11] Буянов Б.Я., Верба В.А. Использование модуля dnn библиотеки OpenCV 3.3 для распознавания объектов // Инновационные технологии в кинематографе и образовании: Сборник IV Междунар. науч.-практ. конф. / ВГИК. М., 2017. С. 47–56. [Buyanov B.Ya., Verba V.A. Using the dnn module of the OpenCV 3.3 library for object recognition // Innovative technologies in cinema

and education: Collection IV Intern. scientific-practical conf. / VGIK. Moscow, 2017. P. 47–56.]

[12] Буянов Б.Я., Верба В.А. Многомерный статистический и когнитивный анализ крупномасштабных систем // Вызовы глобального мира, Вестник ИМТП. 2014. № 4. С. 23–26. [Buyanov B.Ya., Verba V.A. Multivariate Statistical and Cognitive Analysis of Large-Scale Systems // Challenges of the Global World, IMTP Bulletin. 2014. N 4. P. 23–26].

[13] Буянов Б.Я., Верба В.А. Некоторые вопросы определения пространства состояний параметров сложных систем // Системный анализ в проектировании и управлении: Сб. науч. тр. по итогам XXII Междунар. науч.-практ. конф. СПб.: СПбГТУ, 2018. С. 224–229. [Buyanov B.Ya., Verba V.A. Some issues of complex systems state space parameters determining // System analysis in design and management: Sat. scientific tr. on the results of XXII International. scientific-practical conf. SPb .: SPbSTU, 2018. P. 224–229].

[14] Буянов Б.Я., Верба В.А. Мультиагентные модели сложных социо-технических систем // Системный анализ в проектировании и управлении: Сб. науч. тр. по итогам XXII Междунар. науч.-практ. конф. СПб.: СПбГТУ, 2016. С. 155–159. [Buyanov B.Ya., Verba V.A. Multi-agent models of complex socio-technical systems // System analysis in design and management: Sat. scientific tr. on the results of XXII International. scientific-practical conf. SPb.: SPbSTU, 2016. P. 155–159].

IIoT Competencies Support in the Master's Program "Automation of technological Processes and Production"

Vasilij A. Usachev[1], Vjacheslav I. Voronov [2]
Department of Intelligent Systems of Control and Automation
Moscow Technical University of Communications and Informatics
Moscow, Russia

[1]usvas.01@gmail.com, [2]vorvi@mail.ru

Abstract. **The department of ISCA MTUCI is training specialists to manage industrial Internet of things systems, based on a number of disciplines that form students' competences for working with IIoT systems. The article discusses the features of teaching one of these disciplines "Intellectual databases and data warehouses" using the freely distributed software DBMS Hadoop.**

Key words: **Industrial Internet of Things, Big Data, Hadoop.**

I. INTRODUCTION

The professors of the ISCA department developed a number of disciplines on the direction "Automation of technological processes and production" ("Intelligent automated information management systems" profile) that form students' competences for working with IIoT systems, such as "Intelligent databases and data warehouses", "Intellectual methods data processing", "machine learning and big data", etc. [1].

The purpose of mastering the discipline "Intellectual databases and data warehouses" by students is to study the methods of building knowledge bases and data warehouses and their application possibilities for intellectualizing automated processes for storing and processing information.

Tasks of mastering the discipline are: the study of the principles of building knowledge bases and data warehouses; mastering the methods of constructing queries in the NoSQL language; mastering the methods of data mining for automating the storage and processing of information [2].

The total complexity of the discipline, studied in the second semester, is 4 credit units. The discipline provides lectures, practical exercises, laboratory work, independent work, the exam.

The main sections of the discipline are: introduction to the industrial Internet of things technology, Big Data, Data Mining, methods for creating, processing and storing data in NoSQL databases and data warehouses, intellectual analysis tools and their application in databases and data warehouses, developing applications based on Apache Hadoop, ensuring information security of Big Data class systems.

II. ABOUT THE CONTENT OF LECTURE COURSE

As part of the lecture course, attention is focused on the fact that the industrial Internet of things consists of three main components (Fig. 1): a set of intelligent, network-enabled products, product systems and other "things" connected through a communication infrastructure similar to the Internet infrastructure.

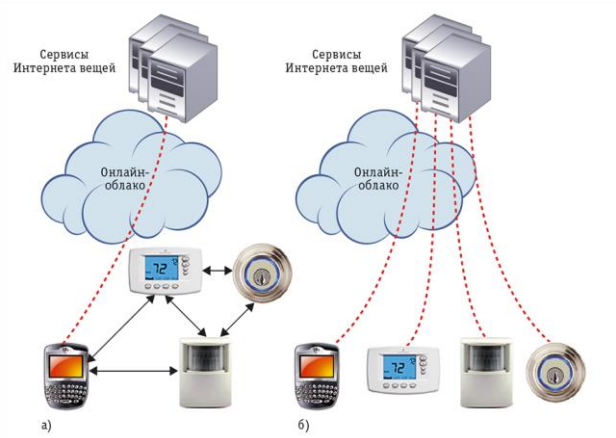


Fig. 1. Structure of the Internet of Things [3]

When developing and implementing the Internet of Things at the enterprise, special attention is paid to the critical infrastructure and the need for real-time analysis of all information coming from numerous sensors, controllers and devices [4]. Data on the state of the product, its operation and the environment are transmitted in real time to the monitoring systems in order to ensure the management, maintenance and updating of the product, as well as the efficiency of the entire system.

When analyzing solutions within the Industrial Internet of Things (IIoT), examples of implementation in the housing and utilities sector and in the mining sector are of interest.

Perhaps the most striking example of IIoT is the experience of the Moscow United Energy Company (MUEC). MUEC, together with the cellular operator MTS, is implementing a project for monitoring energy consumption [5], within which a unified automated system for controlling and recording the transfer of thermal energy and hot water has been created. For this purpose, MUEC installed 47 thousand consumption meters in municipal houses and social facilities, which continuously transmit data to the server of the MUEC Central Energy Accounting System.

The company also launched a system that allows remotely, via the Internet, to take readings of heat consumption in residential buildings in Moscow. Such meters are installed on 23,000 MUEC facilities. The system allows power engineers to quickly find out about failures in heat supply and predict possible breakdowns.

The introduction of intelligent technologies, taking into account the length of linear objects, leads to an increase in reliability and lower operating costs. This allows you to go to the management of the network "on state", and not to carry out repairs in accordance with strict regulatory deadlines.

Another example of the use of IIoT in real production is the experience of Alrosa, which manages several mining and processing plants. Alrosa and the Sum of Technologies system integrator implemented a project to create an automated information management system for the production of the Aikhalsky Minerals processing plants (Yakutia Republic) [6]. With the introduction of the system, JSC Alrosa received a single source of data on the operation of process equipment and a tool for analyzing the efficiency of production activities.

In the course of the project, automatic data collection was organized from process equipment control systems, energy accounting systems and information systems used to manage the mining and transport complex, drilling operations, underground mining, equipment maintenance planning and repair. Production information is consolidated in a unified repository, acting as a single source of technological data for production services and information systems of JSC Alrosa. Detailed data on key performance indicators are displayed on video walls located in the control rooms of the Aikhalsky Minerals processing plants and in the Administration of JSC Alrosa and transferred to the production services for analysis.

Based on the information collected, the system generates analytical reports for the operational control of the implementation of planned production indicators by the mining and processing divisions, as well as for the analysis of technical and operational indicators of mining and transport, mining and process equipment.

Regardless of the type of enterprise activity, the multitude of connected sensors and devices generates a large flow of the most diverse information that needs to be collected, processed and provided an analysis of the state of production as a whole and its components [4].

III. DBMS NoSQl (Hadoop)

To handle the flow of information resulting from the advent of the Internet of Things, it becomes more convenient to use non-familiar relational DBMS such as Oracle, Microsoft SQL Server. Now these products are being supplanted by free tools for working with unstructured data. One of them is Hadoop. The university does not always have the opportunity to work with paid software, so Hadoop is great for use in research tasks. Hadoop contains tools such as:

- Hadoop Distributed File System (HDFS) is a distributed file system that allows you to store information of almost unlimited size;
- Hadoop YARN is a framework for managing cluster resources and managing tasks, including the MapReduce framework;
- Hive is a tool for SQL-like queries on large data (turns SQL queries into a series of MapReduce tasks).

At the heart of Hadoop is the MapReduce paradigm.

The work of MapReduce is based on splitting data processing into two phases: the phase of the map and the phase of convolution reduce. Each phase uses as input and output data a key-value pair, the types of which are selected by the programmer. The map function is simple. We extract the fields we need. The map function is also well suited for eliminating unwanted entries: missing, questionable or erroneous values are filtered out here. The mapping function can run on every machine in the Hadoop cluster. Thus, the possibility of parallel processing of the initial information on several machines appears [7]. The output of the mapping function is processed by the MapReduce infrastructure before it is passed to the reduce function. During this process, key-value pairs are sorted and grouped by key. In the phase of reduce, the processing of values takes place: the calculations need to be made to solve the problem.

In the course of mastering the discipline "Intellectual databases and data warehouses" undergraduates get practical skills in solving problems in the field of data analysis, write the program code of MapReduce tasks, and gain experience in working with the job tracker and task tracker. A MapReduce job is a unit of work that a client wants to do: it consists of input data, a MapReduce program, and configuration information [8]. To perform a job, Hadoop splits it into tasks, which are divided into two types: map tasks and reduce tasks.

Practical acquaintance with big data begins with a course on the Udacity portal and the virtual machine from Cloudera [9]. In this course, students are provided with various data sets: information about purchases, information about records on the forum, logs of the web server. Undergraduates must calculate the average cost of goods sold or determine the city in which they paid the most for Visa cards, find the most active forum participant for a certain period of time or the most visited page on the site.

The system requirements of the Udacity virtual machine are not large, so its launch is possible on all modern computers. Also, when teaching students the discipline used more powerful tools. The department uses a virtual machine

image taken from the Cloudera website [10]. This image differs from Udacity in a richer toolbox. The virtual machine has all the necessary components of the Hadoop ecosystem. In this machine, a full-fledged Hadoop cluster is already deployed and configured. Users do not need to install and configure anything themselves. Also, there are already data sets for learning, and therefore students can start writing their MapReduce tasks immediately after starting the virtual machine. It is possible to upload your data to the Hadoop file system (HDFS) and process it as if they were located in a real Hadoop cluster. The volume of the processed data is not limited. The system requirements of this virtual machine are significantly larger and not every computer is suitable for this purpose. Especially to ensure the educational process, the department has a server that runs virtual machines in an amount sufficient to perform laboratory and coursework by undergraduates. Access to machines is remote.

In the work program of the direction 15.04.04 "Automation of technological processes and production" there is a laboratory course. Some of the works from the course are described in [11].

IV. ABOUT THE LABORATORY PRACTICE

Laboratory work 1. Study of the construction of queries and ways to output a relational database

The goal of the work is to learn how to perform SQL queries in the Cloudera environment with Hadoop tools.

Progress. You need to download data to the HDFS file system using Apache Sqoop. To do this, you need to open a terminal in Cloudera and execute the appropriate command.

This action may take some time. The MapReduce task is launched to export data from the MySQL database and import it into HDFS. Also, tables are created to represent HDFS files in Impala with the appropriate schema. After the import is complete, you need to check whether it has occurred correctly.

Laboratory work 2. Correlation of structured data with unstructured data

The purpose of the work is to gain the skill of working with structured data and unstructured data.

Progress. It is necessary to find the most viewed products in the online store; find out if they are the best selling

Since Hadoop can store unstructured and structured data without changing the entire database, you can also receive, store, and process the web event log. This allows you to find out what site visitors actually viewed most often.

For this we need data on web visits. The most common way to track site navigation is to use Apache Flume. Flume is a scalable tool that allows real-time tracking of routes, filtering, combining and performing "mini-operations" on the data of these paths. For convenience, some sample access data has already been preloaded in Cloudera in log files.

Laboratory work 3. Strong analysis of relationships using Spark

The objective is familiarization with the service Spark to speed up and simplify data processing.

Progress. The Spark service uses a similar concept of the 'map' and 'reduce' operations (the 'join' and 'groupBy' operations are just special variations of the 'reduce' operation). The key advantage of using Spark is that the code takes up less space and intermediate results can be stored in memory, which generally allows for iterations to be much faster.

Using MapReduce is a good option for tasks that use data that cannot fit in memory (for example, petabytes of data). This work uses Spark-on-YARN, which means that MapReduce and Spark (like many components of CDH) have a common resource manager, which makes it easier to allocate resources among a large number of users.

It is necessary to analyze the relationship with the use of Spark and to determine the products most often ordered together. A tool in CDH for quick analysis of object relationships is Apache Spark. For this work, Sparkjob is used to give an idea of the relationships of objects.

Laboratory work 4. Interactive study of event logs

The objective is index data using any of the indexing options provided by Cloudera Search.

Working process. You can choose to batch index data using the MapReduce Indexing tool or, as in our example below, expand the Apache Flume configuration, which already received web log data and event placement in Apache Solr for real-time indexing.

Web log data is a standard web server log.

Solr organizes data in the same way as a SQL database. Each record is called a document and consists of fields defined by the schema: the same as a row in a database table. The difference is that the data in Solr are usually more poorly structured. You can start indexing in real time with Cloudera Search and Flume on the web server's log data and use the Hue search user interface to explore it by creating a search index. Typically, when you deploy a new search pattern, four steps are taken:

- Create an empty configuration. In the Cloudera virtual machine, you will not need to perform steps 1 or 2, since the configuration and the schema file are already included in the cluster.
- Editing the scheme. The most common area that may be of interest is the <fields><fields /> section. From this area we can determine the fields that are present and searchable in the index.
- Download configuration. This operation may take several minutes.
- Creating a collection.

Laboratory work 5. Data Visualization

The objective is to display graphically the correlations found in previous laboratory works.

Familiarize yourself with the Hue web interface for visualizing data and creating your own dashboard.

Create a dashboard and use it to visualize the results obtained in previous works.

V. COURSE WORKS

At the end of the training discipline “Intellectual databases and data warehouses,” students performed term papers. They were given a new unfamiliar data set. Undergraduates could take the proposed data for work or find their own, but in this case they had to compile competent, interesting requests to their data. All calculations must be implemented using MapReduce [8]. The proposed input data contains the following fields: creation-date, education, employment, experience, industry, jobname, location, salary_max, salary_min, schedule. It is also worth noting that the data set includes 422434 rows with information on vacancies.

Example job assignment for a course work on working with big data:

1. Based on the “schedule” field, calculate the percentage for each “work schedule”. To do this, it is necessary to calculate the number of vacancies related to each “work schedule”, then build a pie chart. Comment on the result.
2. Based on the “location” field, count how many vacancies are related to each area. This post consists of the full address of the employer, for example: <location> Smolensk region, Dorogobuzh district, Dorogobuzh, Sovetskaya street, 1 </ location>. However, to accomplish the task, it suffices to single out the region (region, territory, republic, etc.), that is, the information located up to the “first comma”. In the above example will remain: "Smolensk region.” The distribution of the number of vacancies by region should be displayed in the table. In addition, print the total number of vacancies. Comment on the result.
3. Based on the “location” and “salary_min”, “salary_max” fields, calculate the minimum, maximum and average salary in the region. The obtained data must be displayed in the table. Comment on the result.
4. On the basis of the “location” field for the “Sverdlovsk Region”, bring out the top 10 most sought-after professions in the “jobname” field. The obtained data must be displayed in the table. Comment on the result.
5. Determine the maximum salary for the “Ecologist” of the “jobname” field if the “education” field is “Highest” and the experience of the “experience” field is more than or equal to 3 years.

If a field does not contain information, that is, is empty, it is not necessary to take it into account. This note does not apply to the “schedule” and “employment” fields.

One of the undergraduates for the course work was selected data from a social survey conducted in the United States [12]. When questioning people asked for different information: gender, age, education, marital status, type of employment, number of working hours per week, salary and other parameters. The question was raised about the average length of the working week depending on age. During the decision, the student obtained an interesting result (Fig. 2).

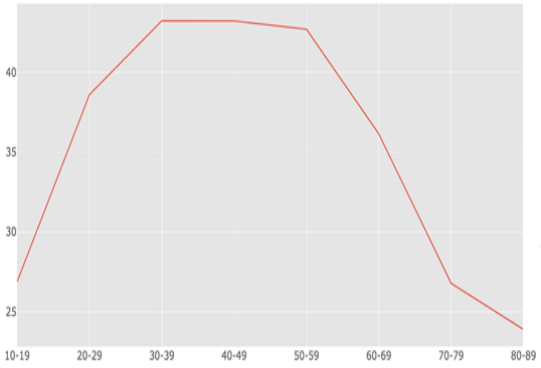


Fig. 2. The average duration of the work week depending on age

The resulting graph is similar to the image of the Gaussian distribution and says that the level of employment depends on the knowledge, experience of the specialist, as well as on the physical condition of citizens. Interest can cause the level of employment of people from 80 to 90 years. A smoother descent can mean that busy people at this age are high professionals and the employer needs such personnel, despite the age of these employees.

VI. CONCLUSIONS

The article presents the program of the discipline “Intelligent databases and data warehouses”, which gives undergraduates practical skills to work with the toolkit of big data processing in a virtual machine from Cloudera.

Using Hadoop with MapReduce and Hive requires virtually no deep knowledge of programming. Using modern tools, you can perform analysis and forecasting, as well as use visualization to simulate problems, which significantly speeds up and facilitates work with big data [1; 8; 11].

Training in the discipline of such a work program is one of the stages in preparing students for the direction 15.04.04 process automation and production process to participate in the WorldSkills IT standards championship in terms of machine learning and big data and Industrial Internet of Things.

VII. REFERENCES

1. Горячев Д.В., Воронов В.И. Большие данные и машинное обучение // Технологии информационного общества: Матер. XII Междунар. отраслевой науч.-техн. конф. М., 2018. С. 327–328. [Goryachev D.V. Big Data and Machine Learning // Information Society Technologies: Proceedings of 12th the international scientific-practical conference. Moscow, 2018. P. 327–328].
2. Voronov V.I., Voronova L.I. Features of realization master’s program “Automation of technological processes and manufactures” // International Journal of Applied and Fundamental Research. 2016. № 2. URL: www.science-sd.com/464-25196.

3. Уонт Р., Шилит Б., Дженсон С. Механизмы Интернета вещей. // Открытые системы. СУБД. 2015. № 1. [Wont R., Shilit B., Jenson S. The mechanisms of the Internet of Things. // Open systems. DBMS. 2015. N 1.]. URL: https://www.osp.ru/os/2015/01/13045328/.

4. Безумнов Д.Н., Воронова Л.И. О развитии и стандартизации технологии Интернета вещей // Технологии информационного общества: Матер. XII Междунар. отраслевой науч.-техн. конф. М., 2018. С. 293–294. [Bezumnov D.N., Voronova L.I. On the development and standardization of the Internet of things technology // Information Society Technologies: Proceedings of XII Intern. branch scientific.-tech. conf. Moscow, 2018. P. 293–294].

5. Применение IoT в энергетике: опыт МОЭК // IoT World Russia Summit. [The use of IoT in the energy sector: the experience of the JSC «MIPC» // IoT World Russia Summit]. URL: http://iotworldsummit.ru/novosti/primenenie-iot-v-energetike.

6. АК «Алроса» завершила проект по внедрению информационной системы управления производством Айхальского ГОК // CNews. [AK “Alrosa” completed the project on the implementation of the production management information system of the Aikhalsky GOK // CNews]. URL: http://www.cnews.ru/news/line/2018-09-05_ak_alrosa_zavershila_proekt_po_vnedreniyu_informatsionnoj.

7. Воронов В.И., Воронова Л.И., Генчель К.В. Применение параллельных алгоритмов в нейронной сети для распознавания жестового языка // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2018). VII Междунар. науч.-техн. и науч.-метод. конф.: Сб. науч. ст.: В 4 т. / Под ред. С.В. Бачевского. М., 2018. С. 207–212. [Voronov V.I., Voronova L.I., Genchel K.V. The use of parallel algorithms in the neural network to recognize the sign language // Actual problems of information and telecommunications in science and education (APINO 2018). VII International scientific tech. and scientific method. Conf.: In 4 t. / Ed. S.V. Bachevsky. M., 2018. P. 207–212].

8. Воронова Л.И., Воронов В.И. Big Data. Методы и средства анализа: Учеб. пос. М., 2016. [Voronova K.I., Voronov V.I. Big Data. methods and tools for analysis. Moscow, 2016]

9. Intro to Hadoop and MapReduce // Udacity. URL: https://classroom.udacity.com/courses/ud617.

10. Downloading a Cloudera QuickStart VM. // Cloudera. URL: https://www.cloudera.com/ downloads/quickstart_vms/5-13.html.

11. Воронов В.И., Усачев В.А. Компетенция “машинное обучение и большие данные” // Приоритетные направления развития науки и образования / Под общ. ред. Г.Ю. Гуляева. Пенза, 2017. С. 97–108. [Usachev V.A., Voronov V.I. The competence “Machine learning and Big Data” // Priorities for development of science and education / Ed. G.Yu. Gulayev. Pensa, 2017. P. 97–108].

12. Adult Data Set // UCI Machine Learning Repository. URL: https://archive.ics.uci.edu/ml/datasets/adult.

Specialists' Training of in Industrial Internet of Things Systems Development

Danil N. Bezumnov[1], Lilia I. Voronova[2]
Department of Intelligent Control Systems and Automation
Moscow Technical University of Communications and Informatics
Moscow, Russia
[1]danbez@ya.ru,[2]voronova.lilia@ya.ru

***Abstract*. Training of qualified industrial Internet of Things systems developers is a key factor in the implementation of the program "Digital economy of the Russian Federation". In this paper the concepts "Internet of Things" and "industrial Internet of things" are compared. Competences of technical specialist in the development of Internet of things systems are defined. Experience of the Department of Intelligent control systems and automation (MTUCI) in training bachelors and masters in "Automation of technological processes and production" и "Management in technical systems". Laboratory workshops is aimed at providing the students with competences of the developers of It-systems are described.**

***Key words.* Internet of things, industrial Internet of things, digital economy, MTUCI, education.**

I. INTERNET OF THINGS

Network technologies are rapidly penetrating into the life of modern man. More devices become able to transmit data streams about users, environment, equipment condition to the network. Now they can be found not only in complex industries, but also in everyday life, on the streets of megacities and medium-sized cities. The adoption of the program "Digital economy of the Russian Federation" in July 2017 contributes to an increase in the pace of digitalization of everyday life of a person and society as a whole, and it's continuously related to the concept of the Internet of things (IoT) [1].

Recommendation Y. 2060 of the International telecommunication Union (ITU-T) defines the IoT as a global infrastructure for the information society that enables the delivery of more complex services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies (Fig. 1).

In the case of IoT, things are objects of the physical or information world that can be identified and integrated into the communication network.

In the discussion of IoT market, there is the identification of this phenomenon with technological solutions that support machine-to-machine (M2M) interaction, such as telemetry or monitoring the status of production facilities. Solutions in this area have existed for a long time and are actively used in engineering, transport, energy, mining, trade and logistics. M2M technologies are used in physical security and surveillance systems.

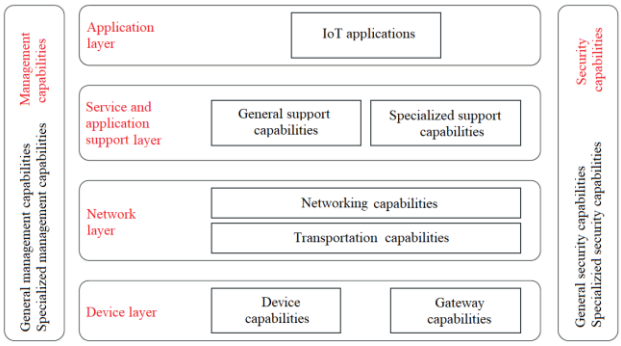


Fig. 1. IoT reference model according to Recommendation Y.2060 of the ITU-T [2, c. 7]

The use of M2M provides sufficient reliable information for decision-making, but requires human involvement for subsequent operations.

The process of transition from M2M to the IoT implies that the information obtained in the course of data mining will allow to make decisions faster and more reliable, to influence the processes without human involvement. It is the analytics of a large amount of data created by different devices that brings the process optimization to another level.

One of the technological vectors of the digital economy is the transition to the industrial Internet of things (IIoT) [3]. IIoT brings together approaches that can seriously change the industry. Some of its elements already existed before this concept took shape, simply were not put together [4].

II. PROFESSIONAL COMPETENCIES

The use of the industrial IIoT involves the creation of a comprehensive solution that combines information and production processes. This is a fairly new task for many professionals, and many factors need to be taken into account when solving it, including industry standards and processes, technological safety and regulatory framework [5].

The developer of IoT systems should have the following competencies:

- work with "things" (microcomputers and sensors): connection, administration and configuration of microcomputers and microcontrollers (Raspberry, Arduino, etc.); programming of microcomputers in C, C++ languages;

connection of microcomputers and microcontrollers with various analog and digital sensors, ADC, DAC; integration of microcomputers and microcontrollers with the Internet and data transfer to the server using wired and wireless technologies (WiFi, Bluetooth);

- creating web applications: application of IDE-integrated web development systems, for example, Microsoft Visual Studio; creating static and dynamic web pages; design of active web pages, placing controls on them; creating web services, setting up data exchange between applications;
- IoT applications development: using the IDE in the IoT area (for example, PTC ThingWorx); setting up communication between devices and IoT application; organization of data transmission and storage; applying basic knowledge to big data and machine learning; configuration of transmission of control signals from a web app to the active IoT device [6].

III. ABOUT THE EDUCATIONAL PROCESS AT THE ICSA DEPARTMENT

Modern students must have skills in developing IoT systems; it is not even a fashion trend, but an urgent need dictated by the fourth industrial revolution and the widespread introduction of IOT and IIoT concepts.

The Department of intelligent control systems and automation (ICSA) of Moscow Technical University of Communications and Informatics (MTUCI) provides training of bachelors in areas 15.03.04 — "Automation of technological processes and production", profile "Industrial Internet of things and robotics", 27.03.04 — "Management in technical systems", profile "Management in cyber physical systems" and training of masters in areas 15.04.04 — "Automation of technological processes and production", the program "Intelligent automated information management systems", 27.04.04 —"Management in technical systems", the program" Intelligent data analysis in technical systems" [7; 8].

Department educational-methodical and scientific work covers modern areas of machine learning, development of control systems using data mining methods [9–11].

As part of the educational programs of the Department, part of the disciplines aimed at obtaining knowledge in the field of digital control systems, skills to program microcontrollers and microprocessors, design skills of complex distributed technical systems. Examples of such disciplines are "Cyberphysical systems and the Internet of things", "Microprocessors in robotic control systems", "Industrial Internet of things Technologies", "Digital technologies of Smart City", "Design of cyberphysical systems", educational and pre-diploma practice.

Educational meccanos "Svyaznoy" and "Malina Z" on the basis of the Arduino module and a single-Board microcomputer Raspberry respectively are used as the material and technical base for the construction of systems for the collection and processing of information from sensors and control of executive devices in laboratory workshops.

The initial acquaintance of students with programming microcontrollers and design of technical systems begins in the first year of the bachelor's degree in the laboratory workshop on the discipline "Cyberphysical systems and the Internet of things".

The first part of the laboratory workshop is aimed at teaching students to create simple microprocessor systems for collecting and processing information.

Laboratory work "Creation of control algorithms for actuators" includes reading information from various sensors, its processing in accordance with a given algorithm and development of control commands for actuators. The state of devices (Fig. 2) can be changed by students directly in the laboratory, which adds clarity to the laboratory work.

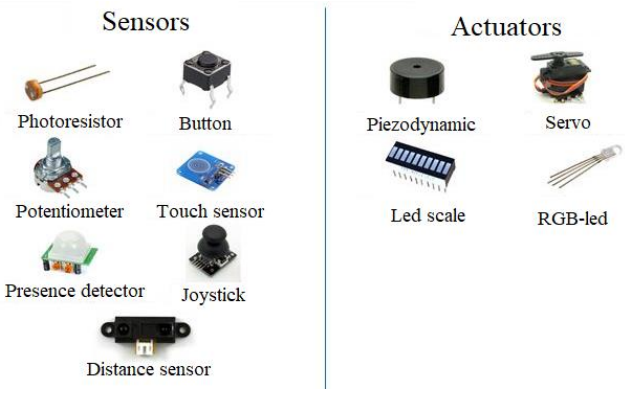


Fig. 2. Element base of laboratory work
"Creation of control algorithms for executive devices"

The purpose of the work is to consolidate knowledge of electrical engineering: analog-digital conversion, pulse width modulation, voltage divider circuit; skills of writing programs with conditions and cycles, data transfer to the PC by serial connection.

During the work, students learn arithmetic and logical operations, functions of reading and writing analog signal *analogRead()* and *analogWrite()*, proportional transfer of values *map()*, limiting the value in the range of values *constrain()*, creating conditions *if()*, cycles *for()*, gain skills in working with arrays, knowledge of methods of debugging programs to eliminate time delays in the system.

Getting students competencies in the development of IoT applications takes place in the laboratory workshop on the discipline "Microprocessors in control systems" [12].

Students are invited to develop a control system of light sources and blinds in the room during the laboratory work "Development of lighting control system" (Fig. 3).

The developed system should automatically maintain the level of illumination at a value set by the user. At the same time, user interaction with the system is performed through an application created in the RemoteXY environment (Fig. 4).

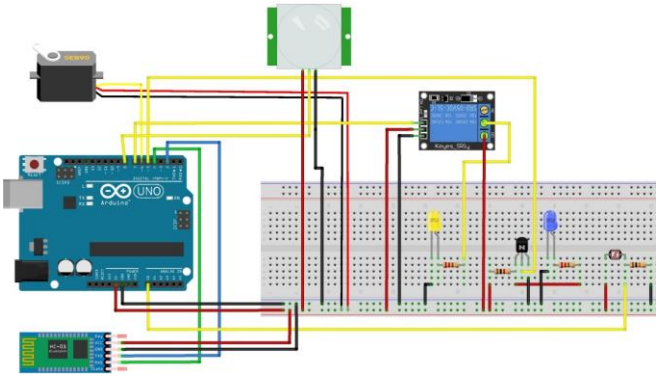


Fig. 3. Wiring diagram of lighting control system

Fig. 4. Mobile application interface for lighting control

Laboratory workshop on disciplines “Fundamentals of mechatronics and robotics”, “Microprocessors in robotic control systems”, “Technological processes in robotic logistics systems” includes the work of students with models of real robotic systems.

The prototype of the robot manipulator with elements of computer vision (Fig. 5) allows the automatic capture of objects at a distance achievable for it.

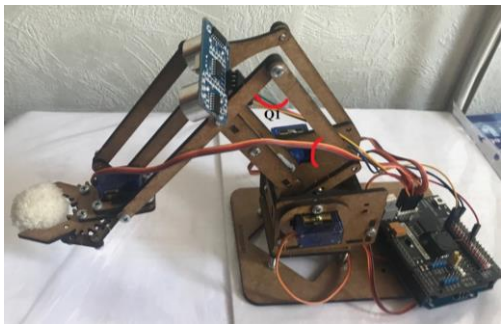


Fig. 5. The prototype of the robot manipulator with elements of computer vision

In the process of switching elements of the system and the development of the algorithm of the system, students solve a wide class of problems: from the inverse problem of kinematics to the creation of the graphical interface of the dispatching system.

The prototype of the output terminal of the automated warehouse of the manufacturer (Fig. 6) is used in the performance of a cycle of laboratory work, during which students need to design and implement a complex algorithm for the supply of goods from stores located in a row at the top of the layout and equipped with dumpers, and their consistent stacking in the truck body using a “boom” controlled by a servo drive.

Fig. 6. The prototype of the output terminal of the automated warehouse of the manufacturer

During the design of the algorithm, students use the methods of Queuing theory, and during its implementation write a program that controls the actuators at the physical level, and the operator interface of the output terminal.

The above-mentioned laboratory workshops allow students to get the full range of competencies required by the modern developer of Internet of things systems, including for industrial applications.

IV. SAFETY AND RISKS

Problems of fault tolerance and safety of the created systems come to the first place when testing them at the logical and physical levels [13].

Libraries and user applications may contain errors that lead to vulnerabilities. Within the security of the IoT, students need to evaluate what is the probability that any given IoT devices have errors. What is the probability that some of these errors represent vulnerabilities? How likely are these vulnerabilities to be exploited now or in the future?

These issues are discussed in the performance and demonstration of laboratory work, lectures and seminars.

V. SUMMARY

The IoT is one of the leading and actively developing technological concepts that can radically change all spheres of human life, including industry.

Training of specialists capable of preparing the technological base for the implementation of the program “Digital economy of the Russian Federation” is a priority for technical universities.

The ICSA Department of MTUCI implements bachelor's and master's programs, in which students get the key competencies of developers of Internet of things systems with a focus on robotics and industry.

For comprehensive and all-round professional development, students are taught all stages of IoT systems development, from design and assembly of "field" level equipment to creation of dispatching systems, testing of created programs, analysis of their safety and fault tolerance.

VI. REFERENCES

[1] Безумнов Д.Н., Воронова Л.И. О поддержке дисциплин, включающих изучение «умного дома», с использованием конструкторов на базе Arduino и Raspberry // Приоритетные направления развития науки и образования / Под общ. ред. Г. Ю. Гуляева. Пенза, 2017. С. 109–118. [Bezumnov D.N., Voronova L.I. Disciplines supplying associated with the study of "smart house", using Arduino and Raspberry // Priorities for development of science and education / Ed. G.Yu. Gulayev. Pensa, 2017. P. 97–108].

[2] Рекомендация МСЭ-Т Y. 2060. Обзор интернета вещей // Iotas. [Recommendation ITU-T Y.4000/Y.2060. Overview of the Internet of things. 2012. // Iotas.] URL: https://iotas.ru/files/documents/wg/T-REC-Y.2060-201206-I!!PDF-R.pdf.

[3] Хабаров И.Е., Киричек Р.В. Анализ использования промышленных протоколов в рамках концепции промышленного интернета вещей // Интернет вещей и 5G (INTHITEN 2017). СПб.: СПбГУТ, 2017. С. 181–185. [Khabarov I.Ye., Kirichek R.V. Analysis of the use of industrial protocols within the concept of the industrial Internet of things // Internet of things and 5G (INTHITEN 2017). Saint-Petersburg: SPSUT, 2017. P. 181–185.]. URL: http://inthiten.org/doc/2017/181-185.pdf.

[4] Коминек Д., Местари Д., Колдер Б. и др. Что такое промышленный «интернет вещей»? // ControlEngineering Россия. 2016. № 4. С. 104-105. []Kominek D., Mestari D., Kolder B. What is the industrial "Internet of things"? // Control Engineering Russia. 2016. № 4. P. 104-105]. URL: http://controleng.ru/wp-content/uploads/64104.pdf.

[5] Семеновская Е. Индустриальный интернет вещей. Перспективы российского рынка // Ростелеком. URL: https://www.company.rt.ru/projects/IIoT/study_IDC.pdf.

[6] Техническое описание компетенции «Интернет вещей» // Региональный координационный центр WorldSkillsRussia. URL: http://mospolytech.ru/storage/files/ws/internet/WSR2016_UM_IoT_TD.pdf.

[7] Воронов В.И., Усачев В.А. Компетенция «Машинное обучение и большие данные» // Приоритетные направления развития науки и образования / Под общ. ред. Г.Ю. Гуляева. Пенза, 2017. С. 97–108. [Usachev V.A., Voronov V.I. The competence "Machine learning and Big Data" // Priorities for development of science and education / Ed. G.Yu. Gulayev. Pensa, 2017. P. 97–108].

[8] Voronov V.I., Voronova L.I. Features of realization master's program "Automation of technological processes and manufactures" // International journal of applied and fundamental research. 2016. № 2. URL: www.science-sd.com/464-25196.

[9] Безумнов Д.Н., Воронова Л.И. О развитии и стандартизации технологии Интернета вещей // Технологии информационного общества: Матер. XII Междунар. отраслевой науч.-техн. конф. М., 2018. С. 293–294. [Bezumnov D.N., Voronova L.I. On the development and standardization of the Internet of Things technology // Information Society Technologies: Proceedings XII International. branch scientific.-tech. conf. Moscow, 2018. P. 293–294].

[10] Воронова Л.И., Воронов В.И. Big data. Методы и средства анализа: Учеб. пос. М., 2016. [Voronova K.I., Voronov V.I. Big Data. Methods and tools for analysis. Moscow, 2016].

[11] Соколов В.П., Безумнов Д.Н. Математическая модель полета квадрокоптера // Приоритетные направления развития науки и образования / Под общ. ред. Г. Ю. Гуляева. Пенза, 2017. С. 130–142. [Sokolov V.P., Bezumnov D.N. Mathematical model of a quadrocopter flight // Priorities for development of science and education / Ed. G.Yu. Gulayev. Pensa, 2017. P. 97–108].

[12] Безумнов Д.Н., Воронова Л.И. Микропроцессоры в системах управления: Учеб.-метод. пос. по выполнению лабораторных работ / МТУСИ. М., 2018. [Bezumnov D.N., Voronova L.I. Microprocessors in control systems: workshop on laboratory work / MTUCI. Moscow, 2018].

[13] Врагова Е.В., Воронова Л.И. Методы и средства защиты от botnet's (зомби-сетей) // Телекоммуникации и информационные технологии. 2018. Т. 5, № 1. С. 112–116. [Vragova Ye. V., Voronova L.I. Methods and means of protection against botnet's (zombie networks) // Telecommunications and information technology. 2018. Vol. 5, № 1. P. 112–116.].

Effective Precoding Technique for Multiuser MIMO Systems*

Taoufik Ben Rejeb
Department of Information Security
Moscow Technical University of Communications and Informatics
Moscow, Russia
benrejebt@yandex.ru

***Abstract*. Using of quantized channel state information on transmitter (CSIT) side for adaptive signal processing on base station (BS) side also called as precoding provides to reduce feedback channel bandwidth in comparison with full CSIT. There are many solutions for feedback channel quantization but the aim of this paper is to note the performance of Grassmannian manifold for codebook design.**

***Keywords:* Multiuser MIMO, MU-MIMO, CLTD-OSIC precoding, Grassmann manifold, quantizaion codebook.**

I. INTRODUCTION

The current radio communication standards include the spatial processing of signals implying the use of MIMO (Multiple-Input-Multiple-Output) multi-antenna systems [1]. The use of MIMO technology in communication systems makes it possible to significantly increase the capacity, but involves the necessity of solving many problems of signal processing. The task of adaptive signal processing for the simultaneous operation for several users on the same allocated frequency-time resources falls on the computing power of both the BS and users sides.

Precoding performed on the BS side based on the CSIT obtained with feedback channel allows to calculate the necessary weights for the most efficient transmission of information over the communication channel.

II. SYSTEM MODEL

We consider a multiuser MU-MIMO system with spatial division multiple access (SDMA) [1], which consist of multiple antenna BS and several users where each of them has multiple antenna receivers. Let's introduce the following notations: N — number of antennas on the BS side; M — sum number of antennas on the user side; k — index number of user ($k = 1, 2, \ldots, K$); K — total number of users in a multiuser system; M_k — number of antennas on k-th user side $\mathbf{H}_k$ — complex channel matrix of dimension $M \times N$, where each element of $\mathbf{H}_k$ is a complex channel multiplier between user antennas and BS antennas; $\mathbf{H} = \left[\mathbf{H}_1^T \quad \mathbf{H}_2^T \quad \ldots \quad \mathbf{H}_K^T\right]$ — general matrix of MIMO channel between all users and BS of dimension $K \times N$, which consists of channel matrices $\mathbf{H}_k$; $\mathbf{T}_k$ — precoding matrix r dimension of $M_k \times N$; $\mathbf{T} = \left[\mathbf{T}_1 \quad \mathbf{T}_2 \quad \ldots \quad \mathbf{T}_K\right]$ — general precoding matrix of BS dimension of $N \times K$, which consist of precoding matrices $\mathbf{T}_k$ for K users, where each of them contains the weighting factors for signals transmitted be BS. The structure of general BS's precoding matrix consisting of precoding vectors is shown in Figure 1.

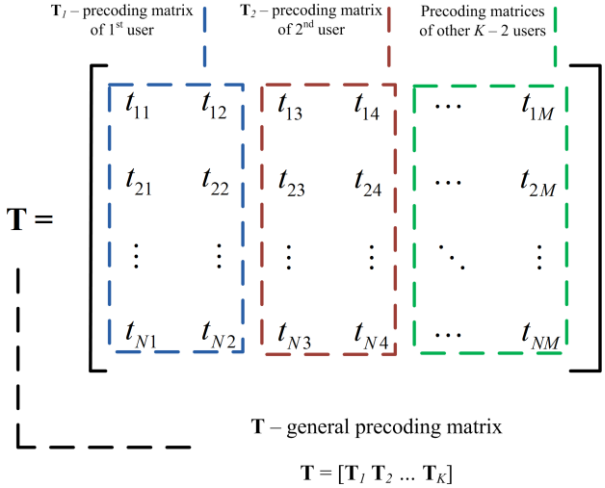


Fig. 1. Structure of general BS's precoding matrix

The main task of a base station is simultaneous data transmission for all K users using N transmit antennas and the same time-frequency resources. The precoding procedure is performed on the BS side, where the weights for the transmitting signals for each user are given by the common precoding matrix T. The calculation of the matrix T is based on CSIT obtained by feedback channel from the each user to the BS.

The signal observed at the k-th receiver can be represented as follows [1]:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{T}_k \mathbf{s}_k + \sum_{j=1, j\neq k}^{K} \mathbf{H}_k \mathbf{T}_j \mathbf{s}_j + \mathbf{n}_k , \qquad (1)$$

where $\mathbf{y}_k$ — vector of received symbols observed at the k-th receiver; $\mathbf{s}_k$ — symbol vector which was transmitted from BS to k-th user; $\mathbf{T}j$ — precoding matrix of other j-th user, where $j \neq k$; $\mathbf{n}_k$ represents the vector of additive noise at the k-th receiver. Second part of (1) represents interfering signals of other users of the system with respect to the signal of the k-th user.

The precoding procedure requires the feedback channel having sufficient bandwidth for periodically transmitting CSIT to BS's side. The formation of such a feedback channel may require resources (time slots or frequency bands) that are maximally used for data transmission channels.

Since the presence of feedback channels is necessary for using the precoding procedure, the actual task is to find ways to reduce the amount of resources (temporary or frequency) needed to organize such channels.

III. LIMITED FEEDBACK

Precoding based on partial channel information was used in known standards of wireless communication systems such as LTE and LTE-Advanced, where we can find term "codebook-based precoding" [4]. The use of such precoding technique requires a codebook, known on the BS and AT side. The codebook is a set of vectors (codewords) consisting of complex channel multipliers. With a known codebook user compares the channel estimate with each of the codebook vectors, and transmits through the feedback channel to the BS only the index number of the selected vector. Quantization of the channel state information allows the transmission of only certain index numbers through the feedback channel represented by B bits, which determine the result of the calculation of the precoding vectors on the BS side. The number of bits B depends of the codebook F used on the user side for quantizing the channel state information. For transmission of quantized channel information to the BS side, the CDI (Channel Direction Information) indicator [6] is used. For estimation of signal level on the user side, the CQI (channel quality information) indicators are used. To evaluate the signal power, user terminals use the SINR — Signal Interference + Noise Ratio [4].

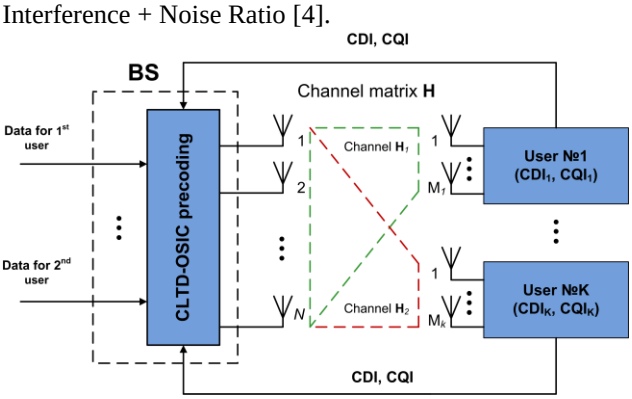


Fig. 2. MU-MIMO system with limited feedback

CODEBOOK DESIGN AND SELECTION CRITERIA

User's codebook can be represented as an $L \times N$ matrix F consisting of L/M_k matrices dimension of $M_k \times N$:

$$\mathrm{F} = \left[\mathbf{F}_1 \quad \mathbf{F}_2 \quad \ldots \quad \mathbf{F}_L \right], L = 2^B , \qquad (2)$$

where $L = 2^B$; l — the index number of the codeword. An illustration of user's codebook structure is shown on Figure 3.

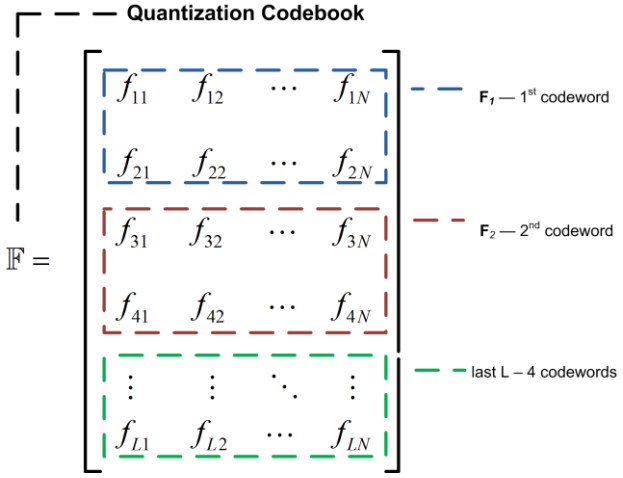


Fig. 3. User's codebook structure

The subscriber terminal codebook is an $L \times N$ dimension matrix, where each row matrix of this codebook is a $L \times N$ codeword. The codewords are used by subscriber terminals as quantized values of channel state information (Fig. 4). The size of the user's codebook depends of the number of bits B transmitted on the feedback channel from the user to BS. High-dimensional codebook requires a larger value of B, and therefore a wider feedback channel, which is not always possible.

Availability of quantized channel vector at base station side makes it possible to calculate or select the precoding matrix. There are two approaches to feedback quantization:

- error calculation between the channel state information and codeword and deciding of precoding matrix on the BS side [10];
- selection the precoding vector on the user side by using different criteria [10].

In the case of the first approach, the following difference criterion is applicable:

$$\mathbf{F}_l = \arg\max_{\mathbf{F}_l \in \mathrm{F}} \left\| \mathbf{H}_k \tilde{\mathbf{F}}_l \right\|_F \qquad (3)$$

Authors in [9] propose to generate codebooks of users based on Grassmann manifold. Grassmannian $\mathbb{G}_{K,N}$ or Grassmann manifold is the set of subspaces of dimension K in N-dimensional complex space. Using of Grassmann manifold packing provide to geometrically represent codewords of codebook F as matrices of various dimensions in a given multidimensional space. Using of Grassmannian package make it possible to arrange codewords represented by subspaces in a multidimensional space using various metrics.

There are different distance metrics between subspaces [9], but for the case of user's equipment with

multiple antenna, the codebook will consist of Grassmanian subspaces (matrices). In this case, the distance metric is applicable:

- Fubini-Study metric*:*

$$d_{FS}(\mathbf{F}_l, \mathbf{F}_m) = \arccos\left|\det\left(\mathbf{F}_m^H \mathbf{F}_l\right)\right|, \quad (4)$$

- Chordal distance metric:

$$d_{ch}(\mathbf{F}_l, \mathbf{F}_m) = \frac{1}{\sqrt{2}}\left\|\mathbf{F}_l\mathbf{F}_l^H - \mathbf{F}_m\mathbf{F}_m^H\right\|_F. \quad (5)$$

IV. SIMULATION RESULTS

The results of computer simulation, presented in Table, allow to compare the results of noise immunity of MU-MIMO systems with proposed CLTD-OSIC precoding technique using grassmannian manifold and MMSE precoding based on full channel matrix knowledge.

SIMULATION PARAMETERS

Channel	**MIMO rayleight fading channel**
Antenna configuration	4x2x2
Number of antennas on BS side	4
Number of antennas on user side	2
Number of users	2
Transceiver architecture	V-BLAST
Channel encoding	Convolution encoding (rate 1/2)
Precoding technique	CLTD-OSIC
Number of feedback bits	6 bits
Demodulator	MMSE
Modulation	QPSK
Number of experiments	100000

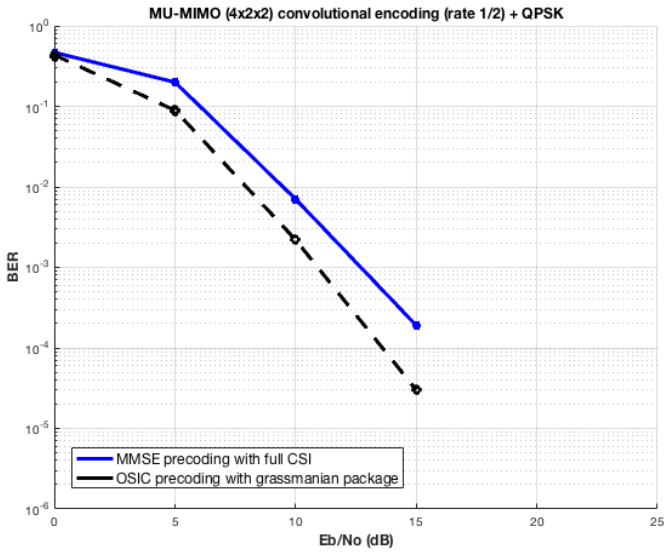


Fig. 4. Simulation results

As shown on Figure 4, using of proposed CLTD-OSIC precoding technique based on Grassmannian manifold provides to gain 2 dB of the energy efficiency in comparison with known MMSE precoding technique based on full channel matrix knowledge.

V. REFERENCES

[1] Бакулин М.Г., Варукина Л.А., Крейнделин В.Б. Технология MIMO: принципы и алгоритмы. М.: Горячая линия — Телеком, 2014. [Bakulin M.G., Varukina L.A., Kreyndelin V.B. MIMO Technology: Principles and Algorithms. Moscow: Hotline — Telecom, 2014].

[2] Kreyndelin V., Smirnov A., Ben Rejeb T. Effective precoding and demodulation techniques for 5G communication systems // Proc. Systems of Signals Generating and Processing in the Field of on Board Communications. 2018. P. 1–6.

[3] Brown T., De Carvalho E., Kyritsi P. Practical guide to the MIMO radio channel: with MATLAB examples. Chichester: John & Wiley Sons Ltd, 2012.

[4] Sibille A., Oestges C., Zanella A. MIMO: From Theory to Implementation. : Elsevier Ltd., 2011.

[5] Yoo T., Jindal N., Goldsmith A. Multi-Antenna Broadcast Channels with Limited Feedback and User Selection // IEEE Journal Sel. Areas in Communications. 2007. P. 1478–1491.

[6] Santipach W., Honig M. Asymptotic capacity of beamforming with limited feedback // Proc. IEEE Int. Symp. Inform. Theory (ISIT). 2004. July. P. 290.

[7] Brown III D.R., Love D.J. MIMO Nullforming with RVQ Limited Feedback and Channel Estimation Errors // Proceedings of the 48th Asilomar Conference on Signals, Systems, and Computers. Pacific Grove, CA, November 2–5, 2014. Pacific Grove, 2014.

[8] Yoo T., Jindal N., Goldsmith A. Multi-Antenna Broadcast Channels with Limited Feedback and User Selection // IEEE Journal on Selected Areas in Communications. 2007. Vol. 25, N 7.

[9] Medra A., Davidson T. Flexible codebook design for limited feedback downlink systems via smooth optimization on the Grassmannian manifold // IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications. 2012.

[10] Tsoulos G. MIMO System Technology for Wireless Communications. Boca Raton: CRC Press, 2006.

Filtration of Linear Recurrent Sequences with Random Delay and Random Initial Phase

M.G. Bakulin
Department of Theory of Electrical Circuits

V.B. Kreyndelin[2]
Department of Theory of Electrical Circuits

D.Yu. Pankratov[3]
Department of Systems and Networks of Radio Communications and Radio Broadcasting

Moscow Technical University of Communications and Informatics
[1]m.g.bakulin@gmail.com, [2]vitkr@gmail.com, [3]dpankr@mail.ru

***Abstract*. Radio communication systems, especially mobile communication systems, are improving very fast towards 5G systems. As the requirements for radio communication systems increase the requirements for synchronization also increase. Linear recurrent sequences are widely used in radio communications systems and in this paper filtration algorithm for such sequences are considered. The results of computer simulation of the algorithms for various signal-to-noise ratios and accumulation time are presented. The synthesized algorithms can be classified as quasi-noncoherent (subincoherent) algorithms, as opposed to noncoherent algorithms that use averaging over the entire region of phase uncertainty, or from coherent, which use the exact phase of the received signal.**

***Key words:* radio communication systems, filtration algorithms, Markov chains, linear recurrent sequences, synchronization algorithms, quasi-noncoherent algorithms, noncoherent algorithms, quasi-coherent algorithms, coherent algorithms.**

VI. INTRODUCTION

Pseudo-random signals based on linear recurrent sequences (LRS) [1–4] are widely used in various information systems [5]. LRS synchronization in such systems is usually performed by multichannel correlators or cyclic search schemes [6–9]. With a long sequence length and a wide area of uncertainty over the delay, the implementation of these approaches requires significant technical and time resources.

Known algorithms for receiving LRS that solve the problem of accelerated synchronization are based on the Ward method [10]. This method allows to reduce the time of search and synchronization of LRS. In Ward's method, the LRS generator is started by a sequence of previously received characters. However, this method has acceptable characteristics only at a sufficiently high signal-to-noise ratio (SNR). Decreasing the SNR significantly degrades the effectiveness of the Ward method. To improve the efficiency of the Ward method while decreasing the SNR, different modifications were proposed [11–14].

The known algorithms of filtration for linear recurrent sequences using the Markov properties of these LRSs allow to improve the synchronization performance [11–14]. However, these methods belong to coherent reception methods and require precise knowledge of the phase of the received radio signal. It is known [15] that coherent methods provide the highest noise immunity during LRS receiving. However, in reality the received signal always has random initial phase. Therefore, the use of coherent methods of reception requires an estimation of this phase and for obtaining the phase estimate the use of appropriate algorithms and additional time are required. This ultimately increases the total time for synchronization.

In this paper, we solve the problem of LRS filtration algorithm synthesizing for accelerated synchronization, which does not require knowledge of the phase of the received radio signal. The proposed algorithm can be used for accelerated synchronization and, as the simulation results show, remains operable even for negative SNR values.

VII. PROBLEM STATEMENT

Let the LRS be described by the following recurrent relation:

$$b_n = \sum_{i=1}^{k} \oplus \delta_i b_{n-i}, \delta_i \in \{0;1\}, b_n \in \{0;1\}, i \in 1,...,k \quad (1)$$

where k is the length of LRS register; n is discrete time; sign $\oplus$ means modulo 2 sum.

It is known [16] that the operations of modulo 2 addition in the alphabet {0; 1} correspond to the operations of modulo 2 addition in the alphabet {1; –1}. Based on this, from (1) we can get:

$$B_n = \prod_{i=1}^{k} \left((1-\delta_i) + \delta_i B_{n-i} \right) \qquad (2)$$

where $B_n = 1 - 2b_n, B_n \in \{1; -1\}$

Expressions (1) and (2) describe the rule for LRS generation and are valid for any cyclic shift. Therefore, the problem of LRS synchronization, which consists in estimating the number of cyclic shift (discrete delay) of the observed sequence, can be considered as the problem of estimating the current LRS segment of length k characters, i.e. k-dimensional vector estimation $\mathbf{B}_n = [B_n \quad \cdots \quad B_{n-k+1}]$.

For the sake of simplicity of the further consideration, let us assume that LRS register has only one tap, i.e.:

$$\delta_i = \begin{cases} 1, & i = l \\ 0, & i \neq l \end{cases}, \quad i = 1,...k, 1 < l < k \qquad (3)$$

We assume that the pseudo-random binary sequence (PRBS) is transmitted using binary phase shift keying (BPSK). Taking into account (1)–(3) we can write the following model:

$$\begin{cases} B_n = B_{n-l} B_{n-k} \\ y_{c,n} = B_n \cos\varphi + \eta_{c,n}, \quad \eta_{c,n} \sim N(0, \sigma_\eta^2), \\ y_{s,n} = B_n \sin\varphi + \eta_{s,n}, \quad \eta_{s,n} \sim N(0, \sigma_\eta^2) \end{cases} \qquad (4)$$

where φ is random phase, but it does not change on the observation interval and uniformly distributed on the interval $[-\pi; \pi]$; η_{nc}, η_{ns} are uncorrelated sequences of Gaussian random variables with zero means and variances σ_η^2 ; y_{nc}, y_{ns} are observed signals. The equation of state in model (4) (upper equation) is written taking into account the fact that LRS can be considered as a degenerate case of a Markov chain.

The task of synchronizing LRS is to estimate the current segment $\mathbf{B}_n \sim [B_n \quad \cdots \quad B_{n-k+1}]$ from the observation $\mathbf{y}_n = (y_1, y_2, ...y_n)$, where $y_n \sim [y_{c,n} \quad y_{s,n}]$.

Model (4) can be transformed and represented in the following form:

$$\begin{cases} B_n = B_{n-l} B_{n-k} \\ y_{c,n} = B_n \theta_c |\cos\varphi| + \eta_{c,n} \\ y_{s,n} = B_n \theta_s |\sin\varphi| + \eta_{s,n} \end{cases} \qquad (5)$$

where $\theta_c = \text{sign}(\cos(\varphi)) \in \{-1; 1\}$;

$\theta_s = \text{sign}(\sin(\varphi)) \in \{-1; 1\}$ are additional symbols that take into account the position of the phase vector in the corresponding half-planes.

Let us introduce the notation for the following sequences:

$$B_{c,n} \sim B_n \theta_c, \; B_{s,n} = B_n \theta_s \qquad (6)$$

It is not difficult to show that $B_{c,n} = B_{c,n-l} B_{c,n-k} \theta_c$, $B_{s,n} = B_{s,n-l} B_{s,n-k} \theta_s$.

We divide the equations of state and observation (5) into two independent systems:

$$\begin{cases} B_{c,n} = B_{c,n-l} B_{c,n-k} \theta_{c,n-1} \\ \theta_{c,n} = \theta_{c,n-1} \\ y_{c,n} = B_{c,n} \cos\varphi_c + \eta_{c,n} \end{cases} \qquad (7)$$

and

$$\begin{cases} B_{s,n} = B_{s,n-l} B_{s,n-k} \theta_{s,n-1} \\ \theta_{s,n} = \theta_{s,n-1} \\ y_{s,n} = B_{s,n} \sin\varphi_s + \eta_{s,n} \end{cases} \qquad (8)$$

where φ_c is random phase uniformly distributed on the interval $[-\pi/2; \pi/2]$; φ_s is random phase uniformly distributed on the interval $[0; \pi]$.

Thus, the task of synchronizing LRS can be divided into the following subtasks:

1. Using the systems of equations of state and observation (7) and (8), estimate the k-dimensional segments of two independent LRSs $\mathbf{B}_{c,n}$, $\mathbf{B}_{s,n}$ and random binary values $\theta_{c,n}$, $\theta_{s,n}$.
2. Using estimates of segments $\mathbf{B}_{c,n}$, $\mathbf{B}_{s,n}$ and random binary values $\theta_{c,n}$, $\theta_{s,n}$ estimate the segment of source LRS $\mathbf{B}_n$.

III. SYNTHESIS OF THE ALGORITHM

Consider first the synthesis of synchronization algorithm of sequence $B_{n,c}$ using model (7).

Using the Bayes formula [15], we can obtain an expression for the joint a posteriori distribution of the vector $\mathbf{B}_{c,n} \sim [B_{c,n}, B_{c,n-1}, ..., B_{c,n-k+1}]$ and parameters $\theta_{c,n}$, at the *n*-th filtration step:

$$\begin{aligned} &p(\mathbf{B}_{c,n}, \theta_{c,n} / \mathbf{y}_{c,n}) = \\ &= D(y_{c,n}) \cdot p(\mathbf{B}_{c,n}, \theta_{c,n} / \mathbf{y}_{c,n-1}) \cdot p(\mathrm{y}_{c,n} / B_{c,n}) \end{aligned} \qquad (9)$$

This equation takes into account that the current observation $y_{c,n}$ depends only on the current LRS character $B_{c,n}$ and does not depend on the rest of the estimated characters.

The relation for the likelihood function can be written as follows:

$$p(\mathrm{y}_n / B_{c,n}) = \int_{-\pi/2}^{\pi/2} p(\mathrm{y}_n / B_{c,n}, \varphi_c) \cdot p(\varphi_c) d\varphi_c . \quad (10)$$

where $p(\varphi_c) = \frac{1}{\pi}$ for $-\frac{\pi}{2} \le \varphi_c \le \frac{\pi}{2}$.

From observation equation of system (7), we can write the expression for the likelihood function:

$$p(\mathrm{y}_n / B_{c,n}, \varphi_c) =$$
$$= C \exp\left(-\frac{1}{2\sigma_\eta^2} \left(y_{c,n} - B_{c,n} \cos(\varphi_c) \right)^2 \right) =$$
$$= \frac{1}{2} C \exp\left(-\frac{y_{c,n}^2}{2\sigma_\eta^2} \right) \exp\left(-\frac{\cos(\varphi_c)^2}{2\sigma_\eta^2} \right) \cdot \quad (11)$$
$$\cdot \left(\cosh\left(\cos(\varphi_c) \right) + B_{c,n} \sinh\left(\cos(\varphi_c) \right) \right)$$

Substituting (11) into (10) after averaging, we get

$$p(\mathrm{y}_n / B_{c,n}, \varphi_c) =$$
$$= C \exp\left(-\frac{1}{2\sigma_\eta^2} \left(y_{c,n} - B_{c,n} \cos(\varphi_c) \right)^2 \right) =$$
$$= \frac{1}{2} C \exp\left(-\frac{y_{c,n}^2}{2\sigma_\eta^2} \right) \exp\left(-\frac{\cos(\varphi_c)^2}{2\sigma_\eta^2} \right) \cdot \quad (12)$$
$$\cdot \left(\cosh\left(\frac{y_{c,n}}{\sigma_\eta^2} \cos(\varphi_c) \right) + B_{c,n} \sinh\left(\frac{y_{c,n}}{\sigma_\eta^2} \cos(\varphi_c) \right) \right)$$
$$= \frac{1}{2} C_2 \left(1 + B_{c,n} T_{c,n} \right)$$

where

$$T_{c,n} \square T_c \left(\frac{y_{c,n}}{\sigma_\eta^2} \right) =$$
$$= \frac{\int_{-\pi/2}^{\pi/2} \exp\left(-\frac{\cos(\varphi_c)^2}{2\sigma_\eta^2} \right) \sinh\left(\frac{y_{c,n}}{\sigma_\eta^2} \cos(\varphi_c) \right) d\varphi_c}{\int_{-\pi/2}^{\pi/2} \exp\left(-\frac{\cos(\varphi_c)^2}{2\sigma_\eta^2} \right) \cosh\left(\frac{y_{c,n}}{\sigma_\eta^2} \cos(\varphi_c) \right) d\varphi_c}$$

Assuming a posteriori independence of the vector elements $\mathbf{B}_{c,n-1}$ between themselves and и value $\theta_{c,n-1}$ at n-1-th filtration step we can write:

$$p(\mathbf{B}_{c,n-1}, \theta_{c,n-1} / \mathbf{y}_{c,n-1}) =$$
$$= p(\theta_{c,n-1} / \mathbf{y}_{c,n-1}) \prod_{i=1}^{k} p(B_{c,n-i} / \mathbf{y}_{c,n-1})$$

Ошибка! Закладка не определена.

Since $B_{n-i} = \pm 1, \quad i = 1, \ldots k$ and $\theta_{c,n-1} = \pm 1$ are random binary values, the corresponding a posteriori probabilities are described by the following expressions [16]:

$$p\left(\theta_{c,n-1} / \mathbf{y}_{c,n-1} \right) = \frac{1}{2} \left(1 + \hat{\theta}_{c,n-1} \theta_{c,n-1} \right)$$
$$p\left(B_{c,n-i} / \mathbf{y}_{c,n-1} \right) = \quad (13)$$
$$= \frac{1}{2} \left(1 + \hat{B}_{c,n-i}^{(n-1)} B_{c,n-i} \right), \; i = 1, \ldots k$$

where $\hat{\theta}_{c,n-1} = E\left\{ \theta_{c,n-1} / \mathbf{y}_{c,n-1} \right\}$,

$\hat{B}_{c,n-i}^{(n-1)} = E\left\{ B_{c,n-i} / \mathbf{y}_{c,n-1} \right\}$ estimates of $\theta_{c,n-1}$ and $B_{c,n-i}$, $i = 1, \ldots k$ values on n-1-th filtration step, optimal by the criterion of the minimum mean square error.

Taking into account the equation of state in expression (7) and expressions (14), we can write:

$$p(\mathbf{B}_{c,n}, \theta_{c,n} / \mathbf{y}_{c,n-1}) = \frac{1}{2} \left(1 + \hat{\theta}_{c,n-1} \theta_{c,n} \right) \cdot$$
$$\cdot \frac{1}{2} \left(\left(1 + \hat{B}_{c,n-k}^{(n-1)} B_{c,n-l} B_{c,n} \theta_{c,n} \right) \right) \cdot \quad (14)$$
$$\cdot \prod_{i=1}^{k-1} \frac{1}{2} \left(1 + \hat{B}_{c,n-i}^{(n-1)} B_{c,n-i} \right)$$

Substituting (15) and (12) into (9) and transforming, we obtain a recurrent algorithm for calculating estimates:

$$\hat{B}_{c,n-r}^{(n)} = \begin{cases} \hat{B}_{c,n-r}^{(n-1)}, \; r \ne l, \; r \ne 0 \\ \dfrac{\hat{B}_{c,n-l}^{(n-1)} + T_{c,n} \hat{\theta}_{c,n-1} \hat{B}_{c,n-k}^{(n-1)}}{1 + T_{c,n} \hat{\theta}_{c,n-1} \hat{B}_{c,n-l}^{(n-1)} \hat{B}_{c,n-k}^{(n-1)}}, \; r = l, \quad r = 0, \ldots k-1 \\ \dfrac{T_{c,n} + \hat{\theta}_{c,n-1} \hat{B}_{c,n-l}^{(n-1)} \hat{B}_{c,n-k}^{(n-1)}}{1 + T_{c,n} \hat{\theta}_{c,n-1} \hat{B}_{c,n-l}^{(n-1)} \hat{B}_{c,n-k}^{(n-1)}}, \; r = 0 \end{cases}$$
$$\hat{\theta}_{c,n} = \frac{\hat{\theta}_{c,n-1} + T_{c,n} \hat{B}_{c,n-l}^{(n-1)} \hat{B}_{c,n-k}^{(n-1)}}{1 + T_{c,n} \hat{\theta}_{c,n-1} \hat{B}_{c,n-l}^{(n-1)} \hat{B}_{c,n-k}^{(n-1)}} \quad (15)$$

Using the well-known relation for hyperbolic functions, the resulting algorithm can be transformed to a somewhat simpler form:

$$\hat{B}_{c,n-r}^{(n)} = \begin{cases} \hat{B}_{c,n-r}^{(n-1)}, \ r \neq l, \ r \neq 0 \\ \tanh(\operatorname{atanh}\left(\hat{B}_{c,n-l}^{(n-1)}\right) + \\ + \operatorname{atanh}\left(T_{c,n}\hat{\theta}_{c,n-1}\hat{B}_{c,n-k}^{(n-1)}\right)), \ r = l, r = 0,...k-1 \\ \tanh(\operatorname{atanh}\left(T_{c,n}\right) + \\ + \operatorname{atanh}\left(\hat{\theta}_{c,n-1}\hat{B}_{c,n-l}^{(n-1)}\hat{B}_{c,n-k}^{(n-1)}\right)), \ r = 0 \end{cases}$$

$$\hat{\theta}_{c,n} = \tanh\left(\operatorname{atanh}\left(\hat{\theta}_{c,n-1}\right) + \operatorname{atanh}\left(T_{c,n}\hat{B}_{c,n-l}^{(n-1)}\hat{B}_{c,n-k}^{(n-1)}\right)\right) \quad (16)$$

Introducing the notation $\hat{\beta}_{c,n-r}^{(n)} = \operatorname{atanh}\left(\hat{B}_{c,n-r}^{(n)}\right)$, $\hat{\vartheta}_{c,n} = \operatorname{atanh}\left(\hat{\theta}_{c,n}\right)$, $\tau_{c,n} = \operatorname{atanh}\left(T_{c,n}\right)$ we get the following algorithm:

$$\hat{\beta}_{c,n-r}^{(n)} = \begin{cases} \hat{\beta}_{c,n-r}^{(n-1)}, \ r \neq l, \ r \neq 0 \\ \hat{\beta}_{c,n-r}^{(n)} + \operatorname{atanh}(\tanh\left(\tau_{c,n}\right) \cdot \\ \cdot \tanh\left(\hat{\vartheta}_{c,n-1}\right)\tanh\left(\hat{\beta}_{c,n-k}^{(n-1)}\right)), \ r = l, r = 0,...k-1 \\ \tau_{c,n} + \operatorname{atanh}(\tanh\left(\hat{\vartheta}_{c,n-1}\right) \cdot \\ \cdot \tanh\left(\hat{\beta}_{c,n-l}^{(n-1)}\right)\tanh\left(\hat{\beta}_{c,n-k}^{(n-1)}\right)), \ r = 0 \end{cases}$$

$$\hat{\vartheta}_{c,n} = \hat{\vartheta}_{c,n-1} + \operatorname{atanh}\left(\tanh\left(\tau_{c,n}\right)\tanh\left(\hat{\beta}_{c,n-l}^{(n-1)}\right)\tanh\left(\hat{\beta}_{c,n-k}^{(n-1)}\right)\right) \quad (17)$$

A further simplification of the algorithm is based on the following approximation [10]:

$$\operatorname{atanh}\left(\prod_i \tanh\left(x_i\right)\right) \approx \approx \min_i\left(\left|x_i\right|\right)\prod_i \operatorname{sign}\left(x_i\right) \equiv F\left(x_1, x_2, ...\right) \quad (18)$$

Finally, we get

$$\hat{\beta}_{c,n-r}^{(n)} = \begin{cases} \hat{\beta}_{c,n-r}^{(n-1)}, \ r \neq l, \ r \neq 0 \\ \hat{\beta}_{c,n-r}^{(n)} + \\ + F\left(\tau_{c,n}, \hat{\vartheta}_{c,n-1}, \hat{\beta}_{c,n-k}^{(n-1)}\right), \ r = l, r = 0,...k-1 \\ \tau_{c,n} + F\left(\hat{\vartheta}_{c,n-1}, \hat{\beta}_{c,n-l}^{(n-1)}, \hat{\beta}_{c,n-k}^{(n-1)}\right), \ r = 0 \end{cases} \quad (19)$$

$$\hat{\vartheta}_{c,n} = \hat{\vartheta}_{c,n-1} + F\left(\tau_{c,n}, \hat{\beta}_{c,n-l}^{(n-1)}, \hat{\beta}_{c,n-k}^{(n-1)}\right)$$

As an approximation, you can use the following algorithm:

$$\hat{\beta}_{c,n-r}^{(n)} = \begin{cases} \hat{\beta}_{c,n-r}^{(n-1)}, \ r \neq l, \ r \neq 0 \\ \hat{\beta}_{c,n-r}^{(n)} + \\ + F\left(y_{c,n}, \hat{\vartheta}_{c,n-1}, \hat{\beta}_{c,n-k}^{(n-1)}\right), \ r = l, r = 0,...k-1 \\ y_{c,n} + F\left(\hat{\vartheta}_{c,n-1}, \hat{\beta}_{c,n-l}^{(n-1)}, \hat{\beta}_{c,n-k}^{(n-1)}\right), \ r = 0 \end{cases} \quad (20)$$

$$\hat{\vartheta}_{c,n} = \hat{\vartheta}_{c,n-1} + F\left(y_{c,n}, \hat{\beta}_{c,n-l}^{(n-1)}, \hat{\beta}_{c,n-k}^{(n-1)}\right)$$

This approximation is based on the relation, that $\tau_{c,n} \equiv y_{c,n}$.

Similarly, an estimation algorithm of $\mathbf{B}_{s,n}$ and $\theta_{s,n}$ for model (8) can be obtained. Omitting the intermediate transformations, we can write

$$\hat{\beta}_{s,n-r}^{(n)} = \begin{cases} \hat{\beta}_{s,n-r}^{(n-1)}, \ r \neq l, \ r \neq 0 \\ \hat{\beta}_{s,n-r}^{(n)} + \\ + F\left(y_{s,n}, \hat{\vartheta}_{s,n-1}, \hat{\beta}_{s,n-k}^{(n-1)}\right), \ r = l, r = 0,...k-1 \\ y_{s,n} + F\left(\hat{\vartheta}_{s,n-1}, \hat{\beta}_{s,n-l}^{(n-1)}, \hat{\beta}_{s,n-k}^{(n-1)}\right), \ r = 0 \end{cases} \quad (21)$$

$$\hat{\vartheta}_{s,n} = \hat{\vartheta}_{s,n-1} + F\left(y_{s,n}, \hat{\beta}_{s,n-l}^{(n-1)}, \hat{\beta}_{s,n-k}^{(n-1)}\right)$$

where the following notation and definitions of the variables are used: $\hat{\beta}_{s,n-r}^{(n)} \equiv \operatorname{atanh}\left(\hat{B}_{s,n-r}^{(n)}\right)$, $\hat{\vartheta}_{s,n} \equiv \operatorname{atanh}\left(\hat{\theta}_{s,n}\right)$

The following initial conditions are used in the recurrent algorithms (20) and (22):

$$\begin{aligned} &\hat{\beta}_{c,-r} = \hat{\beta}_{s,-r} = 0, \ r = 1,...,k \\ &\hat{\vartheta}_{c,0} = \hat{\vartheta}_{s,0} = 0 \end{aligned} \quad . \quad (22)$$

It should be noted that the synthesized algorithms estimate the received sequences for each quadrature component up to the sign. In this case, each of them uses averaging over the initial phase in the range of uncertainty with a width of π, i.e. equal to half the width of the general uncertainty range 2π. Therefore, these algorithms can be classified as quasi-noncoherent (subincoherent) algorithms, as opposed to noncoherent algorithms that use averaging over the entire region of phase uncertainty, or from coherent and quasi-coherent, which use the exact phase or its estimation, respectively.

The second stage of LRS synchronization is to obtain an estimate of the segment of the initial sequence using the results obtained at the first stage. Consider the following approach. As a result of the work of the algorithms (20) and (22), we have the following a posteriori distributions

$$\begin{aligned} p_{ps}\left(B_{c,n}\right) &= \frac{1}{2}\left(1 + B_{c,n}\hat{B}_{c,n}^{(n)}\right) \\ p_{ps}\left(B_{s,n}\right) &= \frac{1}{2}\left(1 + B_{s,n}\hat{B}_{s,n}^{(n)}\right) \end{aligned} \quad (23)$$

From these a posteriori distributions, equivalent likelihood functions can be obtained:

$$\Lambda\left(B_{c,n}\right)=C\frac{p_{ps}\left(B_{c,n}\right)}{p_{pr}\left(B_{c,n}\right)}=C\left(1+B_{c,n}\hat{B}_{c,n}^{(n)}\right)$$
$$\Lambda\left(B_{s,n}\right)=C\frac{p_{ps}\left(B_{s,n}\right)}{p_{pr}\left(B_{s,n}\right)}=C\left(1+B_{s,n}\hat{B}_{s,n}^{(n)}\right) \quad (24)$$

where the distribution of equiprobable binary symbols was used as a priori distribution of random binary values $B_{c.n}, B_{s,n}$:

$$p_{pr}\left(B_{c,n}\right)=p_{pr}\left(B_{s,n}\right)=\frac{1}{2}.$$

Given the definition of sequences $B_{c.n}, B_{s,n}$ in (6) and equivalent likelihood functions (25), we can write the following state and observation model:

$$\begin{cases} B_n = B_{n-l}B_{n-k} \\ \hat{B}_{c,n} = B_n\theta_{c,n} \\ \hat{B}_{s,n} = B_n\theta_{s,n} \end{cases} \quad (25)$$

As a priori distribution for random binary values, the following a posteriori distributions can be used:

$$p_{ps}\left(\theta_{c,n}\right)=\frac{1}{2}\left(1+\theta_{c,n}\hat{\theta}_{c,n-1}^{(n-1)}\right)$$
$$p_{ps}\left(\theta_{s,n}\right)=\frac{1}{2}\left(1+\theta_{s,n}\hat{\theta}_{s,n-1}^{(n-1)}\right) \quad (26)$$

After the transformation, we get

$$\hat{\beta}_{n-r}^{(n)}=\begin{cases}\hat{\beta}_{n-r}^{(n-1)}, & r\neq l,\ r\neq 0\\ \hat{\beta}_{n-l}^{(n-1)}+F\left(Z_n,\hat{\beta}_{n-k}^{(n-1)}\right), & r=l,\quad r=0,1,\ldots k-1\\ F\left(\hat{\beta}_{n-l}^{(n-1)},\hat{\beta}_{n-k}^{(n-1)}\right)+Z_n, & r=0\end{cases}$$
$$Z_n=F\left(\hat{\vartheta}_{s,n-1},\hat{\beta}_{s,n}^{(n)}\right)+F\left(\hat{\vartheta}_{c,n-1},\hat{\beta}_{c,n}^{(n)}\right) \quad (27)$$

The initial conditions for the algorithm (28) are as follows:

$$\hat{\beta}_{-r}=0,\ r=1,\ldots,k \quad (28)$$

Fig. 1 shows a block diagram of the synthesized filtration algorithm of LRS with a random delay and a random initial phase (algorithms (21), (22), (28)). It contains two identical channels for processing quadrature components (algorithms (21) and (22)) that implement the first processing stage, the blocks combining the processing results of the first stage, and the LRS filtration block that implements the second stage of LRS filtration. It should be noted that the blocks marked with a dashed line in this scheme implement a coherent filtration algorithm of LRS, which coincides with the algorithm described in [12]. The structure of the signal processing unit of the in-phase quadrature component of the first stage is shown in Fig. 2. In this scheme, the dashed line circled the blocks, estimating the sign of the phase multiplier of the quadrature component ϑ_c.

The processing unit of another quadrature component has an identical block diagram.

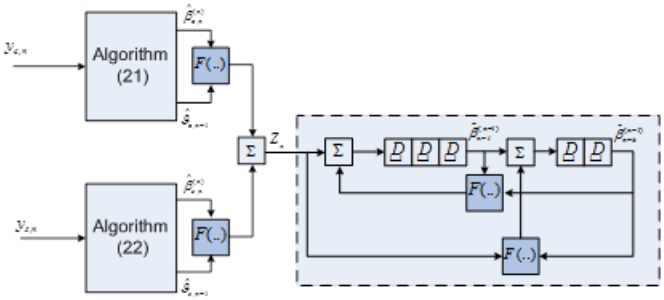


Fig. 1. Block diagram of the filtration algorithm of LRS with random delay and random initial phase

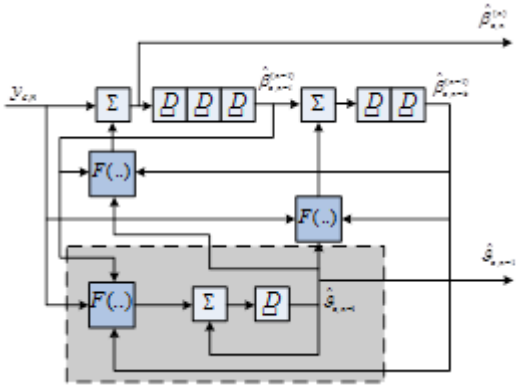


Fig. 2. Block diagram of a processing algorithm of one quadrature component (21)

Thus, the synthesized filtration algorithm of LRS with random delay and random initial phase is a two-stage combination of a coherent algorithm and algorithms for individual filtration of signals of quadrature components, which are LRS filtration algorithms with sign inversion. According to its classification, it refers to quasi-noncoherent algorithms.

IV. SIMULATION AND ANALYSIS OF RESULTS

The characteristics of the developed algorithms were investigated by statistical simulation. During the simulation, LRS with a period of $M=2^{11}-1$ (characteristic polynomial (100000000101) was used.

Figure 3 shows the dependence of correct estimation probability (Prob) of the current state of the forming shift register on the accumulation time (N) for the synthesized algorithm (21)–(23), (26) ('subincoherent'). The same figure shows the characteristics of the filtration algorithm of LRS with a known phase (the coherent algorithm — 'coherent'), described in [11; 12]. Figure 4 shows the dependence of correct estimation probability (Prob) of the current state of the forming shift register on the signal-to-noise ratio (SNR) for the same algorithms.

It can be seen from the above figures that with a sample length of 600 characters of the LRS and a signal-to-noise ratio of 8 dB, the characteristics of the proposed quasi-noncoherent algorithm practically coincide with the characteristics of the coherent algorithm. With a larger sample length, the proposed structure of the algorithm provides a greater probability of correct synchronization than a coherent algorithm. This is explained by the fact that in the proposed algorithm, in addition to evaluating the signs of the quadrature components in the first stage, the second stage of LRS filtration is used that finally leads to improved filtration quality.

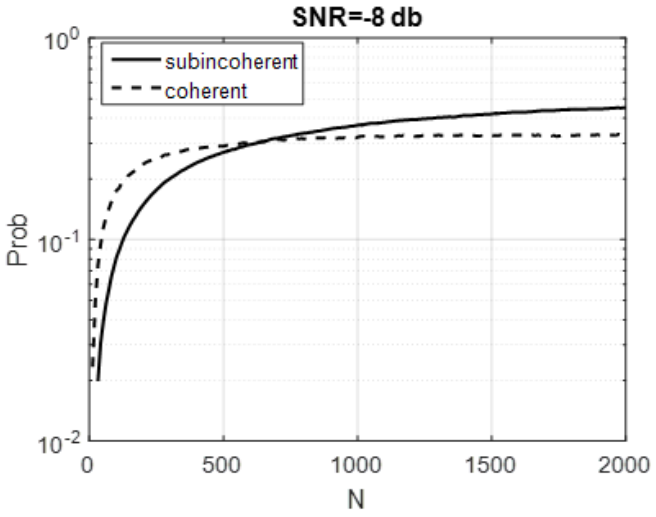


Fig. 3. Dependences of correct estimation probability (Prob) of the current state of the forming shift register on the accumulation time (N) for the synthesized algorithm (21)–(23), (26) (‘subincoherent’) and coherent algorithm from [11] (‘coherent’)

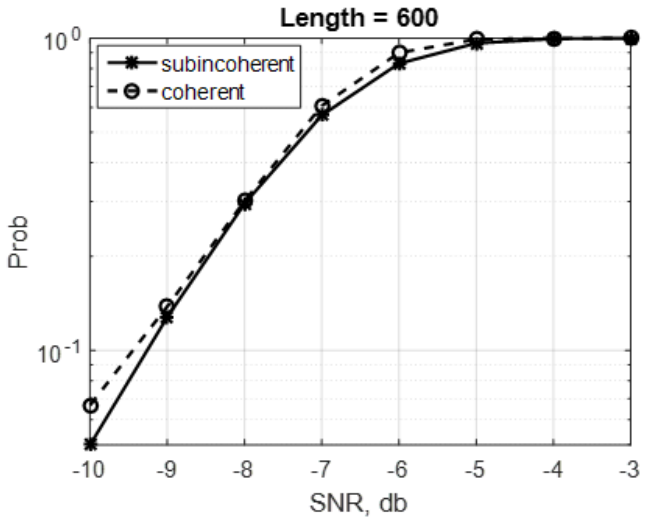


Figure 4. Dependences of correct estimation probability (Prob) of the current state of the forming shift register on the signal-to-noise ratio (SNR) of the synthesized algorithm (21)–(23), (26) (‘subincoherent’) and coherent from [11]

Thus, LRS filtration algorithms (21)–(23), (28)–(29) synthesized in this work, synchronize over any segment of LRS at a random signal phase, so they can be used, in particular, in communication systems using signal ensemble based on LRS segments [2; 6].

V. REFERENCES

[1] Теория и применение псевдослучайных сигналов / А.И. Алексеев, А. Г. Шереметев, Г.И, Тузов и др. М.: Наука, 1969. [Theory and Application of Pseudorandom Signals / A.I. Alekseev, A.G. Sheremetiev, G.I. Tuzov et al. Moscow: Nauka, 1969].

[2] Варакин Л.Е, Системы связи с шумоподобными сигналами. М.: Радио и связь, 1985. [Varakin L.E. Communication Systems with Pseudonoise Signals. Moscow: Radio and communication, 1985].

[3] Диксон Р.К. Широкополосные системы / Пер. В.И. Журавлев. М.: Связь, 1979., [Dickson R.K. Broadband Systems / Transl. from English, ed. by V.I. Zhuravlev. Moscow: Communication, 1979].

[4] Ипатов В. Широкополосные системы и кодовое разделение сигналов. М.: Техносфера, 2007. [Ipatov V. Broadband Systems and Code Division of Signals. Moscow: Tekhnosfera, 2007].

[5] Тепляков И.М., Рощин Б.В., Фомин А.И. Радиосистемы передачи информации. М.: Радио и связь, 1982. [Teplyakov I.M., Roshchin B.V., Fomin A.I. et al Radio Information Systems. Moscow: Radio and communication, 1982].

[6] Журавлев В.И. Поиск и синхронизация в широкополосных системах. М.: Радио и связь, 1986. [Zhuravlev V.I. Search and Synchronization in Broadband Systems. Moscow: Radio and communication, 1986].

[7] Петров Е.П., Алешкин Е.А. Метод подавления подобных помех с неизвестной амплитудой и задержкой сигнала в системах связи с шумоподобными сигналами // T-Comm: Телекоммуникации и транспорт. 2016. Т. 10, N 11. С. 34–39. [Petrov E.P., Aleshkin E.A. Method of suppression of signal-similar interference with unknown amplitude and time delay in communication systems with noise-shaped signals // T-Comm: Telecommunications and Transportation. 2016. Vol. 10. N 11. P. 34–39].

[8] Борисов В.И., Шестопалов В.И., Лимарев А.Е. и др. Оценка эффективности синхронизации по задержке в широкополосных системах связи с множественным доступом // Радиотехника. 2012. № 8. С. 4–16. [Borisov V.I., Shestopalov V.I., Limarev A.E. et al. Estimation of the effectiveness of synchronization by delay in broadband communication systems with multiple access // Radiotekhnika. 2012, N 8. P. 4–16].

[9] Шахтарин Б.И., Сизых В.В., Сидоркина Ю.А. и др. Синхронизация в радиосвязи и навигации. М.: Горячая линия — Телеком, 2011. [Shakhtarin B.I., Sizykh V.V., Sidorkina Yu.A. et al. Synchronization in Radio Communication and Radio Navigation. Moscow: Hotline-Telecom, 2011].

[10] Ward R. Acquisition of Pseudonoise Signals by Sequential Estimation // IEEE Transactions on Communications. 1965. Vol. 13, N 4. P. 475–483.

[11] Бакулин М.Г., Крейнделин В.Б. Метод приема псевдослучайных сигналов с неизвестной задержкой // Известия высших учебных заведений. Радиоэлектроника. 1991. Т. 34, №4. С. 47–51. [Bakulin M.G., Kreyndelin V.B. The method of receiving pseudo-random signals with an unknown delay // News of higher educational institutions. Radio Electronics. 1991. Vol. 34, N 4. P. 47–51].

[12] Бакулин М.Г., Крейнделин В.Б., Терехов А.Л. Спектральный алгоритм фильтрации линейных реккурентных последовательностей // Радиотехника. 1994. №6. С. 66–74. [Bakulin M.G., Kreyndelin V.B., Terekhov A.L., A filtration algorithm for linear recurrent sequences from a mixture with noise // Radiotekhnika. 1994. N 6. P. 66–74].

[13] Петров Е.П., Частиков А.В., Харина Н.Л. и др Алгоритм и структура устройства быстрого поиска шумоподобных сигналов // T-Comm: Телекоммуникации и транспорт. 2013. №4, С. 38–41. [Petrov E.P., Chastikov A.V., Kharina N.L. et al. The algorithm and the structure of the device to quickly search for noise-like signals // T-Comm: Telecommunications and Transport. 2013. N 4. P.. 38–41].

[14] Прозоров Д.Е. Метод последовательной оценки псевдослучайных сигналов на основе модели многосвязной цепи Маркова // Журнал радиоэлектроники: электронный журнал. 2013. №10 [Prozorov D.E. Method of sequential estimation of pseudo-random signals based on the model of a multiply connected Markov chain // Journal of Radio Electronics: Electronic Journal. 2013. N 10].

[15] Тихонов В.И. Оптимальный прием сигналов. М.: Радио и связь, 1983. [Tikhonov V.I. Optimal Reception of Signals. Moscow: Radio and communication, 1983].

[16] Бакулин М.Г. Фильтрация цифровых последовательностей с использованием моментного описания распределения вероятностей // Электросвязь. 2000. № 1. С. 24–27. [Bakulin M.G. Filtration of digital sequences using a moment description of the probability distribution // Electrical Communications. 2000. N 1. P. 24–27].

Digital Filters' Quality Improvement Technique with Given Frequency Response Requirements

V.B. Kreyndelin[1]
E.D. Grigorieva[2]
Department of Theory of Electrical Circuits
Moscow Technical University of Communication and Informatics
[1]vitkrend@gmail.com, [2]ed.grigorieva@yandex.ru

Abstract. **A method of the synthesis of digital filters based on transformation of complex response function of the analog filter prototype by replacement of a complex variable *p* by a complex variable *z* is considered. This method allows to synthesize digital IIR-filters with given requirements to frequency response that increases communication systems performance.**

Key words: **bilinear transformation method, synthesis of filters with infinite impulse response, the pre-emphasis of the frequency scale, extended bilinear transformation, precision filters.**

I. INTRODUCTION

Digital filtering is being widely used for development of various data transmission systems and devices. Often very strict requirements concerning the accuracy of implementation of the required amplitude-frequency characteristic are given to digital filters. Well-known bilinear transform is being used for digital IIR-filters synthesis [1;2].

However, for bilinear transform frequency response of the digital filter is not the same as one of analog prototype filter. Traditionally, to achieve acceptable match between filter-prototype response and one of digital filter, the pre-distortion method is used [2; 6]. Unfortunately, pre-distortions might not achieve required match between frequency response of the filter-prototype and digital filter.

Here we propose an approach to digital filters synthesis that let to avoid pre-distortions. It lets to implement precision filters with required frequency response. This approach is based on the generalization of known bilinear transform.

II. DIGITAL FILTERS SYNTHESIS, BASED ON TRADITIONAL BILINEAR TRANSFORM

Z-transform is widely used for study of discrete signals and linear discrete systems, instead of discrete Laplacian transform. It can be obtained via variables change in the discrete Laplacian transform [1;2]:

$$z = e^{pT} \tag{1}$$

where T is sampling time period.

A method of transform of transfer function of analog prototype filter $H_a(p)$ into the system function of the digital filter $H(z)$, based on mapping of the complex variable p into complex variable z, is widely used for digital filters synthesis. Main idea of this method is in the following [1; 4].

Based on (1), we can obtain the following relation between p and z:

$$p = \frac{1}{T}\ln(z) \tag{2}$$

.

It should be note, that at immediate substitution of (2) into (1), it is possible to get required transfer function of the digital filter. However, this function, unfortunately, will be transcendent function. Moreover, transfer function of implementable digital filter this function can be represented in the form of the ratio of two polynomials [5]. To meet this requirement, the equation (2) must be changed by its fractional-rational approximation (in the form of the ratio of two polynomials).

Let use known expansion of the $\ln(z)$ function [**Ошибка! Источник ссылки не найден.**]:

$$\ln(z) = 2\cdot\sum_{k=1}^{\infty}\frac{1}{2k-1}\cdot\left(\frac{z-1}{z+1}\right)^{2k-1} \tag{3}$$

Usually, only the first term is used in the expansion (3), and higher order terms are not be used. As a result, we get the following approximation, taking into account (2):

$$p = \frac{2}{T}\cdot\frac{z-1}{z+1} \tag{4}$$

Based on (4), it is not difficult to get known bilinear *z*-transform [1**Ошибка! Источник ссылки не найден.**; 4–6]:

$$p = \frac{2}{T}\cdot\frac{1-z^{-1}}{1+z^{-1}} \tag{5}$$

The system function $H(z)$ of the digital filter can be obtained via the transfer function $H_a(p)$ of the analog prototype filter, based on (5) substitution:

$$H_a(p) \overset{p=\frac{2}{T}\cdot\frac{1-z^{-1}}{1+z^{-1}}}{\rightarrow} H(z) \quad (6)$$

Transform (5) is fractional-rational first order function from z^{-1}. This transform provides unambiguous mapping of *p*-plane into *z*-plane. Digital filter is stable, if its analog prototype is stable. Optimality features of the analog prototype are kept in *z*-domain, because of unambiguous mapping of frequency axe $j\Omega$ into unit circle $e^{j\omega T}$.

Relationship between analog Ω and digital ω frequencies is actually non-linear, i.e., frequency scale is being deformed. Let find this relation, based on (5). An operator of $p=\sigma+j\Omega$ for frequency axe has the form of $p=j\Omega$, since $\sigma=0$ and $z=e^{j\omega T}$. Therefore, only frequency axe $j\Omega$ is considered in *p*-plane, and unit circle $e^{j\omega T}$ is considered in *z*-plane [4]:

$$j\Omega=p=\frac{2}{T}\cdot\left(\frac{e^{j\omega T}-1}{e^{j\omega T}+1}\right)=j\cdot\frac{2}{T}\cdot tg\frac{\omega T}{2},$$

откуда следует, что

$$\Omega=\frac{2}{T}\cdot tg\frac{\omega T}{2} \quad (7)$$

Connection between the scale of normalized frequencies of analog prototype filter Ω and the scale of normalized frequencies of the digital filter ω at ordinary bilinear *z*-transform is nonlinear (taking into account (7)). It is not possible to neglect this non-linearity, so we need to recalculate frequencies in digital domain into analog domain, and from analog domain into digital domain, to compensate frequency distortions.

Let consider proposed generalization of the traditional bilinear *z*-transform (5).

III. GENERALIZATION OF KNOWN BILINEAR TRANSFORM AND ITS APPLICATION TO DIGITAL FILTERS SYNTHESIS

To reduce non-linearity of the relationship between the scale of frequencies of digital filter ω and the scale of frequency of analog prototype filter Ω, we propose not to limit by only the first term of (3) series, but to take into account several terms for an approximation of *p* by complex variable *z*.

We have the following expression:

$$j\Omega=p=\frac{1}{T}\cdot\ln(z)\approx\frac{2}{T}\cdot\sum_{k=1}^{N}\frac{1}{2k-1}\cdot\left(\frac{z-1}{z+1}\right)^{2k-1} \quad (8)$$

Let introduce the following notation [7;8]:

$$f_N(z)=\frac{2}{T}\cdot\sum_{k=1}^{N}\frac{1}{2k-1}\cdot\left(\frac{z-1}{z+1}\right)^{2k-1} \quad (8)$$

Use the substitution:

$$z=e^{j\omega T}.$$

Then, we obtain:

$$\frac{z-1}{z+1}=j\cdot tg\frac{\omega T}{2}. \quad (10)$$

Where we have, taking into account (8):

$$\Omega=\frac{2}{T}\cdot\sum_{k=1}^{N}\frac{1}{2k-1}\cdot(-1)^{k-1}\cdot\left(tg\frac{\omega T}{2}\right)^{2k-1} \quad (11)$$

Expression is the relationship between analog Ω and digital ω frequencies for the case of generalized bilinear transform (8).

Based on (9), we get the system function of the digital filter:

$$H_a(p) \overset{p=f_N(z)}{\rightarrow} H(z) \quad (12)$$

The order of the system function $H(z)$ and the order of the digital filter are equal to $n\cdot(2\cdot N-1)$, where *n*– is the order of an analog prototype filter; N is the number of terms in (3). For traditional bilinear *z*-transform (5) we have *N* = 1.

For example, for $N=3$ (3-d order approximation) we have:

$$j\Omega=p=\frac{1}{T}\cdot\ln(z)\approx\frac{2}{T}\cdot\sum_{k=1}^{3}\frac{1}{2k-1}\cdot\left(\frac{z-1}{z+1}\right)^{2k-1}= \quad (13)$$

$$=\frac{2}{T}\cdot\left(\left(\frac{z-1}{z+1}\right)+\frac{1}{3}\cdot\left(\frac{z-1}{z+1}\right)^{3}+\frac{1}{5}\cdot\left(\frac{z-1}{z+1}\right)^{5}\right)$$

After substitution $z=e^{j\omega T}$ into (13) we get, using (10):

$$j\Omega=p=\frac{2}{T}\cdot\left(j\cdot tg\frac{\omega T}{2}+\frac{1}{3}\cdot\left(j\cdot tg\frac{\omega T}{2}\right)^{3}+\frac{1}{5}\cdot\left(j\cdot tg\frac{\omega T}{2}\right)^{5}\right), \quad (29)$$

Where for *N = 3*, we obtain the relation between scales of frequencies of analog prototype filter Ω and digital filter ω:

$$\Omega=\frac{2}{T}\cdot\left(tg\frac{\omega T}{2}-\frac{1}{3}\cdot\left(tg\frac{\omega T}{2}\right)^{3}+\frac{1}{5}\cdot\left(tg\frac{\omega T}{2}\right)^{5}\right) \quad (15)$$

For $N=2$ (2-nd order approximation) we have:

$$\Omega=\frac{2}{T}\cdot\left(tg\frac{\omega T}{2}-\frac{1}{3}\cdot\left(tg\frac{\omega T}{2}\right)^{3}\right) \quad (16)$$

For large N relationship between scales of frequencies of both analog and digital filters converges to linear dependency in the range of normalized frequencies in filter band pass [8].

IV. EXAMPLES OF THE DIGITAL FILTERS SYNTHESIS

To study of proposed method, we have provided the synthesis of two low-pass Butterworth digital filters, those are met the requirements from [**Ошибка! Источник ссылки не найден.**]:

- boundary frequency of pass band is $F_{pass}=1$ Hz;

- boundary frequency of non-pass band is $F_{stop} = 2,414$ Hz;
- sampling frequency is $F_s = 7$ Hz; sampling time period is $T = \frac{1}{F_s}$;
- non-uniformity of attenuation in pass band $\Delta A = 0,5$ dB;
- minimal accepted attenuation $A_{\min} = 20$ dB.

Initially, let calculate the system function of the digital filter $H_1(z)$, using ordinary bilinear z-transform (5), that matches to the case of $N = 1$, than let calculate the system function $H_3(z)$, using proposed transform in the case of $N = 3$. In both cases an analog prototype filter is the same. Calculation is the following.

1. Bound frequencies Ω_i of the analog prototype filter (without predistortions) have the following values, those are the same with normalized vs. bound frequency Ω_{pass} frequencies:

$$\Omega_p = \frac{\Omega_{pass}}{\Omega_{pass}} = 1; \quad \Omega_s = \frac{\Omega_{stop}}{\Omega_{pass}} = 3,4 .$$

2. Calculation of the order of analog prototype filter [2; 6]:

$$n = \frac{lg\left(\frac{10^{0,1\cdot A_{\min}} - 1}{10^{0,1\cdot \Delta A} - 1}\right)}{2\cdot lg(\Omega_s)} = 2,453 .$$

Since filter order must be integer, we get $n = 3$.

3. Poles $p(k)$ of analog prototype filter **are to be calculated, $k = 1,\ldots n$; n is the number of poles of** analog prototype filter, that is the same as its order **[2Ошибка! Источник ссылки не найден.; 10;11]:**

$$p(k) = \frac{1}{\sqrt[n]{10^{0,1\cdot\Delta A} - 1}} \cdot \left(-sin\left(\frac{\pi\cdot(2k-1)}{2n}\right) + j\cdot cos\left(\frac{\pi\cdot(2k-1)}{2n}\right)\right) \quad (17)$$

Based on (17)**, we calculate the following values of poles of** analog prototype filter**:**

$$p(1) = -0,71 + j\cdot 1,23; \; p(2) = -1,42;$$
$$p(3) = -0,71 - j\cdot 1,23.$$

4. Transfer function of analog prototype filter **is the following:**

$$H_a(p) = \frac{1}{\sqrt{10^{0,1\cdot\Delta A} - 1}\cdot(p - p(1))\cdot(p - p(2))\cdot(p - p(3))} =$$
$$= 2,863\cdot\frac{1}{(p + 1,42)}\cdot\frac{1}{\left(p^2 + 1,419\cdot p + 2,016\right)} \quad \textbf{(18)}$$

Frequency response of analog prototype filter is given below:

$$H_a(f) = 2,863\cdot\frac{1}{j\frac{f}{f_{pass}} + 1,42}\cdot\frac{1}{\left(j\frac{f}{f_{pass}}\right)^2 + 1,419\cdot j\frac{f}{f_{pass}} + 2,016} \quad \textbf{(19)}$$

5. For the case of ordinary bilinear *z*-transform, the substitution of (5) into (18) **is being done, and after some simple transformations we get the following expression for the system function $H_1(z)$** of digital filter:

$$H_a(p) \overset{p = f_1(z)}{\rightarrow} H_1(z),$$

$$H_1(z) = 0,095\cdot\left(\frac{z+1}{z-0,17}\right)\cdot\left(\frac{z^2 + 2z + 1}{z^2 - 0,448z + 0,359}\right). \quad (20)$$

Frequency response of digital filter, using bilinear *z*-transform, is given below:

$$H_1(f) = 0,095\cdot\left(\frac{e^{j\cdot 2\pi\cdot f\cdot T} + 1}{e^{j\cdot 2\pi\cdot f\cdot T} - 0,17}\right)\cdot$$
$$\cdot\left(\frac{e^{j\cdot 2\pi\cdot f\cdot T\cdot 2} + 2\cdot e^{j\cdot 2\pi\cdot f\cdot T} + 1}{e^{j\cdot 2\pi\cdot f\cdot T\cdot 2} - 0,448\cdot e^{j\cdot 2\pi\cdot f\cdot T} + 0,359}\right) \quad (21)$$

Expression (21) is given by the substitution of (1) into (20), taking into account normalization via sampling frequency.

The synthesis of the digital filter, based on generalized bilinear *z*-transform ($N = 3$) is similar. Using items 1…4, we obtain the same transfer function of analog prototype filter(18). But the next stage is different.

Let use the following transform

$$p = f_3(z) = \frac{2}{T}\cdot\left(\left(\frac{z-1}{z+1}\right) + \frac{1}{3}\cdot\left(\frac{z-1}{z+1}\right)^3 + \frac{1}{5}\cdot\left(\frac{z-1}{z+1}\right)^5\right) \quad (22)$$

After substitution (22) into (18) we obtain:

$$H_a(p) \overset{p = f_3(z)}{\rightarrow} H_3(z),$$

We get system function $H_3(z)$ of digital filter (3-d order generalization):

$$H_3(z) = 0,033\cdot\frac{z^5 + 5z^4 + 10z^3 + 10z^2 + 5z + 1}{z^5 + 2,25z^4 + 4,5z^3 + 1,5z^2 + 0,72z - 0,4}\cdot$$

$$\cdot \frac{\left(z^{10}+10z^{9}+45z^{8}+120z^{7}+210z^{6}+\right.}{\left.+252z^{5}+210z^{4}+120z^{3}+45z^{2}+10z+1\right)} \qquad (23)$$
$$\frac{\left(z^{10}+4z^{9}+13z^{8}+21z^{7}+28,9z^{6}+\right.}{\left.+19,5z^{5}+15,3z^{4}+5,8z^{3}+3,8z^{2}+0,8z+0,4\right)}$$

6. Polynomials of 2-nd degree with real coefficients are being generated, based on conjugated zeros and conjugated poles. These polynomials are further used for both nominators and denominators of chains.

Frequency response of the digital filter, based on generalized *z*-transform, can be obtained by the exchange of $z \to e^{j\cdot 2\pi\cdot f\cdot T}$ in (23).

Let compare frequency response of received digital filters between each other and with frequency response of analog prototype filter with 3-d order.

Frequency responses of attenuation are shown by the Fig.1 for analog prototype filter and synthesized digital filters:

- $\left|A1(f)\right|$ is for digital filter, based on an ordinary bilinear z-transform;
- $\left|A3(f)\right|$ is for digital filter, based on generalized bilinear z-transform;
- $\left|Aa(f)\right|$ is for analog prototype filter of 3-d order.

Filter cutoff frequency of the filter, using an ordinary bilinear *z*-transform, is equal to $F_{cut} = 1{,}37$. Filter cutoff frequency of analog prototype filter, and obtained digital filter (using 3-d order bilinear z-transform) are same and it is equal to $F_{cut} = 1{,}42$.

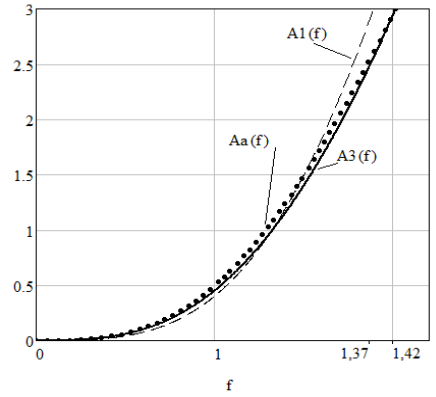


Fig. 1. Curves of deviation of filters in the passband

From Fig. 1 one can see that frequency response of the digital filter, based on generalized bilinear *z*-transform, is quite close to one of analog prototype filter in working frequency domain. It lets to avoid pre-distortions and to improve accuracy of frequency response of digital filter.

V. CONCLUSION

Based on described results, one can see that implementation of proposed generalized bilinear *z*-transform lets to get digital filter with frequency response very close to ones of analog prototype filter. It lets avoid predistortions those are required for traditional bilinear *z*-transform.

Increase of the order of digital filter (and its complexity increase) is the disadvantage of proposed method. Practically, a tradeoff should be found between frequency response accuracy and filter complexity.

VI. REFERENCES

1. Оппенгейм А., Шафер Р. Цифровая обработка сигналов. 2-е изд. М.: Техносфера, 2007, [Oppengame A., Shafer R. Digital signal processing. 2nd ed. Moscow: Technosphera, 2007].
2. Соболев В.Н. Теория электрических цепей. М.: Горячая линия — Телеком, 2014. [Sobolev V.N. Theory of electric circuits. Moscow: Goryachaya Linia — Telecom, 2014].
3. Фихтенгольц Г. М. Курс дифференциального и интегрального исчисления. 8-е изд. М.: Физматлит, 2003. Т 1. [Fihtengolts G.M. Course of differential and integral calculus. 8th ed. Moscow: Fizmatlit, 2003. Vol. I].
4. Баскаков С.И. Радиотехнические цепи и сигналы. М.: Высшая школа, 2000. [Baskakov S.I. Radiotechnical circuits and signals. Moscow: Visshaya shkola, 2000].
5. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. М.: Мир, 1978. [Rabiner L. R., Gold V. Theory and application of digital signal processing. Moscow: Mir, 1978].
6. Смит С. Научно-техническое руководство по цифровой обработке сигналов. М.: Додэка, 2008. [Smit S. Digital signal processing. Practical manual for engineers and scientists. Moscow: Dodeka XXI, 2008].
7. КрейнделинВ.Б., Григорьева Е.Д. Развитие метода билинейного преобразования для синтеза цифро- вых фильтров // Международная научно-техническая конференция “INTERMATIC-2017”. Ноябрь 2017. М., 2017. С. 1183–1185. [Kreyndelin V.B., Grigorieva E.D. Development of the method of bilinear transform for digital filters synthesis. Materials of International Conference «INTERMATIC–2017». Fundamental problems of radioelectronic devices development. November, 2017. Moscow, 2017. P. 1183-1185.
8. Крейнделин В.Б., Григорьева Е.Д. Методика повышения показателей качества цифровых фильтров в системах связи // Труды Международной научно-технической конференции «Телекоммуникационные и вычислительные системы – 2017». М., 2017. С. 162–166. [Kreyndelin V.B., Grigorieva E.D. Method of quality increase of digital filters in communication systems // Materials of International Conference “Telecommunication and computational systems – 2017’. Moscow, 2017. P. 162–166].
9. Kwaha B.J., Kolawole E.A., Batu A.M. The design and implementation of a digital infinite impulse response (IIR) lowpass Butterworth filter – A comparison of Matlab and Bilinear transformation methods // Indian Journal of Science and Technology. 2011. Vol. 4, N 4 P. 451–455.
10. Марченко А.П. Частотные фильтры: пассивные, активные и цифровые. М.: Горячая линия — Телеком, 2017. [Marchenko A.P. Digital filters: passive, active and digital. Moscow: Goryachaya Linia — Telecom, 2017].
11. Смирнов Н.И., Фриск В.В, Теория электрических цепей. М.: Горячая линия — Телеком, 2018. [Smirnov N.I., Frisk V.V. Theory of electric circuits: conspect of Lectures. Moscow: Goryachaya Linia — Telecom, 2018].

Timing Quality for New Generation Networks

S.V. Melnik[1]
Moscow Technical University of Communications and Informatics;
Central Science Research Telecommunication Institute

N.I. Smirnov[2]
Moscow Technical University of Communications and Informatics
Moscow, Russia
[1]svmelnik@bk.ru, [2]smirnov.ni@yandex.ru

***Abstract*. New generation networks deployment has one of the perspective directions the concept of the Internet of Everything (IoE) [3]. Rendering services in the IoE networks require providing a stable time scale with an accuracy of definition not less than 1 ms, and at operation of applications of the tactile Internet not less than 1 microsec. As in the current legislation are absent obligatory requirements to the accuracy of definition of time in the IoE networks, the mechanism of additional certification is required. Additional certification of production and services is the mechanism for check of correctness of the made decisions and current state of projects on production of new hi-tech production and services. Existence of the certificate of Digital Innovations system can be criterion of providing a level of quality and necessary amount of works by results of the project.**

***Key words:* digital economy, certification, type approval.**

I. INTRODUCTION

5G mobile communication networks, technologies of the Industrial Internet of Things (IIoT), Tactile Internet of Things (TIoT), Software Defined networks (SDN) with Network Functions Virtualization (NFV), and Big Data systems have to become a technological basis for implementation of the "Digital Economy of the Russian Federation" program. All these technologies unite Internet of Everything (IoE) concept. For sure, work of the technologies connected with high reliability and providing the minimum delay at distribution of signals tough standards to providing a timeline are required. So, in networks of the IIoT, at remote control of the infrastructure equipment and objects, at the organization of the pilotless movement of vehicles at speeds up to 60 km/h, ensuring accuracy of a timeline 1 ms is required. Ensuring management of the pilotless movement of vehicles at speeds more than 80 km/h, controls of unmanned aerial vehicles and railway objects require providing a time scale with an accuracy of 1 microsec. There is one more aspect of use of high-precision devices for ensuring exact time is a TIoT. In such appendices as a telemedicine, in case of carrying out remote surgical influence, the accuracy of time of 1 microsec on condition of the maximum delay of transfer of a package of 1 ms is required.

If to consider all these technological directions in the light of the operating regulatory legal base, there are collisions. It is connected with that the main existing normative legal documents, such as the law No. 126-FZ "About communication" of 07.07.2003 and the law No. 184-FZ of 27.12.2002 "On technical regulation" are accepted more than 15 years ago. Then nobody could assume that in the near future technologies of pilotless driving of cars will begin to take root that the large cities will become almost completely covered not only networks of mobile communication with high-speed data transmission, but also full video surveillance with possibility of a face recognition. Become objects of interaction of communication networks not only natural and legal entities, but also inanimate objects. Information processing becomes spaced (Cloudy technologies – Cloud computing) with part of functions which are transferred to local networks (Foggy calculations – Fog computing).

It turns out that presentation of obligatory requirements to means of communication at the level of technical characteristics (Network Performance) isn't enough for ensuring integrity, stability and safety of networks of new generation. It is necessary to impose additional requirements which have to concern not only means of communication, but also services of a digital network on level of service (Quality of Service), and also at the level of external perception of communication services (Quality of Experience).

It is clear that we are expected by deep processing of current laws and bylaws, but this process can take a lot of time, and it will always lag behind the real requirements connected with the advancing technological development. As the adopted laws and bylaws have to rely on the operating accepted norms and standards and can't advance standardization.

The collision consists that in the absence of the existing obligatory requirements, laws and bylaws, it is necessary to realize the most part of the "Digital Economy of the Russian Federation" program and to be sure that all invested funds did not go to waste. It is about means from the state budget and the reporting for their expenditure are regulated by the current legislation.

Certainly, implementation of the "Digital Economy of the Russian Federation" program requires scientific approach. First of all it is necessary to develop the concept of creation of a

modern national network of telecommunication within implementation of the program of development of Digital economy of the Russian Federation. This concept differs from the general program of development of communication networks of the Russian Federation in that it concerns the certain direction which is the "Digital Economy of the Russian Federation" program with public financing. The mechanism of public financing means other level of control and supervision of work than financing from commercial structures. Therefore, measures of installation of requirements to networks, means and communication services on the basis of which the decision on their suitability for implementation of the "Digital Economy of the Russian Federation" program are made have to be developed.

Before emergence of the approved international standards to networks of mobile communication 5G, IIoT, TIoT, SDN/NFV, BigData and to other technologies of new generation, it is impossible to accept requirements for obligatory certification of production and services. Meanwhile, skilled and commercial operation of the corresponding networks and constructions is necessary to avoid possible technological lag. It is necessary to provide the mechanism of additional certification of production and services which will extend only on the objects planned for use at implementation of the "Digital Economy of the Russian Federation" program.

Additional certification allows to impose requirements to again developed products and services before emergence of obligatory requirements.

For assistance of implementation of the "Digital Economy of the Russian Federation" program approved by the order of the Government of the Russian Federation of July 28, 2017 No. 1632-r [2] under the auspices of the International academy of communication system of certification "Digital Innovations" is created.

Additional certification allows not only to estimate efficiency of the measures undertaken for implementation of the Program but also to certify quality of synchronization of time.

Additional certification of production and services is the mechanism for check of correctness of the made decisions and current state of projects on production of new hi-tech production and services.

Existence of the certificate of Digital Innovations system can be a criterion of performance of quality and amount of works by results of the project, and also qualities of definition of time for the IoE networks.

It is expedient to use the certificate of Digital Innovations system, as means of confirmation of the characteristics of production which are not regulated by the existing obligatory requirements [3].

Requirements at additional certification can be both softer, and more rigid than at obligatory.

Certification in Digital Innovations system is more flexible and universal tool than obligatory confirmation of compliance. Certification in Digital Innovations system is irreplaceable at implementation of the concept of creation of the new environment for production and services within the program of digital economy of the Russian Federation [4–7].

II. *THE CENTRALIZED SCHEME OF TIME SYNCHRONIZATION*

The problem of time synchronization can be solved by accession to the remote server of time having an exit to the source synchronized with the State standard, for example, of GNSS GLONASS. The centralized architecture of time synchronization in a capillary network is shown in Figure 1. However, such method of the decision determines that it is necessary to guarantee that when receiving signals of time, the distortion involving increase in probability of a mistake will not arise in each network node.

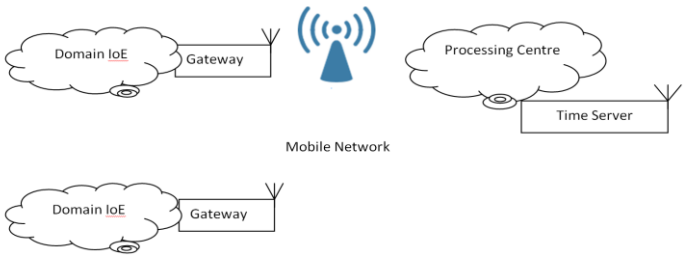


Fig. 1. The centralized scheme of time synchronization

III. *THE DISTRIBUTED SCHEME OF TIME SYNCHRONIZATION*

In a network with a large number of IoE domains, some of which can be on considerable removal from the center of information processing, the centralized scheme of providing the uniform temporary is not effective.

By analogy with the methods stated in [6; 7], it is possible to offer the distributed architecture for time synchronization in a IoE network (Fig. 2).

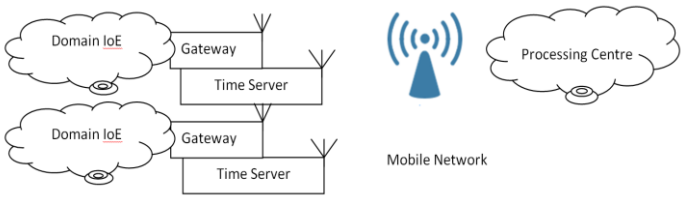


Fig. 2. The distributed scheme of time synchronization

In the distributed scheme of ensuring unity and accuracy of a time scale in a IoE network, we use the compact server of time with the GPS/GLONASS receiver connected to a lock of the IoE domain. At such version of the solution of a problem of time synchronization, we can guarantee the accuracy of an order of units of nanoseconds. This accuracy of temporary synchronization has enough for the solution of any technological task in which interests the capillary network is under construction. This mechanism, especially effective, will be when using in systems of a dual purpose which impose increased requirements on reliability.

IV. *TIMING QUALITY DETECTION AND USING FAST BLOCKCHAIN ALGORITHM*

In system of certification "Digital Innovations" essentially new technique of certification of quality of time synchronization is developed. This technique is based on the block chain algorithm when processing the temporary tags generated by the equipment by STC KOMSET. When the quality control mechanism of synchronization initializing, each tag which is generated by the server of time by STC KOMSET is added to the block in the field of Time. Blocks are processed in system of processing of inquiries, a control system of servers and on transit hosts of the estimated company. Thus, it is formed Blockchain, excluding falsification of data on determination of quality of

synchronization. In the measurements scheme we use new kind of blockchain algorithm in Standard Blockchain algorithm using timing from Internet clock with accuracy 0,1 sec. Usually Blockchain looks like in the Fig. 3.

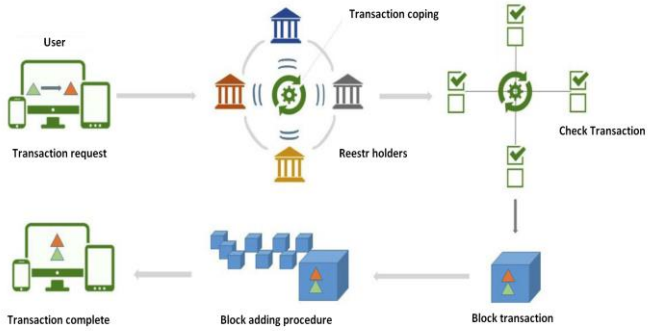


Fig. 3. Standard Blockchain algorithm

At this algorithm using following blocks (see Table).

Block format for Standard Blockchain algorithm

Field	Description	Size
BLOCK_ID	Block Identifier	4 bytes
TIME	Time	4 bytes
USER_ID	User Block Identifier	5bytes
LEVEL	Miner level	2bytes
SIGN	Signature (TYPE, BLOCK_ID, PREV_BLOCK_HASH, TIME, USER_ID, LEVEL, MRKL_ROOT	128 - 512bytes
TRANSACTIONS	Transactions	Less 3Mb

Cicling request choosing blocks with new information each 120 sec. and modified the chain. In this case we used prescision timing with accuracy 1 ms from STC KOMSET high prescision time servers, we can make blockcain cicling 100 times faster. There is new Fast Blockchain Algorithm at the Figure 4.

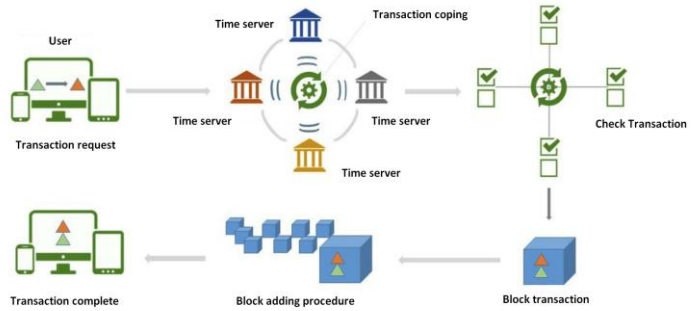


Fig. 4. Fast Blockchain algorithm

Fast Blockchain algorithm can provide more efficient blockchain processing. As the part of tegs, we use time lables from the STC KOMSET servers as was mentioned before.

The certificate of conformity is issued as a result of check in system of certification “Digital Innovations”.

V. CERTIFICATION OF TIMING QUALITY IN CERTIFICATION SYSTEM “DIGITAL INNOVATIONS”

Certification system “Digital Innovations“ purposes are:

- support of the environment of production in which data in a digital form are a key factor of production in all spheres of social and economic activity;
- support of conditions of development of high-tech businesses;
- assistance to prevention of emergence of obstacles and restrictions of development of high-tech businesses;
- support of increase of competitiveness of the russian Federation in the global market.

These purposes require creation of conditions of close interaction of the main key objects:

- the markets and branches of economy where interaction of suppliers and consumers of production and services is carried out;
- platforms and technologies where means for development of the markets and branches of economy are formed;
- creating conditions for development of platforms and technologies and effective interaction of consumers of production and services.

Areas of certification of objects of through digital technologies are:

- systems and objects of the industrial Internet of things;
- systems and objects of the tactile and touch Internet of things;
- components of a robotics and sensorics;
- means and services of technologies of wireless communications;
- means and services of virtual and augmented reality technologies ;
- systems for work with big data;
- systems and objects of neurotechnologies and artificial intelligence;
- systems of the distributed register;
- objects of quantum technologies;
- new production technologies.

Change of the list of such technologies on a measure emergence and developments of new technologies is provided.

Certification of the services directed on support of information infrastructure:

- certification of services of communication networks which provide requirements of economy on collecting and data transmission of the state, business and citizens taking into account the

technical requirements imposed by digital technologies;

- certification of services of data-processing centers which provides granting to the state, to business and citizens of available, steady, safe and economically effective services in storage and data processing on conditions and allows including to export services in storage and data processing;
- certification of services of digital platforms of work with data for ensuring requirements of the power, business and citizens;
- certification of services of system of collecting, the processing, storage and providing spatial data to consumers providing needs of the state, business and citizens for actual and reliable information about spatial objects.

Prevention of calls and threats:

- the problem of ensuring human rights in the digital world;
- the threats of the personality, to business and the state connected with tendencies to creation of difficult hierarchical information and telecommunication systems;
- building of opportunities of external information and technical impact on information infrastructure;
- growth in computer crime.

VI. CONCLUSION

In the IoE networks it is necessary to use a uniform time scale.

Timeline accuracy in the IoE networks has to be 1 ms depending on the transmitted data.

It is expedient to use the distributed mechanism of time synchronization and frequency realized on servers of STC KOMSET to the IoE networks.

The certification mechanism based on algorithm Blockchain is used to synchronize time and frequency.

The certificated quality of time synchronization is issued In system of certification “Digital Innovations”. The technique of an assessment is based on Blockchain algorithm that renders it unsuitable for falsification.

VII. REFERENCES

[1] Указ Президента РФ от 09.05.2017 N 203 «О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы» [The decree of the president of the Russian Federation “About strategy of development of information society in the Russian Federation for 2017–2030”]. http://www.consultant.ru/document/cons_doc_LAW_216363/.

[2] Распоряжение Правительства РФ от 28.07.2017 N 1632-р «Об утверждении программы "Цифровая экономика Российской Федерации"» [The order of the Government of the Russian Federation of July 28, 2017 No. 1632-r “The Digital Economy of the Russian Federation”]. URL: http://www.consultant.ru/document/cons_doc_LAW_221756/.

[3] Оситис А.П., Мельник С.В., Петрова Е.Н., Смирнов Н.И. Применение добровольной сертификации в целях содействия реализации программы «Цифровая экономика Российской Федерации» // Труды Международной научно-технической конференции «Телекоммуникационные и вычислительные системы –2017». М., 2017. С. 117–122. [Ositis A.P., Melnik S.V., Petrova E.N., Smirnov N. I. Application of voluntary certification for assistance of implementation of the “Digital economy of the Russian Federation” program // International Forum of Informatization (IFI) 2017. International congress “Communication technologies and networks”: Collection of works. Moscow, 2017. P. 117–122.

[4] Мельник С.В., Петрова Е.Н., Смирнов Н.И. Технология мобильной связи четвертого поколения LTE. Основные особенности и перспективы внедрения в России // T-Comm. Телекоммуникации и транспорт. 2012. Т. 4, № 7. С. 89–90. [Melnik S.V., Petrova E.N., Smirnov N.I. Fourth-generation mobile communications technology of LTE The main features and prospects of introduction in Russia // T-Comm: Telecommunications and transport. 2010. T. 4 N 7. P. 89–90].

[5] Аджемов А.С., Смирнов Н.И., Мельник С.В. и др. Пути обеспечения единого точного времени на сетях электросвязи России с использованием ГЛОНАСС // Системы синхронизации, формирования и обработки сигналов. 2013. Т. 4. № 3. С. 150–153. [Adzhemov A.S., Smirnov N. I., Melnik S.V. et al. Ways of ensuring uniform exact time on networks of telecommunication of Russia with use of GLONASS // System of synchronization, formation and processing of signals. 2013. T. 4. N 3. P. 150–153].

[6] Мишенков С.Л., Мельник С.В., Петрова Е.Н. и др. Обеспечение точного времени для сектей связи с использованием возможнеостей ГНСС ГЛОНАСС или ГЛОНАСС/GPS. Моделирование сетей мобильной связи нового поколения // T-Comm. Телекоммуникации и транспорт. 2012. Т. 6. № 9. Р. 102–103. [Mishenkov S. L., Melnik S.V., Petrova E.N., Smirnov N. I. ensuring exact time for communication networks with use of opportunities of GNSS GLONASS or GLONASS/GPS. modeling of networks of mobile communication of new generation // T-Comm: Telecommunications and transport. 2012. T. 6, N 9. P. 102–103].

[7] Мишенков С.Л., Смирнов Н.И., Мельник С.В. и др. Перспективы использования модернизируемой глобальной навигационной спутниковой системы ГЛОНАСС в качестве многофункциональной спутниковой системы // T-Comm. Телекоммуникации и транспорт. 2011. Т. 5, № 9. С. 106–109. [Mishenkov S. L., Smirnov N. I., Melnik S.V. et al. Prospects of use of the modernized global navigation satellite GLONASS system as multipurpose satellite system // T-Comm: Telecommunications and transport. 2011. T. 5, N 9. P. 106–109ъ.

Linear Iterative Demodulation Algorithm for MIMO Systems with Large Number of Antennas

D.Yu. Pankratov[1],
Department of Systems and Networks of Radio Communications and Radio Broadcasting

A.G. Stepanova Department of Theory of Electrical Circuits
Moscow Technical University of Communications and Informatics
[1]dpankr@mail.ru; [2]ag.otc@rambler.ru

***Abstract*. The use of Massive MIMO technology (i.e. MIMO technology with large number of antennas) is planned in 5G networks for significant capacity increasing. However, on the way to practical application of Massive MIMO technology in 5G systems, there are many problems, in particular, problems of digital signal processing algorithms synthesis on the receiving side, therefore, the developing task of demodulation algorithms with good quality characteristics and low computational complexity is very important.**

In this paper we discuss linear demodulation algorithms for MIMO systems: Zero-Forcing (ZF) algorithm, minimum mean square error (MMSE) algorithm and Chebyshev iterative algorithms. For these algorithms BER performance characteristics are obtained.

***Key words:* 5G, Massive MIMO, demodulation algorithms, Chebyshev iterative algorithms, BER performance.**

To significantly increase the capacity of modern radio communication systems, Multiple Input Multiple Output technology (MIMO) is widely used [1; 2], which uses multiple transmitting and multiple receiving antennas.

In 5G networks, in order to increase the capacity, it is planned to use Massive MIMO technology (i.e. MIMO technology with a large number of antennas). This technology allows to obtain high spectral efficiency and high energy efficiency of the communication system. However, on the way to practical application of Massive MIMO technology in 5G systems, there are many problems, in particular, the problem of digital signal processing algorithms synthesis, therefore, the task of demodulation algorithms developing with good quality characteristics and low computational complexity is very important.

In this paper we discuss linear demodulation algorithms for MIMO systems: Zero-Forcing (ZF) algorithm, minimum mean square error (MMSE) algorithm and the Chebyshev iterative algorithms. The noise immunity curves of these algorithms for the MIMO system with 4 transmitting and 4 receiving antennas were obtained (bit error ratio (BER) versus signal-to-noise ratio (SNR)).

Figure 1 shows the MIMO block diagram. The model of the signal at the input of the demodulator is the following [2]:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{n}, \qquad (1)$$

where **y** is received signals vector of $M \times 1$ dimension; ***H*** is complex matrix of MIMO radio channel of $M \times M$ dimension; ***s*** is vector of transmitted information symbols of $M \times 1$ dimension; ***n*** is Gaussian random vector of noise of $M \times 1$ dimension. Elements ***hij*** of MIMO channel matrix ***H*** represent the complex transmission coefficients from the *j*-th transmitting antenna to the *i*-th receiving antenna.

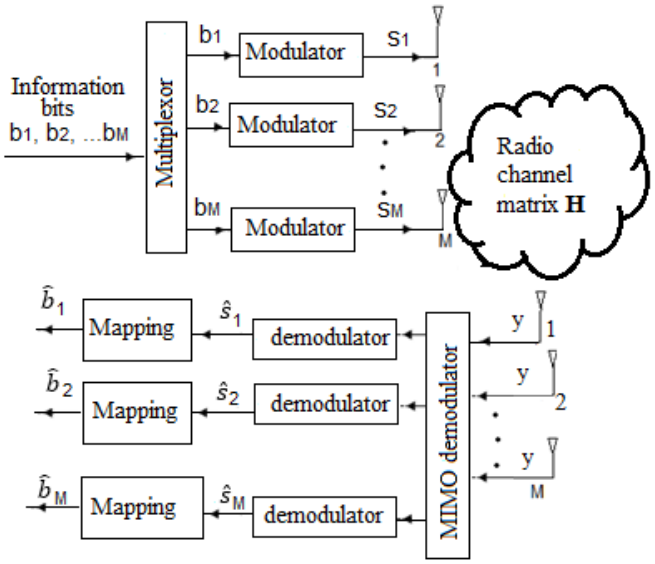


Fig. 1. MIMO system block diagram

The algorithm for MIMO system simulation has several steps:

- generation of bit vector **b** and information symbol vector **S** for all transmitting antennas, generation of MIMO channel matrix ***H***, consisting of complex coefficients ***hij***;
- generation of complex Gaussian noise vector ***n***;
- obtaining of signal and noise mixture **y**;
- demodulation of signal **y** – obtaining of information symbol vector estimation $\hat{\mathbf{S}}$ and corresponding vector $\hat{\mathbf{b}}$;

- comparing of $\hat{\mathbf{b}}$ и $\mathbf{b}$ vectors and errors detection;
- bit error ratio (BER) calculation for given signal-to-noise ratio (SNR).

These steps are performed for a given number of experiments and SNR values to achieve average BER performance for different SNR values. A detailed description of this algorithm can be found in paper [3].

In MIMO systems ZF (Zero Forcing) algorithm can be used for demodulation. The ZF estimation of information symbols can be found as follows [2; 4]:

$$\hat{\mathbf{s}}^{ZF} = \arg\min_{s \in C^I} \|\mathbf{y} - \mathbf{Hs}\|^2 = (\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'\mathbf{y} \quad (2)$$

where C^I is *I*-dimensional continuous complex space; $\hat{\mathbf{s}}$ is estimation of received information symbol vector $\mathbf{S}$; $\boldsymbol{y}$ is received signals vector of $M \times 1$ dimension; $\mathbf{H}$ is complex matrix of MIMO radio channel of $M \times M$ dimension; $(\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'$ is pseudoinverse matrix with respect to the channel matrix $\mathbf{H}$; $\mathbf{H}'$ is Hermitian conjugate matrix with respect to the channel matrix $\mathbf{H}$. From (2) it can be seen that Zero Forcing algorithm of demodulation does not take into account the presence of observation noise $\mathbf{n}$, resulting in a significant loss in noise immunity.

Let us now consider minimum mean-square error (MMSE) demodulator. The estimation $\hat{\boldsymbol{\theta}}^{\mathbf{MMSE}}$, optimal by MMSE criteria, can be found as follows [2, 4]:

$$\hat{\boldsymbol{\theta}}^{\mathbf{MMSE}} = [\mathbf{H}'\mathbf{H} + 2\boldsymbol{\sigma}_{\boldsymbol{\eta}}^2 \cdot \mathbf{1}]^{-1}\mathbf{H}'\mathbf{y}, \quad (3)$$

where $2\boldsymbol{\sigma}_{\boldsymbol{\eta}}^2$ is complex noise dispersion; $\mathbf{1}$ is identity matrix.

Algorithm (3) takes into account the presence of noise in radio channel and, therefore, it has a higher noise immunity compared to ZF algorithm. This can be seen in Figure 2, which shows the dependences of the average BER for ZF and MMSE algorithms for different SNR values. Simulation was performed for MIMO system with 4 transmitting and 4 receiving antennas for 10,000 experiments.

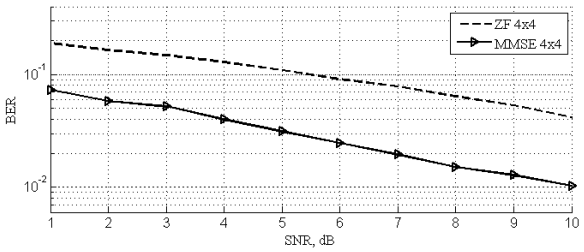


Fig. 2. BER performance for ZF and MMSE demodulation algorithms

To implement the demodulation algorithms considered, it is necessary to calculate the inverse matrix, which is a very difficult task in real-time conditions for systems with large number of antennas. However, it is impossible to calculate in advance the inverse of MIMO channel matrix, which depends on the complex transmission coefficients of radio channel, since matrix $\mathbf{H}$ varies randomly and should be estimated in real-time conditions. Thus, in case of large number of transmitting and receiving antennas, the use of these methods in practice becomes difficult for implementation.

One way to solve this problem is to use iterative methods that have much less computational complexity. These methods iteratively solve a system of linear equations $\mathbf{H}\hat{\mathbf{s}} = \mathbf{y}$, using some initial approximation $\hat{\mathbf{s}}_0$. Approximate solution is consistently calculated during several steps (iterations) [5; 6].

Consider the Chebyshev method. The iteration scheme (4) with variable iteration parameters (5) is called the Chebyshev iteration scheme [5]:

$$\frac{\hat{\mathbf{s}}_i - \hat{\mathbf{s}}_{i-1}}{\tau_i} + \mathbf{H}\hat{\mathbf{s}}_{i-1} = \mathbf{y}, \quad (4)$$

$$\gamma_1\mathbf{E} \le \mathbf{H} \le \gamma_2\mathbf{E},\ \gamma_1 > 0,$$

$$\tau_i = \frac{\tau_0}{1 + \rho_0\mu_i},\ i = 1, 2, ..., n,\ \tau_0 = \frac{2}{\gamma_1 + \gamma_2},$$

$$\rho_0 = \frac{1-\xi}{1+\xi},\ \xi = \frac{\gamma_1}{\gamma_2}, \quad (5)$$

where γ_1, γ_2 are minimum and maximum eigenvalues of matrix $\mathbf{H}$; $\mu_i = \cos\frac{2i-1}{2n}\pi$ is the zeros of the Chebyshev polynomial $\mathrm{T}_n(x) = \cos(n \arccos x)$ on the interval $-1 \le x \le 1$. In this case, the following estimate takes place:

$$\|\mathbf{H}\hat{\mathbf{s}}_n - \mathbf{y}\| \le q_n \|\mathbf{H}\hat{\mathbf{s}}_0 - \mathbf{y}\|,\ q_n \le \varepsilon, \quad (6)$$

where $q_n = \frac{2\rho_1^n}{1 + \rho_1^{2n}}$, $\rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}$. Let us write the expression for the parameters τ_i:

$$\tau_i = 2/[\gamma_2 + \gamma_1 + (\gamma_2 - \gamma_1)(\cos\frac{2i-1}{2n}\pi),$$

$$i = 1, 2, ..., n. \quad (7)$$

Let us rewrite the expression (7) of a sequence of iterative parameters with a given maximum number of iterations $i_{\max}$, associated with the roots of Chebyshev polynomials, in the following form [4]:

$$\tau_i = \left[\frac{\gamma_2 - \gamma_1}{2}\cos\left(\frac{i - 1/2}{i_{\max}}\pi\right) + \frac{\gamma_2 + \gamma_1}{2}\right]^{-1}, \quad (8)$$

where γ_2, γ_1 are maximum and minimum eigenvalues of the channel matrix; $i_{\max}$ is maximum number of iterations, which is specified in the Chebyshev demodulation algorithm. Table 1 shows the simulation algorithm for the considered iterative demodulator with parameters τ_i. The simulation results are the immunity characteristics of the MIMO system for different numbers of antennas.

Table 1

Simulation algorithm for the considered iterative demodulator with parameters τ_i

Step number	Simulation program operations	Variables that are used in the program
1.	Start of signal-to-noise ratio cycle	*SNR*
2.	Start of experiments cycle	*L*=10000
3.	Transmitting antennas cycle	*M*
	Generation of a uniformly distributed random variable (for each antenna)	*x*
	Bit generation (1 or 0) from *x*	*b*
	Modulation (generation of information symbol for each antenna)	*s*
4.	End of the cycle (step 3)	
5.	Formation of a vector from generated information symbols for all transmitting antennas	$\boldsymbol{s}$
6.	MIMO channel matrix generation, consisting of complex transmitting coefficients	$\boldsymbol{H}$
7.	Generation of complex Gaussian noise vector	$\boldsymbol{n}$
8.	Simulation of signal and noise mixture	$\boldsymbol{y} = \boldsymbol{Hs} + \boldsymbol{n}$
9.	Application of the Chebyshev method for demodulation	
10.	Star of iterations cycle	i_max
11.	Obtaining estimate on the *i*-th iteration	$\hat{\mathbf{s}}_i$
12.	End of cycle (step 10)	
13.	Demodulation (obtaining estimate of the received vector of information symbols)	$\hat{\mathbf{s}}$
14.	Obtaining estimate of received bits vector	$\hat{\mathbf{b}}$
15.	Comparing vectors $\hat{\mathbf{b}}$ и $\mathbf{b}$ and detection of errors	*errors*
16.	Number of errors calculation	*sum*
17.	End of cycle (step 2)	
18.	Bit error ratio calculation for each SNR value	*BER*
19.	End of cycle (step 1)	
20.	Plotting bit error ratio versus signal-to-noise ratio for different numbers of antennas	$BER = f(SNR)$

Figure 3 shows the noise immunity dependencies $BER = f(SNR)$ obtained by the minimum mean square error (MMSE) method and the Chebyshev method using expression (8) for the parameters τ_i at different signal-to-noise ratio (SNR) values. The simulation was carried out for MIMO system with 4 transmitting and 4 receiving antennas with a number of experiments equal to 10,000.

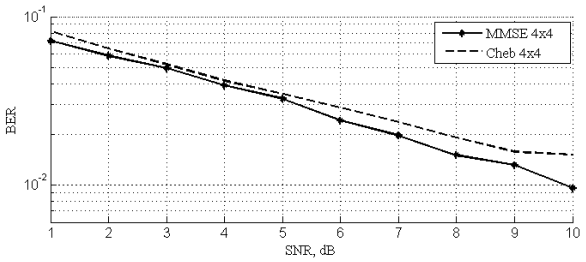


Fig. 3. BER performance of MMSE and Chebyshev demodulation algorithms

For the algorithm (4)–(7) due to computational instability, it is not indifferent in what sequence the zeros of the Chebyshev polynomial are taken. Next, we present a rule for constructing such a sequence of parameters $\{\tau_i\}$ of (8), for which the convergence of the iterative method is monotonic and there is no computational instability. From the sequence, it is necessary to form a permutation containing only the values of the iterative parameters with odd numbers and arrange them in a certain order as follows [4; 5]:

$$\theta_1 = \{1\},\ \theta_{2i-1}^{(2m)} = \theta_i^{(m)},\ \theta_{2i}^{(2m)} = 4m - \theta_{2i-1}^{(2m)} \quad (9)$$

Optimal permutation for 8 iterations is the following [5]: {1; 15; 7;9;3;13;5;11}.

Using the permutations, we can get better convergence with the same number of iterations (Fig. 4).

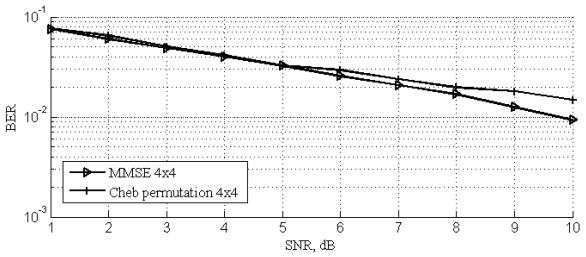


Fig. 4. BER performance of MMSE algorithm and Chebyshev algorithm with permutations of the iterative parameters

The rate of convergence of the iterative process (4)–(7) with the permutation of the iterative parameters in (9) can be found from the following expression [5]:

$$n = n_0(\varepsilon) = \frac{1}{2\sqrt{\xi}} \ln \frac{2}{\varepsilon}, \quad (10)$$

where $n_0(\varepsilon)$ is number of iterations sufficient to solve the system of linear equations with a given accuracy $\varepsilon > 0$.

The Chebyshev method provides fast convergence, but its application requires information about the minimum and maximum eigenvalues of the MIMO channel matrix.

Figure 5 shows the dependencies of noise immunity for the MMSE method and the Chebyshev method with permutations of iterative parameters, in which the eigenvalues of matrix **T** were used when calculating the iteration parameters:

$$\mathbf{T} = \mathbf{H}'\mathbf{H} + 2\sigma_\eta^2 \cdot \mathbf{1}, \qquad (11)$$

where $2\sigma_\eta^2$ is complex noise dispersion; $\mathbf{1}$ is identity matrix; $\mathbf{H}'$ is Hermitian conjugate matrix with respect to matrix $\mathbf{H}$.

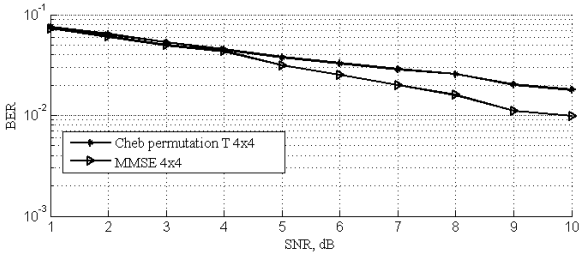


Figure 5. BER performance of MMSE algorithm and Chebyshev algorithm with permutations of the iterative parameters and eigenvalues of matrix **T**

Massive MIMO technology in future 5G systems [7; 8] will provide a significant increase in capacity of such systems. The considered demodulation algorithm based on the Chebyshev method is planned to be used for MIMO systems with large number of antennas, as well as for signals with higher modulation order. The main limitation of Chebyshev iteration method applicability is the need for a priori knowledge of MIMO channel matrix **H** spectrum boundaries (maximum and minimum eigenvalues), and with decreasing accuracy of these values, the convergence of the algorithm slows down. The task of finding and applying estimates of MIMO channel matrix eigenvalues in demodulation algorithms requires additional research.

REFERENCES

[1] Wei Xiang, Kan Zheng, Xuemin (Sherman) Shen. 5G Mobile Communications. [S.l.]: Springer International Publishing, 2017.

[2] Бакулин М.Г., Варукина Л.А., Крейнделин В.Б. Технология MIMO: принципы и алгоритмы. М.: Горячая линия – Телеком, 2014 [Bakulin M.G., Varukina L.A., Kreyndelin V.B. MIMO Technology: Principles and Algorithms. Moscow: Hotline-Telecom, 2014].

[3] Панкратов Д.Ю., Степанова А.Г. (МТУСИ, Москва) Моделирование системы MIMO // Международная научно-техническая конференция “INTERMATIC-2017”. Ноябрь 2017. М.: МИРЕА, 2017. С. 1052–1056. [Pankratov D.Yu., Stepanova A.G. MIMO system simulation // Proceedings of the International Scientific and Technical Conference “INTERMATIC-2017”, November, 2017. Moscow: MIREA, 2017. P. 1052–1056.

[4] Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю. Технологии в системах радиосвязи на пути к 5G. М.: Горячая линия – Телеком, 2018. [Bakulin M.G., Kreyndelin V.B., Pankratov D.Yu. Technologies in Radio Communication Systems on the Way to 5G. Moscow: Hotline-Telecom, 2018].

[5] Самарский А.А. Теория разностных систем. М.: Наука, 1977. [Samarskii A.A. Theory of Difference Schemes. Moscow: Nauka, 1977],

[6] Шлома А.М., Бакулин М.Г., Крейнделин В.Б. Новые алгоритмы формирования и обработки сигналов в системах подвижной связи / Под ред.А.М. Шломы. М.: Горячая линия – Телеком, 2008. [Shloma A. M., Bakulin M. G., Kreyndelin V. B. et al. New Algorithms for Generating and Processing signals in Mobile Communication Systems / Ed. by prof. A.M. Shloma. Moscow: Hotline-Telecom, 2008].

[7] Luo F.-L., Zhang C. Signal processing for 5G: algorithms and implementations. Chichester,: John Wiley & Sons Inc., 2016.

[8] Wunder G., Jung P., Wiedmann F. 5G NOW: Non-orthogonal, Asynchronous Waveforms for Future Mobile Applications // IEEE Communications Magazine. 2014. Vol. 52. N 2.

Analytical Method of Forming the Equivalent Virtual MIMO Channel Matrix for Space-Time Codes of High Dimensions

Andrey Reznev
Scientific and Research Center ATLAS
Moscow, Russia
Andrey.Reznev@gmail.com

***Abstract*. This work provides the analytical method for Equivalent Virtual Channel Matrix (EVCM) synthesis. The approach of Equivalent Virtual Channel Matrix might be applied for any type of STC to convert it into BLAST STC structure. The work demonstrates that manual methods are not applicable for large scale STC structures due to overladen manual derivation. Decomposition of STC codes matrices into multiplication product of channel part and symbols allows picking the symbol vector as the separate part of model equation. Linear transformation converts the multiplication product into BLAST type system with new channel matrix. This virtual channel matrix equation is the analytical equation for EVCM.**

***Key words:* MIMO, BLAST, EVCM, Equivalent Virtual Channel Matrix, Space Time Coding, Golden code, STC.**

I. INTRODUCTION

The group of technologies of high speed mobile Internet access is coming, the marketing umbrella for those ones is "Fifth Generation – 5G". Smooth migration from 4G to 5G supported by radio access technology named Multiple Input – Multiple Output (MIMO), which uses the number of transmitting and receiving antennas within the system. Some of the 3GPP standards declare that possible numbers of antennas may exceed 128 [1].

The necessity to increase number of antennas provides by BLAST which is the common space time coding approach, when one information symbol sends through one antenna simultaneously with other symbols [2; 3]. One of the ways is transmission for several time intervals which might decrease the number of antennas in MIMO.

II. SYSTEM MODEL

Some of the well-known multi-interval STCs might be orthogonal in order to achieve low decoding complexity [2].

Alamouti and Golden structures are STC which use multi-interval transmission, and they both need to be transformed into BLAST system model. The way of transformation replaces the channel matrix to Equivalent Virtual Channel Matrix (EVCM) [2], which is pretty simple for two antennas but requires the overladen manual works for the big MIMO system.

EVCM structure is the composition of channel information and coding method with the dimension which is the result related to number of transmission antennas and transmission intervals.

In case of Alamouti STC, we faced with the system of two equations which give us 2x2 matrix, for classical Golden we can see the system of four equations and 4x4 matrix. The AB or ABBA combination will bring us at least eight complicated equations with 8x8 matrixes. The increased numbers of antennas is the result of Shannon—Hartley theorem when the system throughput multiplies with numbers of paths between transmitting and receiving antennas. 5G speed levels will accelerate the number of antennas [2; 4].

We use system model, see (1) **Ошибка! Источник ссылки не найден.**; [3]:

$$\mathbf{Y} = \mathbf{H} \cdot \mathbf{S}(\boldsymbol{\theta}) + \boldsymbol{\Gamma} \qquad (1)$$

In the system model we have: $\mathbf{Y}$ is received signal matrix, dimension is $M \times L$; $\mathbf{H}$ is channel matrix, dimension $M \times N$; $\mathbf{S}(\boldsymbol{\theta})$ is space time code matrix, dimension $N \times L$; $\boldsymbol{\theta}$ is transmitted information vector, dimension $Q \times 1$; $\boldsymbol{\Gamma}$ is an independent, identically distributed noise matrix, dimension $M \times L$; M is number of receiving antennas; N is number of transmitting antennas; L is number of time intervals for transmitting; Q is number of transmitting information symbols in STC.

Some of the STCs may include symbols and complex conjugate symbols both, the typical example is Alamouti code which is the origin code for numbers of large STCs. We need to consider this to produce correct common method for EVCM synthesis. We generalize the system model (1) to the new equation to add complex conjugate symbols (2):

$$\mathbf{Y} = \mathbf{H} \cdot \mathbf{S}\left(\boldsymbol{\theta}, \boldsymbol{\theta}^*\right) + \boldsymbol{\Gamma} \qquad (2)$$

III. PROPOSED METHOD AND RESULTS

For STC we may have following view (3), where every matrix element depends on the whole raw of information symbols including their conjugate values:

$$\mathbf{S}\left(\boldsymbol{\theta}, \boldsymbol{\theta}^*\right) = \begin{bmatrix} S_{11}\left(\theta_1, \theta_2 \ldots \theta_Q; \theta_1', \theta_2' \ldots \theta_Q'\right) & \cdots & S_{1L}\left(\theta_1, \theta_2 \ldots \theta_Q; \theta_1', \theta_2' \ldots \theta_Q'\right) \\ \vdots & \vdots & \vdots \\ S_{N1}\left(\theta_1, \theta_2 \ldots \theta_Q; \theta_1', \theta_2' \ldots \theta_Q'\right) & \cdots & S_{NL}\left(\theta_1, \theta_2 \ldots \theta_Q; \theta_1', \theta_2' \ldots \theta_Q'\right) \end{bmatrix} \qquad (3)$$

Now every element $S_{i,j}\left(\theta_1,\theta_2\ldots\theta_Q;\theta_1',\theta_2'\ldots\theta_Q'\right)$ of matrix $\mathbf{S}(\boldsymbol{\theta},\boldsymbol{\theta}^*)$ (3) might be represented as (4):

$$S_{i,j}\left(\theta_1,\theta_2\ldots\theta_Q;\theta_1',\theta_2'\ldots\theta_Q'\right)=\sum_{q=1}^{Q}\alpha_{i,j}\cdot\theta_q+\beta_{i,j}\cdot\theta_q', \quad i=1;2;\ldots N,\ j=1;2;\ldots L \tag{4}$$

Those $\alpha_{i,j}(q),\beta_{i,j}(q)$ factors inherit the original STC properties.

Combining (4), (3), we see (5):

$$\mathbf{S}(\boldsymbol{\theta},\boldsymbol{\theta}^*)=\begin{bmatrix}\sum_{q=1}^{Q}\alpha_{11}(q)\cdot\theta_q+\beta_{11}(q)\cdot\theta_q' & \cdots & \sum_{q=1}^{Q}\alpha_{1L}(q)\cdot\theta_q+\beta_{1L}(q)\cdot\theta_q' \\ \vdots & \vdots & \vdots \\ \sum_{q=1}^{Q}\alpha_{N1}(q)\cdot\theta_q+\beta_{N1}(q)\cdot\theta_q' & \cdots & \sum_{q=1}^{Q}\alpha_{NL}(q)\cdot\theta_q+\beta_{NL}(q)\cdot\theta_q'\end{bmatrix}= \tag{5}$$

$$=\sum_{q=1}^{Q}\begin{bmatrix}\alpha_{11}(q) & \cdots & \alpha_{1L}(q)\\ \vdots & \vdots & \vdots \\ \alpha_{N1}(q) & \cdots & \alpha_{NL}(q)\end{bmatrix}\cdot\theta_q+\begin{bmatrix}\beta_{11}(q) & \cdots & \beta_{1L}(q)\\ \vdots & \vdots & \vdots \\ \beta_{N1}(q) & \cdots & \beta_{NL}(q)\end{bmatrix}\cdot\theta_q'$$

Reorganize (5) to the compact view (6):

$$\mathbf{S}(\boldsymbol{\theta},\boldsymbol{\theta}^*)=\sum_{q=1}^{Q}\mathbf{A}_q\cdot\theta_q+\mathbf{B}_q\cdot\theta_q', \text{ where}$$

$$\mathbf{A}_q=\begin{bmatrix}\alpha_{11}(q) & \cdots & \alpha_{1L}(q)\\ \vdots & \vdots & \vdots \\ \alpha_{N1}(q) & \cdots & \alpha_{NL}(q)\end{bmatrix},\quad \mathbf{B}_q=\begin{bmatrix}\beta_{11}(q) & \cdots & \beta_{1L}(q)\\ \vdots & \vdots & \vdots \\ \beta_{N1}(q) & \cdots & \beta_{NL}(q)\end{bmatrix}_{q=1;2;\ldots Q} \tag{6}$$

New matrices $\mathbf{A}_q\ \mathbf{B}_q$ derived from original STC structure. We might use our preparations to transform the signal model to virtual BLAST type system (2).

For our further steps we present the transmitted symbols to make explicit real and imaginary parts (7):

$$\begin{aligned}\theta_q&=\mathrm{Re}(\theta_q)+j\,\mathrm{Im}(\theta_q)=\theta_{Rq}+j\theta_{Iq}\\ \theta_q'&=\mathrm{Re}(\theta_q')-j\,\mathrm{Im}(\theta_q')=\theta_{Rq}-j\theta_{Iq}\end{aligned} \tag{7}$$

Combining (7) and (6), we see (8):

$$\mathbf{S}(\boldsymbol{\theta},\boldsymbol{\theta}^*)=\sum_{q=1}^{Q}\mathbf{A}_q(\theta_{Rq}+j\theta_{Iq})+\mathbf{B}_q(\theta_{Rq}-j\theta_{Iq})= \tag{8}$$

$$=\sum_{q=1}^{Q}(\mathbf{A}_q+\mathbf{B}_q)\cdot\theta_{Rq}+j(\mathbf{A}_q-\mathbf{B}_q)\cdot\theta_{Iq}$$

Apply vectorization to model (2),

$$\mathbf{Y}=(1\otimes\mathbf{H})\cdot vec(\mathbf{S}(\boldsymbol{\theta},\boldsymbol{\theta}^*))+\boldsymbol{\eta} \tag{9}$$

Vectorization result for STC has the new form (10):

$$vec(\mathbf{S}(\boldsymbol{\theta},\boldsymbol{\theta}^*))=vec\left(\sum_{q=1}^{Q}(\mathbf{A}_q+\mathbf{B}_q)\cdot\theta_{Rq}+j(\mathbf{A}_q-\mathbf{B}_q)\cdot\theta_{Iq}\right)= \tag{10}$$

$$=\sum_{q=1}^{Q}\theta_{Rq}\cdot vec(\mathbf{A}_q+\mathbf{B}_q)+j\theta_{Iq}\cdot vec(\mathbf{A}_q-\mathbf{B}_q)$$

Let's mark some variables within (10) and combine multipliers, and see (11):

$$vec(\mathbf{S}(\boldsymbol{\theta},\boldsymbol{\theta}^*))=\mathbf{T}_1\cdot\boldsymbol{\theta}_R+j\cdot\mathbf{T}_2\cdot\boldsymbol{\theta}_I, \text{ where}$$

$$\mathbf{T}_1=\left[vec(\mathbf{A}_1+\mathbf{B}_1);vec(\mathbf{A}_2+\mathbf{B}_2)\ldots vec(\mathbf{A}_Q+\mathbf{B}_Q)\right]$$

$$\mathbf{T}_2=\left[vec(\mathbf{A}_1-\mathbf{B}_1);vec(\mathbf{A}_2-\mathbf{B}_2)\ldots vec(\mathbf{A}_Q-\mathbf{B}_Q)\right] \tag{11}$$

$$\boldsymbol{\theta}_R=\begin{vmatrix}\theta_{R1}\\ \theta_{R2}\\ \vdots\\ \theta_{RQ}\end{vmatrix},\quad \boldsymbol{\theta}_I=\begin{vmatrix}\theta_{I1}\\ \theta_{I2}\\ \vdots\\ \theta_{IQ}\end{vmatrix}$$

Substituting (11) in (9), we can get the new signal model (12):

$$\mathbf{Y}=(1\otimes\mathbf{H})\cdot(\mathbf{T}_1\cdot\boldsymbol{\theta}_R+j\mathbf{T}_2\boldsymbol{\theta}_I)+\boldsymbol{\eta} \tag{12}$$

In order to simplify the equation we may reorganize to (12) to (13)

$$\mathbf{Y}=(1\otimes\mathbf{H})\cdot\mathbf{J}\cdot\boldsymbol{\theta}^{R/I}+\boldsymbol{\eta}, \text{ where} \tag{13}$$

$$\mathbf{J} = \left[\mathbf{T}_1; j\mathbf{T}_2\right]$$

$$\boldsymbol{\theta}^{R/I} = \left|\begin{matrix}\boldsymbol{\theta}^R \\ \boldsymbol{\theta}^I\end{matrix}\right| = \left|\theta_1^R \quad \theta_2^R \quad \ldots \quad \theta_Q^R \quad \theta_1^R \quad \theta_1^R \quad \ldots \quad \theta_1^R\right|^{\mathrm{T}}$$

Analytical equation for EVCM has the view.

$$\mathbf{H}_{EVCM} = \left(\mathbf{1} \otimes \mathbf{H}\right) \cdot \mathbf{J} \qquad (14)$$

EVCM system model transformation for information sequences with complex conjugate symbols doubling the size of EVCM dimension and transmitting vector should be replaced to double size vector constructed from real and imaginary parts of original symbols.

IV. REFERENCES

[1] TS 23.501. System Architecture for the 5G System. ETSI. 2017.

[2] Бакулин М.Г., Варукина Л.А., Крейнделин В.Б. Технология MIMO: принципы и алгоритмы. М.: Горячая линия – Телеком, 2014 [Bakulin M.G., Varukina L.A., Kreyndelin V.B. MIMO Technology: Principles and Algorithms. Moscow: Hotline-Telecom, 2014].

[3] Larsson E. G., Stoica P. Space-Time Block Coding for Wireless Communications. 3rd ed. Cambridge: Cambridge University Press, 2008.

[4] Oesges C., Clerckx B. MIMO Wireless Communications. Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems. Oxford: Academic Press, 2013.

[5] Крейнделин В., Резнев А. Методы формирования пространственно временных матриц для систем MIMO высокой размерности // Электросвязь. 2017. №4. С. 46–50. [Kreindelin V.B., Reznev A.A. Methods for STC matrices synthesis for Large Scale MIMO structures // Electrical Communications. 2017. №4. P. 46–50].

.

New Algorithms for Processing Signals at the Receiver Side for Wireless Communication Systems with Massive MIMO Technology

Aleksey E. Smirnov
Department of Information Security
Moscow Technical University of Communications and Informatics
Moscow, Russia
smirnov.al.ed@gmail.com

***Abstract*. Multiple-Input Multiple-Output systems (MIMO) have been a trending direction of research in the past few years. MIMO technology offers increased throughput compared to single-antenna systems. Extension of MIMO systems is a large-scale MIMO (or massive MIMO) that use antenna arrays with a few hundred antennas. Therefore, the number of operations required for computation of the estimate of the vector of transmitted symbols is increasing with the increased number of antenna arrays. The Minimum Mean Square Error (MMSE) detection technique is the basic one for MIMO systems due to low computational expenses as compared with the Maximum Likelihood (ML) method and a higher resistance to noise as compared to Zero Forcing (ZF) technique. Unfortunately, MMSE demodulation technique cannot be realized for massive MIMO systems due to high level of computational expenses.**

The paper provides us description of new low-complexity as compared to MMSE iterative demodulation technique for massive MIMO systems. To compare a resistance to noise of new iteration demodulation technique with MMSE technique it is developed a simulation model.

***Key words:* MIMO, detection, computational complexity, Neumann series, iterative detector.**

I. INTRODUCTION

The problem of the received signal demodulation in the multi-array systems at the receiver end at the known matrix of the channel is one of the traditional aspects in the sphere of wireless telecommunication. The problem consists in recovering of the transmitted signal at the receiver end from the received vector of counts at the known channel matrix and at known statistical characteristics of the noise [1].

There is a certain number of known algorithms used for the purpose of demodulation, whereas each of the algorithms has its own advantages and drawbacks as compared to the other. These algorithms allow computing the estimate of the vector of transmitted symbols at the receiver end. Using some of them allows obtaining high accuracy of the estimate, however, at that, as a consequence, there occurs high complexity of the signal processing. On the contrary, others possess a lower complexity but the estimate turns out to be less accurate. Depending upon specific conditions and criteria at developing of the telecommunication system, it is necessary to make a choice in favor of one of the algorithms.

II. SYSTEM MODEL

An abstract model of the MIMO system based on which different configurations of the system can be built (Fig. 1).

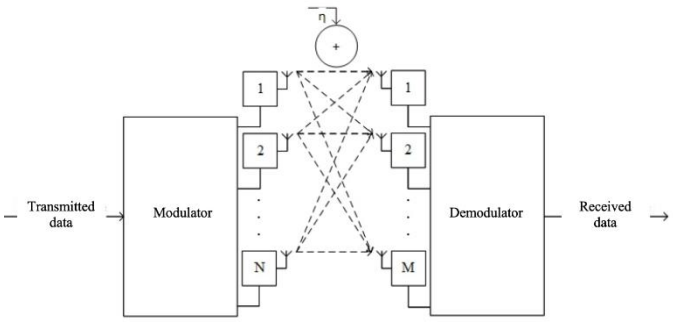


Figure 5. The MIMO system model

Here N is the number of antennas at the transmitter part, M is the number of antennas at the receiver end. At the receiver end, the signals from each of the transmitting antennas are emitted simultaneously and within one and the same frequency band. In the channel these signals are subject to the influence of the Rayleigh fading and the additive Gaussian white noise. Therefore, an additive aggregate of N transmitted signals arrives at each of M receiving arrays.

Mathematical model of the MIMO system has the following representation:

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{x} + \boldsymbol{\eta}, \quad (1)$$

where y is the vector of received complex counts with the dimensionality $M \times 1$; $\mathbf{H}$ is the matrix of the telecommunication channel complex transmission coefficients with the dimensionality $M \times N$, $\mathbf{x}$ is the vector of transmitted informational symbols with the dimensionality $N \times 1$; $\boldsymbol{\eta}$ is the complex random Gaussian noise vector in the telecommunication channel with the dimensionality $M \times 1$.

Three main techniques used for computing the estimate of transmitted symbols [1] are the following:

- ZF;
- MMSE;
- ML.

A. Zero Focrcing (ZF)

ZF demodulator is the best linear demodulation technique in terms of complexity but suffers from noise enhancement. Using this demodulator, the estimate is computed under the equation:

$$\hat{\mathbf{x}}_{ZF} = \mathbf{G}_{ZF} \cdot \mathbf{y} = \left(\mathbf{H}^{'}\mathbf{H}\right)^{-1} \cdot \mathbf{H}^{'} \cdot \mathbf{y} \qquad (2)$$

In order to compute estimate of transmitted symbols using ZF, it is considered the matrix of the channel only but the availability of the noise is disregarded that results in decreasing of the resistance to noise.

B. Minimum Mean Square Error (MMSE)

MMSE is introduced to overcome the negative effect of noise enhancement. This algorithm provides better estimate computation results because availability of the noise is taken into consideration. The estimate obtained as the result of the MMSE demodulator operating execution is represented mathematically as follows:

$$\hat{\mathbf{x}}_{MMSE} = \mathbf{G}_{MMSE} \cdot \mathbf{y} = \left(\mathbf{H}^{'}\mathbf{H} + 2\sigma^2\mathbf{1}\right)^{-1} \cdot \mathbf{H}^{'} \cdot \mathbf{y} \qquad (3)$$

where $2\sigma^2$ is the aggregate dispersion of the real and imaginary components of the Gaussian noise vector and I is identity matrix.

C. Maximum Likelihood (ML)

The ML method possesses the best characteristics of resistance to noise [1]. This demodulation technique minimizes the difference between y and x distorted by the channel. The estimate obtained with the help of this method minimizes the square of the discrepancy norm:

$$\hat{\mathbf{x}}_{ML} = \arg\min_{\mathbf{x}\in\mathbf{X}^N} \left\|\mathbf{y} - \mathbf{H}\cdot\mathbf{x}\right\|^2 \qquad (4)$$

III. COMPUTATIONAL COMPLEXITY

The mathematical algorithms can be classified [1] in accordance with complexity of their execution. Considering that the algorithm is composed of a sequence of strictly determined operations (commands) [1], the algorithms can be classified depending upon the number of operations required for obtaining of the result. The parameter, which is called the computational complexity, is often used for classification of the algorithms. The term is used for determining of the number of elementary arithmetical operations, which must be performed for obtaining the solution to a certain problem.

Let us consider in detail the detection algorithms mentioned above. We shall regard the telecommunication systems, in which the number of arrays at the receiver end and the number of arrays at the transmitter end are equal; that is $N = M$. The number of elementary arithmetical operations required for obtaining of the estimate $\hat{\mathbf{x}}_{MMSE}$ and $\hat{\mathbf{x}}_{ZF}$ depends primarily upon the number of arrays at the received end and at the transmitted part that is upon the dimension of the channel matrix. To find the minimum, according to (4), it is necessary to perform enumeration upon all the possible combinations of the vector of complex informational symbols $\mathbf{x}$. The number of operations depends upon two parameters of the system such as the number of transmitting arrays and the type of modulation used.

Considering the trend to growth of the data transmission rate in the present-day telecommunication systems, signal processing at the receiver end has to be performed [1] during the interval of time equal to the duration of the informational symbol. With the increase of the number of the receiving and transmitting arrays and modulation type the number of mathematical operations, which are required for obtaining of the estimate of the vector of transmitted symbols, is also increasing. Due to the fact, that implementation of the ML method requires execution of a great number of operations, obtaining of the estimate for the vector of transmitted symbols using the ML method cannot be realized practically [1].

Development of wireless communication systems resulted [3] in origination of large-size MIMO systems, or the so-called massive MIMO. Their main difference from the existing multi-array systems is in availability of the antenna arrays, the number of arrays in which exceeds by several times the number of arrays in traditional systems [3][3]. In this condition, an apparent drawback as compared with the traditional multi-array systems is in the fact that increasing of the number of arrays at the receiver and transmitter ends results in increasing of the channel matrix dimensionality, and, consequently, in increasing of the number of mathematical operations [5] necessary for obtaining of the estimate at demodulation. Therefore, with the increased number of arrays the computational complexity while obtaining the estimates $\hat{\mathbf{x}}_{ZF}$ and $\hat{\mathbf{x}}_{MMSE}$ is increasing, however, the advantage of a low complexity as compared with obtaining of the estimate $\hat{\mathbf{x}}_{ML}$ is preserved.

IV. COMPLEXITY OF MMSE DETECTOR

We address to the equation (3) in order to obtain the estimate of MMSE. We shall separately consider all the operations used for obtaining the estimate upon the MMSE algorithm:

- multiplication of the matrix with the dimensionality $N \times N$ by the matrix with the dimensionality $N \times N$;
- multiplication of the matrix with the dimensionality $N \times N$ by the vector-column with the dimensionality N (this operation has to be performed twice);
- summation of two matrices with the dimensionality $N \times N$;
- inversion of the matrix with the dimensionality $N \times N$ with the help of solving of the system of linear algebraic equations (SLAE).

Main contribution in overall computational complexity for obtaining of the estimate $\hat{\mathbf{x}}_{MMSE}$ consists in the operation of multiplication of two matrices and the operation inversion of the matrix.

It is important to notice, that the channel matrix and the vector of received counts contain complex numbers. Therefore, total number of elementary arithmetical operations required for obtaining of the estimate for real multiplications and two real summations is increasing.

V. ITERATIVE DETECTOR FOR MASSIVE MIMO

To recover of the transmitted signal at the receiver end from the received vector in wireless communication systems based on massive MIMO technology, we propose a new iteration detector with lower computational complexity as compared with known MMSE detector.

Let $\mathbf{T} = \left(\mathbf{H}^{'}\mathbf{H} + 2\sigma^2\mathbf{1}\right)$ and $\mathbf{Y} = \mathbf{H}^{'} \cdot \mathbf{y}$, next step we multiple left and right side (3) by T and the estimate of the vector of transmitted is represented as follows:

$$\mathbf{T} \cdot \hat{\mathbf{x}}_{MMSE} = \mathbf{Y} \quad (5)$$

Then the process of obtaining of the estimation $\hat{\mathbf{x}}_{MMSE}$ at the receiver end consists in solving of the given system of linear equations (SOLE). There are so many different methods of solving the SOLE [6]**Ошибка! Источник ссылки не найден.**. One part of them are direct methods: provide an exact solution. These methods have high computational complexity. The second ones are indirect solver. Direct methods for obtaining exact solution of SOLE have the same computational complexity as compared to obtaining of the estimation $\hat{\mathbf{x}}_{MMSE}$ according to equation (3).

To decrease computational complexity for signal processing at the receiver side, we propose to use iterative methods for solving SOLE (5). Using these methods, it is necessary to start with a first approximation and compute iteratively sequence of approximation without ever reaching exact solution. In this case, we have an opportunity to reduce number of iterations for solving SOLE and, consequently, decrease of the number of mathematical operations necessary for obtaining of the estimate at demodulation. The main goal while choosing the best iterative methods for solving particular SOLE is in finding compromise between complexity and accuracy of iterative method.

The stop criteria for these methods are the maximum number of iterations or defined small value of error $\varepsilon < \max\limits_{1 \le i \le N} \left\| \hat{\mathbf{x}}^{i+1}_{PROP_j} - \hat{\mathbf{x}}^{i}_{PROP_j} \right\|$ and we can rewrite equation (5) as:

$$\mathbf{T} \cdot \hat{\mathbf{x}}_{PROP} \approx \mathbf{Y} \quad (6)$$

where i is the iteration number; $j = 1,2 \ldots N$ is the index of vector $\hat{\mathbf{x}}_{PROP}$ component.

To decrease computational complexity of obtaining of the estimation $\hat{\mathbf{x}}_{PROP}$ at the receiver end a lot, it is possible to use iterative methods to solve the SOLE (6) with maximum number of iterations $L = \frac{N}{4}$.

We propose to use iterative BiConjugate Gradient Stabilized (l) (BiCGstab(l)). This approach is described in [7] [8]. To achieve a maximum effect in decreasing computational complexity for obtaining estimation $\hat{\mathbf{x}}_{PROP}$ the maximum number of iterations L should be smaller than matrix $\mathbf{H}$ dimensions N for few times $L << N$.

VI. SIMULATION RESULTS

To compare the results of noise immunity of different demodulation techniques for massive MIMO systems, we perform a simulation model. Simulation parameters are in Table.

Simulation parameters

Parameter	Value
Channel	MIMO rayleigh fading channel
Number of antennas on BS side	64
Number of antennas on user side	64
Transceiver architecture	V-BLAST
Error correcting code type	Convolutional coding (rate 1/2)
Modulation type	16-QAM
Total number of iterations for solving SOLE	16

And as we use convolutional coding in our system, we have one more operation at the receiver side that requests cubic level of computational complexity matrix inversion for calculating correlation matrix for MMSE despite we need to know only main diagonal for soft decision demodulation based on log-likelihood ratio. Its operation describes the following equation:

$$\mathbf{R}_{MMSE} = \left(\mathbf{H}'\mathbf{H} + 2\sigma^2\mathbf{I}\right)^{-1} \cdot 2\sigma^2 \quad (7)$$

Obtaining correlation matrix for MMSE requires more elementary arithmetical operations than obtaining solution of SOLE (6) using BiCGstab(l). We have no need to calculate each element of matrix $\mathbf{R}_{MMSE}$. To avoid direct matrix inversion, we propose to use approximation based on Neumann series. This way we can reduce computational expenses for this operation. Use of Neumann series for decreasing computational complexity of matrix inversion is described in [9].

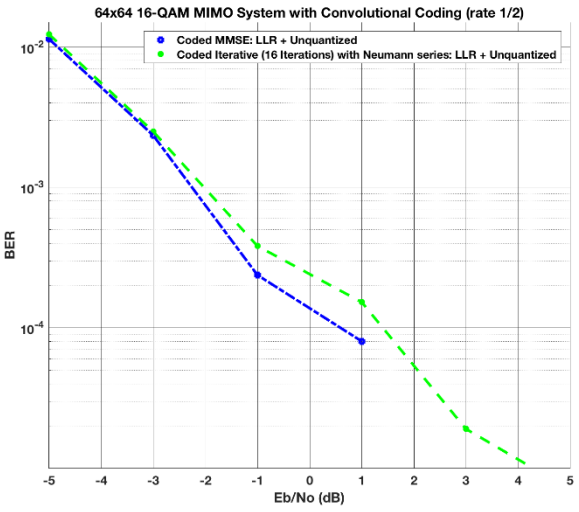


Fig. 6. Noise immunity comparison for MMSE and iterative detectors

The results of simulation (Fig. 2) show that application of iterative detector combined with Neumann series approach for signal processing at the receiver end allows decreasing a total number of the required elementary operations for 5,5 times with

losing in the resistance to noise (about 0,2 dB for BER $=10^{-3}$) for the matrices of complex transmission coefficients of the telecommunication channel with the large dimensionality, which can be found in the promise massive MIMO systems.

VII. REFERENCES

[1] Бакулин М.Г., Варукина Л.А., Крейнделин В.Б. Технология MIMO: принципы и алгоритмы. М.: Горячая линия – Телеком, 2014. [Bakulin M.G., Varukina L.A., Kreyndelin V.B. MIMO technology: principles and algorithms. Moscow: Hot line – Telecom, 2014].

[2] Levitin A.V. Introduction to the design and analysis of algorithms. Reading: Addison-Wesley, 2007.

[3] Larsson E.G., Edfors O., Tufvesson F.et al. Massive MIMO for next generation wireless systems // Communications Magazine, IEEE. 2014. Vol. 52, N 2. P. 186–195.

[4] Rusek F., Persson D., Buon Kiong Lau, et al. Scaling up MIMO: challenges with very large arrays // Signal Proccesing Magazine, IEEE. 2013. Vol. 30, N 1. P. 40–60.

[5] Крейнделин В.Б., Смирнов А.Э. Снижение вычислительной сложности алгоритмов демодуляции в многоантенных системах за счёт использования быстрых алгоритмов // Telecommunications and Radio Engineering. 2016. Том. 75, N 19. С. 1757–1773. [Kreyndelin V.B., Smirnov A.E. Decreasing of computational complexity of demodulation algorithms in multi-antenna systems due to application of fast algorithms // Telecommunications and Radio Engineering. 2016. Vol. 75, N 19. P. 1757–1773].

[6] Самарский А.А., Гулин А.В, Численные методы. М.: Наука, 1989. [Samarskii A.A., Gulin A.V. Numerical methods. Moscow: Nauka, 1989].

[7] Sleijpen G.L.C., van der Vorst H.A., Fokkema D.R. BiCGstab(l) and other hybrid Bi-CG methods // Numerical Algorithms. 1994. Vol. 7. P. 75–109.

[8] Sleijpen G.L.C., Fokkema D.R. BiCGstab(l) for linear equations involving unsymmetric matrices with complex spectrum // Electronic Transactions on Numerical Analysis. 1993. Vol. 7. P. 11–32.

[9] Monteiro F. A., Rosário F., Rodrigues A. Fast matrix inversion updates for massive MIMO detection and precoding // Signal Proccesing Letters, IEEE. 2016. Vol. 23, N 1. P. 75–79.

CyberSecurity Concept For New Generation Telecommunication Networks

Nikolay Ulpe
JSC InfoCES
Moscow, Russia

Sergey Melnik
Moscow Technical University of Communications and Informatics (MTUCI)
Central Science Research Telecommunication Institute (CSRTI)
Moscow, Russia
sv.melnik@bk.ru

Abstract. **The article describes the main perspective directions of development of Worldwide network. Potentially number of connections to the Internet will increase to 70 trillion by 2025, reconstruction and modernization of modern communication networks and new approach to ensuring information security. Information security becomes a cornerstone of creation of communication networks of future generations and how data and communication networks will be protected safety of Society and State will depend. If today objects, potentially vulnerable from the Internet, are computers and the maximum harm which hacker attack can cause is a temporary suspension of work of automated control systems and access to information, any element of Digital economy can become object of attack in networks of the Industrial Internet of things. It is necessary for prevention of potential threats: to classify these threats; to have the uniform formal description of threats and the warning and correcting influence; to have the uniform concept of the organization of information security. This article brings up questions which creation of the uniform concept of information security of Russia with the purpose to minimize potential danger of harmful impact on significant objects of protection has to form the basis and to exclude possibility of network attacks to critical objects.**

Key words: **Internet, Security, IoT, IIoE, IoE, 5G, Modeling, Timing.**

VIII. INTRODUCTION

Internet of Everything (IoE) is the new concept for information exchange in telecommunication networks. This concept includes all main fields: Internet of Tings (IoT), Industrial Internet of Tings (IIoT) and Human data exchange networking. It is not only data exchange between any probes and big data centre. It is very important to have sufficient type approval and cyber security mechanism for this new concept. The report consists of the universal incident object description exchange format based on Recommendation ITU-T X.1541. This method describes all possible incidents based on standard classes. This is the basement for cyber security modeling and building new protected software solutions for secure data exchange. We used 27 basic classes ITU-T X.1541.

It is well understood that more and more devices are connected to worldwide network the Internet every year. If today the majority of connections corresponds to the animated users, i.e. the main sources of data are people, in the near future inanimate objects will become the main sources. For a long time networks of intermachine interaction of M2M which are gradually built according to requirements to creation of networks of the Internet of Thing (IoT) develop. In parallel there is a development of the uniform principles of remote control of important objects of infrastructure according to requirements of creation of networks of the Industrial Internet of prophetic (IIoT). Development in the field of the tactile Internet (TIoT) became more active in the latest time. For each of these directions it is required to establish necessary and sufficient measures for information security.

The Internet of Everything (IoE) or the General Internet is the new concept of exchange of information in communication networks including all above-named directions in total with information from subscribers of a communication network.

The main principles when determining measures of ensuring information security are:

- ranking of information sources;
- classification of protected objects;
- the uniform formal description of the threats arising in case of unauthorized access to information and in case of malicious change of information;
- the uniform formal description of the influence preventing unauthorized access and change of information;
- the uniform formal description of the influence correcting the unauthorized access taking place and change of information.

Extensive works on standardization are conducted for the purpose of ensuring uniform approach to information security. The main organizations which carry out these works are:

- International union of Telecommunication (ITU-T);
- International organization for standardization (ISO);

- Cloud Security Alliance (CSA), the organization of information security for cloudy structures..

Networks of the Internet of things unite a large number of the various sensors communicating on the public Internet. Security of objects of networks of the IoT has a class low and average. Networks of IIoT make the objects relating to infrastructure of digital economy. Security of objects of networks of the IIoT has a high class. Objects of the tactile Internet is an opportunity to feel the object which is at distance. Sensors which transfer its characteristics which are reproduced on the reception end are connected to object and create effect of a touch. Such mechanisms are used, for example, at remote inspection of the patient in a telemedicine. In its current stage of development there are devices for carrying out remote operations.

Considering high responsibility at information transfer, it is necessary to provide mechanisms of a guarantee of appropriate information security in future General Internet already today. The international standards in the field of cybersafety are developed and worked.

In Russia works on creation of standards for voluntary certification of objects of information security which have to become part of the uniform concept are also conducted.

II. INFORMATION SECURITY STANDARDIZATION

Groups of scientists and engineers worldwide work today on ensuring information security for IoT/IIoT/TIoT. In the field of information security treat the main aspects of standardization:

- risks assessment;
- general description of information on incidents;
- Web applications protection;
- detection of anomalies in a network traffic, "botnet" detection/protection;
- application of metrics and indexes for assessment of threats and measures of protection.

The main standards in the field of information security of IoT are developed in the International Union of Telecommunication (ITU-T), the research commissions (SG 13, SG17), in Cloud Security Alliance (CSA) and in ISO/IEC - research committee JTC1 SWG 5 and the SC 27.WG2 working group. There are general activities of the research commission of SG17 Q4 "Cyber security".

III. CYBEX METHODOLOGY (STRUCTURED CYBERSECURITY EXCHANGE, X.1500)

The principle of the organization of the protected data exchange between consumers is the cornerstone of counteraction to threats of information security. This exchange is specified by the recommendation of ITU-T X.1500 [1].

The methodology of Structured Cybersecurity Exchange (CYBEX) is developed for the protected data exchange between consumers. The methodology of CYBEX represents the five-level block model of data exchange shown in figure 1.

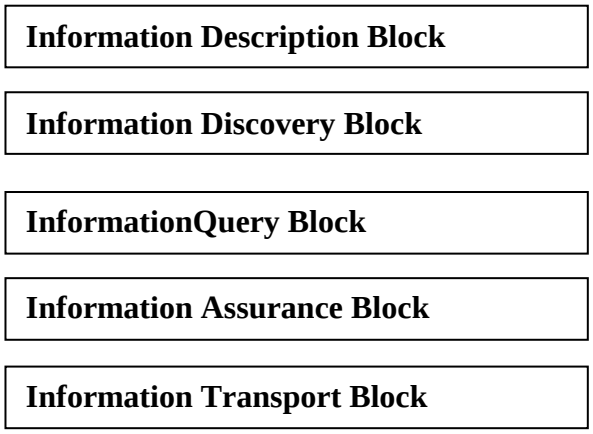


Fig. 1. Functional CYBEX blocks

IV. INFORMATION DESCRIPTION BLOCK

This functional description of information which is required to be transferred between consumers includes a format and language of the description. Formats and languages of the description information are provided in 21 international standards [2]. Standards of the block of the description of information include components on:

- format of the description of vulnerabilities of CVE – X.1520 [4];
- the catalog of descriptions of threats of CWE – X.1524 [5];
- format of the description of the CPE platforms – X.1528 [6];
- metrics for the quantitative assessment of vulnerabilities of CVSS – X.1521 [7];
- the catalog of patterns of attacks and methods of protection of CAPEC – X.1544 [8];
- language of the description of criteria of the vulnerable software of OVAL – X.1526 [9];
- format of the description of incidents of computer safety of IODEF – X.1541 [10].

V. INFORMATION DISCOVERY BLOCK

This functional block is responsible for definition and research of information on a source. There are two paradigms such as centralized maintaining a database of the entrusted sources (OID [10]) and the decentralized mechanism. On the Internet both of these mechanisms are used. The centralized mechanism of maintaining a database of the entrusted sources is based on entering of information on safety into a special database. Each country has the digital identifier in OID. The OID 2.48 identifier is appropriated to the safe site.

Together with the centralized the decentralized method of identification of safe sources RDF World Wide Web Consortium (W3C) [11] is used. RDF works at a basis of own search algorithm includes a method of an automatic assessment of safety of sources.

The Countermeasure Knowledge Base accumulates information on countermeasures that corresponds to cyber risks. To describe information in the knowledge base, CYBEX introduces the Common Vulnerability Scoring System (CVSS), Common Weakness Scoring System (CWSS), Open Vulnerability and Assessment Language (OVAL), eXtensible Configuration Checklist Description Format (XCCDF).

The Product & Service Knowledge Base accumulates information on products and services. To describe information in this knowledge base, CYBEX introduces Common Platform Enumeration (CPE) and Common Configuration Enumeration (CCE). CPE provides a structured naming scheme for information technology systems, languages to describe it. These formats and languages are platforms, and packages, while CCE provides unique identifiers to system configuration issues to facilitate fast and accurate correlation of configuration data across multiple information sources and tools. That knowledge on cyber risks and countermeasures are often linked to specific products and services. For instance, a CVE is linked to CPE identifiers and CVSS scores in NVD.

VI. INFORMATION QUERY BLOCK

The block of a query of CYBEX represents specially developed programming language of the protected inquiries of CYBEX X.chirp. The specification allows to use additional fields for providing and monitoring of parameters of information security in standard language of inquiries of SQL.

VII. INFORMATION ASSURANCE BLOCK

According to methodology of CYBEX, three standards for certification of the protected sources are provided: X.evcert, X.eaa and ETSI TS 102042 V2.0. The standards X.evcert, X.eaa describe the algorithms similar to a digital signature, with that difference that this signature (digital certificate) can be appropriated to automatically entrusted source. The ETSI TS 102042 V2.0 standard provides the mechanism of existence of the organization for certification which can carry out recognition of certificates of the sites of the public Internet under the responsibility. It describes these requirements for certification authorities issuing public key certificates.

VIII. INFORMATION TRANSPORT BLOCK

The block of transportation of information is based on the protocols standardized by X.cybex-tp. The general description of transport protocols of the protected information transfer is provided in the X.cybex-beep specification. Besides, there are protocols described in the ETSI TS102232-1 standard. All these protocols of the protected information transfer on the public Internet use means of cryptography (enciphering). Length of a code depends on the importance of information parcel and on the class of security, CYBEX defined in the block of the description of information.

Albeit other protocols can be used for this transport, currently only the BEEP protocols are being investigated. Other candidate protocols, such as SOAP, exist but no draft recommendation for such protocols have been presented yet. From the viewpoint of forensics, ETSI TS102232-1 is also introduced here. This provides assurance of forensics information delivery to law enforcement and security authorities.

IX. INFORMATION SECURITY ORGANIZATION

In general process of ensuring information security is based on maintaining automatically updated databases so-called “the entrusted sources”. The mechanism of the automated certification of sources works. Contents of the site and a stream of inquiries from it are analyzed by the special software then the site is brought in a database with assignment to it the certificate of a safe source [12].

Thus, development completely specialized, including domestic operating systems is unfairly expensive and a little effective. It is expedient to use the software with an open code (OpenSource). This approach is used today by all large corporations for creation of the systems which are a trade secret. Using Open source software, it is possible to apply all available libraries to interfaces, input-output of data, etc. and to concentrate forces and means only on development critical, from the point of view of the developed technology, procedures which code, naturally reveals to nobody.

Thus, it turns out to cut down significantly terms and expenses on development demanded, having at the same time increased its reliability, thanks to a large number of already debugged fragments. At that we should also add that it is much easier to find the programmers owning the means relating to group of an open code comparing those who owns highly specialized knowledge. Thus, if we want to organize production of the professional software, but not “garage programming”, this way is preferable.

X. TIMING FOR CYBER SEQURITY

For the mechanism of digital certification in the protected networks, the synchronous time scale and a binding to it of all inquiries and answers are used. For ensuring unity and accuracy of this process it is expedient to use Timinator, the new modular server of time of production (KOMCET). This product is intended for work in future networks of mobile communication 5G and has to become one of the most effective remedies of high-precision synchronization of time and frequency with use of GLONASS/GPS/COMPAS/GALILEO together with the high-precision built-in generator and implementation of the PTP protocol.

XI. ACKNOWLEDGMENT

It is necessary to develop the uniform concept of ensuring information security of the general Internet taking into account requirements of the international standards and standards of the legislation of the Russian Federation. This concept has to include classification of objects of protection, importance of potential threats and methods ensuring the actions preventing threats of information security in total with the correcting influence in case of realization of this or that threat.

It is necessary to develop system of voluntary certification in the field of ensuring information security and national standards for this system.

XII. REFERENCES

[1] Overview of cybersecurity information exchange. Recommendation ITU-T X.1500. 04/2011.

[2] Takahashi T., Kadobayashi Y., Fujiwara H. Ontological approach toward cybersecurity in cloud computing // Proceedings of the 3rd International Conference on Security of Information and Networks. 2010. P. 100–109.

[3] Common vulnerabilities and exposures. Recommendation ITU-T X.1520. 01/2014.

[4] Common weakness enumeration. Recommendation ITU-T X.1524. 03/2012.

[5] Common platform enumeration. Recommendation ITU-T X.1528. 09/2012.

[6] Common vulnerability scoring system 3.0. Recommendation ITU-T X.1521. 03/2016.

[7] Common attack pattern enumeration and classification. Recommendation ITU-T X.1544. 04/2013.

[8] Language for the open definition of vulnerabilities and for the assessment of a system state. Recommendation ITU-T X.1526. 01/2014.

[9] Incident object description exchange format. Recommendation ITU-T X.1541. 09/2012.

[10] International Telecommunication Union. Information technology - Open Systems Interconnection - Procedures for the operation of OSI Registration Authorities: General procedures and top arcs of the International Object Identifier tree. X.660. August 2008.

[11] The World Wide Web Consortium (W3C). Resource Description Framework (RDF). 2010. URL: http://www.w3.org/RDF/.

[12] Tom C. Security and IoT in IEEE standards // IEEE-Standards, February 2016.

Experimental Study of Wireless Music And Speech Communication with the Use of Technology of Single-Wire Lines

Valery V. Frisk
Department of Theory of Electrical Circuits
Moscow Technical University of Communications and Informatics
Moscow, Russia
frisk@mail.ru

Abstract. **The paper presents the results of experimental studies on the wireless transmission of music and speech using the technology of single-wire lines. The ability to transmit energy, speech and music over a single-wire line allows this technology to be used to create a wireless transmission line.**

Keywords: **single-wire transmission line, wireless transmission line, step-up transformer, a step-down transformer, eddy current, mobile phone, microphone.**

I. INTRODUCTION

Theoretical and experimental studies have shown the importance of developing devices based on technologies using eddy currents [1–3]. At present, another technology using single-wire power lines has received growing attention [4; 5]. An alternating electric current is propagated along a single-wire line in a vortex way [6–11].

II. WIRELESS LINE FOR VOICE AND MUSIC

If a closed aluminum ring is placed at the end of a single-wire line, then an alternating eddy current is formed in it. Let us establish a second analogous ring near the first ring. It produces its own eddy current. We connect an additional single-wire line to the second ring. The vortex current from the second ring will propagate through this additional one-wire line. As a result, a wireless transmission line is formed using two single-wire lines (Fig. 1).

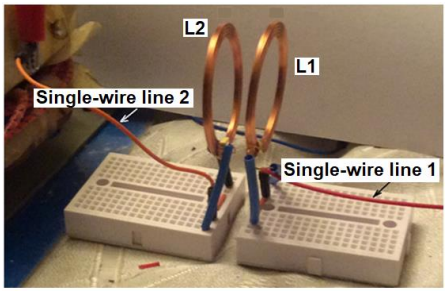


Fig.1. Wireless line for voice and music

III. EXPERIMENTALSETUPFORTRANSFERRINGSPEECHANDMUSICOVERAWIRELESSLINE

The experimental method of transmitting music over a wireless line is as follows. As a music source, we use a mobile phone (Explay A350 TV) in radio mode. The electrical signal is removed from the jacks to connect the headset. We will amplify this signal using an amplifier (TDA2005R). We connect the step-up transformer T1 to the output of the amplifier. We connect the first single-wire line to one of the pins of the high-voltage winding of step-up transformer T1. The second terminal of the transformer winding remains free. At the end of the first one-wire line, the first closed inductor L1 is installed. Near the first coil, we establish a similar second closed inductor L2. The distance between the coils is $x = 10$ mm. At the end of the second one-wire line, a step-down transformer T2 is installed. A second amplifier (TDA1562Q) is connected to the second winding of the step-down transformer T2. The BA1 speaker is connected to this amplifier. The wireless link is formed by two closed inductors L1 and L2 connected to two single-wire lines.

An experiment scheme for analyzing the transfer of music over a wireless link is shown in Figure 2.

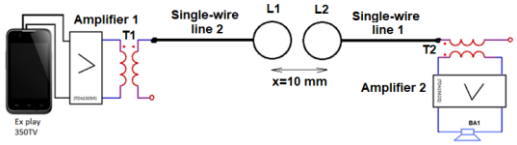


Fig. 2. Diagram for transferring speech and music over a wireless line

The operation of this wireless line is as follows. The signal is generated by the mobile phone tuned, for example, to a radio station. From the headphone outputs, the signal goes to the first amplifier, where it amplifies. The transformer-raising transformer T1 serves to excite eddy currents in a single-wire line up to the required value. As a result, in the first single-wire line, closed transverse eddy currents are excited and propagate, as was shown above. Obviously, a closed eddy current will also appear in the transmitting coil L1. This current forms an alternating magnetic field in the space near L1. In this field there is a closed receiving coil L2. A voltage arises that causes its eddy

current. This current propagates through the second single-wire line. The second transformer T2 lowers the voltage to the required level. The second amplifier amplifies the signal to the level of the desired speaker BA1. This experiment showed that the speaker BA1 quite well reproduced the music.

IV. VOICE OVER WIRELESS WITH MICROPHONE

The method of experimental research for the analysis of voice transmission over a wireless line is as follows., we use the microphone BM1 (Philips SBC MD150/00) instead of a mobile phone. To amplify the signal from the microphone BM1, we connect the microphone amplifier (BM137). For further amplification, the first amplifier (TDA2005R) is used. The voltage increases with the help of a step-up transformer T1. One first wire is connected to one end of the secondary winding T1. The second end of this winding remains free. The first closed inductor L1 is installed at the end of the first one-wire line. Near the first coil, we establish a similar second closed inductor L2. The distance between the coils is x = 10 mm. A step-down transformer T2 is installed at the end of the second one-wire line. A second amplifier (TDA1562Q) is connected to the second winding of the step-down transformer T2. The BA1 speaker is connected to this amplifier. The template is designed so that author affiliations are not repeated each time for multiple authors of the same affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization). This template was designed for two affiliations.

An experimental design for the analysis of voice transmission over a wireless link is shown in Figure 3.

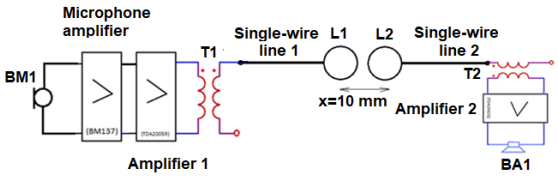


Fig. 3. Voice over wireless with microphone

This experiment showed that the speaker BA1 perfectly reproduces the speech transmitted through the microphone BM1.

As a result of this experiment, a voice signal was transmitted over the wireless line using two single-wire transmission lines.

V. REFERENCES

[1] Дорофеев А.Л. Вихревые токи. М.: Энергия, 1977. [Dorofeev A.L. Eddy currents. Moscow: Energy, 1977].

[2] Lammeraner J., Stafl M. Eddy Currents: Trans. from Czech. Praha, 1964.

[3] Миткевич В.Ф. Магнитный поток и его преобразования. М.; Л.: Изд-во АН СССР, 1946. 358 с. [Mitkevich V.F. Magnetic flux and its transformation. Moscow; Leningrad: Publishing House of the Academy of Sciences of the USSR, 1946].

[4] Sommerfeld A. Electrodynamik. Leipzig: Akademische Verlagsgesellschaft Oeest&Portig K.-G, 1949.

[5] Gubo G. Surface waves and their applications to transmission lines // J. Appl. Phys. 1950. Vol. 21. P. 1119.

[6] Стребков Д.С., Некрасов А.И. Резонансные методы передачи и применения электрической энергии. М.: ГНУ ВИЭСХ, 2008. [Strebkov D.S., Nekrasov A.I. Resonant methods of transmission and application of electrical energy. Moscow: GNU VIESH, 2008].

[7] Фриск В.В. Обнаружение индукционных токов в линии связи // Нелинейный мир. 2013. №8. С. 73–75. [Frisk V.V. Detection of induction currents in the communication line // Nonlinear world. 2013. N 8. P. 73–75].

[8] Frisk V.V. Researches of electric processes in a single-wire line and in the flat condenser // T-Comm. 2014. Vol. 8, N 1. P. 58–59.

[9] Фриск В.В. Индуктивные токи в однопроводной линии связи// Материалы Международной научно-технической конференции «INTERMATIC-2012», 3–7 декабря 2013, Россия, Москва, 2012 г. М.: МИРЭА, 2012. Ч. 5. С. 199–202. [Frisk V.V. Inductive currents in a single-wire communication line // Proceedings of the International Scientific and Technical Conference “INTERMATIC-2012”, December 3–7, 2013, Russia, Moscow, 2012. Moscow: MIREA, 2012. Part 5. P. 199–202].

[10] Фриск В.В. Вихревые токи в однопроводной линии // Труды конференции «Телекоммуникационные и вычислительные системы», 26 ноября 2014 г. М.: Информпресс-94, 2014. С. 107–109. [Frisk V.V. Eddy currents in a single-wire line // Proceedings of the conference “Telecommunication and Computational Systems”. 2014, November 2. Moscow: Informpress-94, 2014. P. 107–109].

[11] Frisk V.V. Speech transmission on a single-wire line // T-Comm. 2015. Vol. 9, N 7. P. 97–99.

Security of Wireless onboard Sensor Network

Andrey N. Petukhov[1], Pavel L. Pilyugin[2]

Department of Critical Information Infrastructures Security
Moscow Technical University of Communications and Informatics
Moscow, Russia
[1]anpetukhov@yandex.ru, [2]ppl@mail.ru

Dmitry A. Ovcharenko

Luxoft Professional LLC

Moscow, Russia

i-want-to-buy-violin@yandex.ru

Abstract. **The paper discusses the construction of a secure sensor network on board a modern civil aircraft using wireless communication technologies. The context of information security (threats, attack scenarios) and the specifics of this context for civil avionics are established. Revealed significant heterogeneity and multidimensional infrastructure on-Board wireless sensor network, the expediency of bringing complex solutions to mobile networks fifth generation (5G) and software defined networking (SDN) to solve security issues onboard wireless sensor network.**

Keywords: **information security, avionics, wireless network, sensor network, software-defined networks.**

I. INTRODUCTION

The existing communication systems of the aircraft, including operational systems on board, as well as various sensors for engines, chassis, range sensors, controllers of state parameters of various components and parts of the aircraft, require complex wiring and the manufacture of wiring harnesses, which adds weight to the aircraft and, in turn, increases fuel consumption. These systems have a well-defined, but still limited reliability and involve double or even triple redundancy, in order to reduce the risk of breakage or defective wiring.

At the same time, wireless technology is already enough to be seriously considered by the aviation industry for use as a means of communication within the aircraft in future generations of aircraft. This is evidenced, in particular, by the launch and development of the Wireless Avionics Intra-Communication project (WAIC) and the working groups of the Aerospace Vehicle Systems Institute (AVSI), which was created by the largest aerospace companies to solve common problems related to wireless avionics [1].

II. WIRELESS ONBOARD SENSOR NETWORK

The main advantages of using wireless onboard sensor communication network (WOSN) are the following:

- reduced volume and complexity of wiring and wiring harness designs with appropriate weight savings and improved overall fuel efficiency (total weight of wires reaches several tons plus up to 30% additional weight to attach the wiring harness to the construction);
- significant expansion of possibilities of reconfiguration of equipment and supplies due to the increased flexibility of installation, use of the mobile equipment;
- reliable control of parameters, moving or rotating units and parts of the aircraft;
- improving the reliability of on-board systems by reducing the level of failures of inconsistent processing of multiple signals of wired redundancy (double or triple) [2].

In addition, the solution ITU-R (WRC-15 Final act) for future WOSN systems defined frequency range 4.2–4.4 GHz [3] (previously it was used only for onboard altimeters). The dedicated frequency range (4.2–4.4 GHz) is designed for safety-related radio communications between avionics components integrated or installed exclusively on board the same aircraft.

Technical (hardware-software) complex of WOSN is characterized by the following defining properties [4]:

- the network is designed for radio communication between avionics components integrated or installed on board the same aircraft;
- radio communication is carried out in an isolated internal network between two or more points on the same aircraft;
- the network covers only applications related to flight control and safety support;

- the network does not support air-to-ground, air-to-satellite or air-to-air (other aircraft) communications;
- the network is not intended to be used for communication or entertainment of passengers in flight;
- network systems operate during all phases of flight, including on the ground;
- aircraft equipped with wireless on-board network systems is operated on a global basis and crosses national borders;
- signals to the wireless network is weakened by the fuselage of the aircraft.

Therefore, the main tasks to be solved in the design of wireless security solutions can be as follows:

- providing network access control, in order to determine which communication participants are on the network, what resources these participants can use and applying access control rules for traffic between these network nodes;
- ensuring the integrity and reliability of the wireless network segment as a critical resource;
- ensuring the security of data transmitted over the network, primarily through security protocols that use algorithms for authentication, encryption and traffic integrity control.

III. FEATURES OF WOSN SECURITY THREATS

When solving these problems and making practical decisions on the information security of the WOSN, it should be borne in mind that this system is operated in conditions, the features of which determine the specifics of the relationship to the threats and vulnerabilities that occur for wireless networks in General. So, on the ground the aircraft is in an area with a high level of physical security. Access to the airfield is limited and controlled, there are systems of control of movement and video surveillance, i.e. there are grounds to assert that the object of protection is in a controlled zone. This complicates the potential for an intruder to get close to the aircraft for the attack on the wireless network. Especially as on the earth, part of critical elements of a network can be inactive. Security measures carried out at the airport during the pre-flight period — baggage control and hand luggage inspection — also limit the possibility of the appearance on board of special equipment, and if it is still delivered there, its use will raise suspicions of the crew.

In flight, the attacker will be limited in time (the maximum duration of the flight in the calculations taken 18 hours), and it may not be enough to collect material and perform the attack. Repeated flight on the same plane will not help the attacker, as each flight can be accompanied by a change of static keys. In principle, it is possible to set the task to raise the algorithmic security to a level that will make inefficient cryptanalysis during the time of flight and thus get rid of many security problems, but most likely this decision will be unacceptably expensive, slow and cumbersome.

The least secure wireless traffic is during connection establishment and exchange. In some cases, in these moments, there may be short-term situations of practical vulnerability of the data transmission system. If you provide for the initialization of the stationary flight operation of the WOSN nodes and communications even before the appearance of passengers in the cabin, then you can deprive a potential attacker of the opportunity to take advantage of a temporary decrease in the level of network security.

Most WOSN nodes are located in fixed locations and have relatively stable activity characteristics. This creates the prerequisites for the use of information about the general stationary picture of electromagnetic radiation on board and makes it possible to quickly indicate the emergence of a new source of radio signal and localize its position in the cabin (short-range radiation sensors in passenger seats, toilets and in the vestibules of the cabin).

These circumstances indicate not so much a reduction in the level of threats to information security, but the specifics of their manifestations and attitude to them in the implementation of specific practical solutions to protect against these threats.

IV. VARIETY OF REQUIREMENTS AND CONDITIONS OF WOSN ELEMENTS

The choice of the underlying communication platform is significantly influenced by the architectural and infrastructural features of the WOSN. The proposed architecture WOSN consists of the following components:

- network node is a network object with the ability to connect and communicate with another network object using a radio interface; a network node can also provide one or more wired interfaces that allow you to interact with objects outside the WOSN;
- gateway node is a network node connecting the WOSN (or parts thereof) to other, usually wired on-board networks, such as the avionics communication network on board the aircraft;
- a leaf node is the network node capable of ensuring the connection between the gateway node and the sensor, actuator or display using the radio interface WOSN (physically the end node may contain the sensor, actuator or display) [2].

When discussing the requirements for the basic IOS communication platform and security features, it is advisable to take into account the data transfer rates and the installation location of the transceiver antennas of the network nodes (on the inner or outer surface of the fuselage). WOSN applications can be divided into two categories that meet the data rate requirements of applications: low-speed applications have data rates below 10 kbit/s, and high-speed applications have data rates above 10 kbit/s. The expected average data rate for low-speed applications inside the aircraft ranges from 10 bps (electrical consumption monitoring) to 800 bps (cabin pressure control) per line. The peak transmission rate on one channel can reach 1 kbit/s (control of electromagnetic radiation on board). For outdoor sensors, the spread of requirements is even greater: the average speed varies from 20 Mbit/s (external door sensors)

to 8 kbit/s (flight control sensors). For high-speed applications inside the aircraft, the expected average speed ranges from 12.5 kbit/s (FADEC engine management interface) to 1.6 Mbit/s (freeze frame in the cockpit or salon) and the peak data rate can reach 4.8 Mbit/s (predictive engine sensors). Transmission speed requirements for high-speed applications outside the aircraft body also vary by several orders of magnitude: for them, the average speed is expected to range from 45 kbit/s (sensors of construction) to 1 Mbit/s (external video cameras) [5].

The prospective WOSN topology is such that the radio transmission of the applications installed inside the aircraft structure and shielded from the outside is provided through wireless subnets, each consisting of a gateway node and one or more end nodes. Each compartment is equipped with at least one gateway node serving all end nodes in the coverage area of that gateway node. Signal attenuation caused by bulkheads or even interior furniture is usually too high for the gateway node to service the compartments, outside of where it is located. Small compartments, such as the cockpit or the electronics compartment, may require only one gateway Assembly, while a large compartment, such as a passenger salon, may require multiple gateway assemblies. For radio applications installed outside the aircraft fuselage and therefore not shielded from the outside, the antennas are installed in places where the attached gateway node can reach all the associated end nodes. For example, the gateway node antenna can be mounted on the top of the fuselage, from where the aircraft edges, wing tips, vertical and horizontal tail and all relevant sensors are viewed. In the other case, the antennas of the gateway units can be installed inside the wheel wells to connect only to the sensors located on the chassis.

Thus, WOSN is characterized by significant heterogeneity, fundamental differences in the operating conditions of the network nodes, configuration dynamics, a variety of schemes and connection routes. All this is accompanied by a gradation of the criticality of applications and infrastructure elements, assuming up to five different levels of qualification requirements of confidence in avionics [6] and, in particular, in its security [7].

V. PROMISING BASIC COMMUNICATION PLATFORM OF WOSN

One of the areas in which solutions are being developed to create a communication platform that meets these conditions is the concept of 5G, a new generation communication network. This concept provides for the creation of a heterogeneous network, which will use different technologies to serve traffic and users of different types on the basis of a combination of developing radio access technologies with new data transmission technologies. It is assumed that in 5G networks devices exchange multiple streams of information simultaneously with nodes of different types, whose task at a particular time is to service this particular device. Each device has its own policy of interaction with the network, taking into account the amount of data transmitted, the amount of allowable delay and other parameters, in particular security requirements. From the who perspective, this means a transition to a network model where the primary is the network node (gateway, the end node), not the base station.

The use of controlled antennas (SDR-technology) capable of changing the radiation pattern will not only increase the speed of data transmission due to the growth of the signal / noise ratio, but also significantly increase the protection of WOSN from jamming and unauthorized listening. To improve the efficiency of network resources use and reliability of its functioning, it is provided to allocate (physically and in time) resources for different types of traffic, and for each network fragment its own data transmission technology can be used (network slicing). Due to the flexibility of the approach, it is possible to fulfill the most diverse and even contradictory requirements of different types of network nodes.

For example, the use of a special slice (ultra-reliable low latency communication) for data transmission with a small delay concept will allow to transfer data with a very short duration of transmission. It will be impossible for the attacker on board the aircraft to intercept such a “radio shot” and, moreover, to simulate it in sync. Since a large number of low-power devices are seen as part of the WSON end nodes, solutions for a slice, which is intended for IoT, are of interest. Following the 5G concept, WSON will become a “layer cake” combining different technologies, the use of each of which will be determined depending on the current requirements of a particular node of the network.

A feature of the 5G solution complex, attractive from the point of view of WOSN security, is a possible implementation of the “device-to-device” technology. Within WOSN, network devices communicate in units or tens of meters from each other, and thanks to this technology, only signal traffic can pass through gateways and wired avionics segments, allows to establish and control connections between network nodes, and the data itself will pass directly between devices.

Today, there is a network technology, based on the separation of management and data traffic. This technology is the software-defined network (SDN). SDN is using special protocols for the interaction of control and transport layers, for instance OpenFlow.. The control layer elements (controllers) can be located in wired avionics segments, they set the current routing scheme, actually controlling the topology (and therefore security) of the network. Elements of the transport layer (switches) carry out packet data transfer between the nodes of the WOSN network (possibly physically combined with them), and create prerequisites for highly effective unification of “classic” network security solutions. In addition, from the point of view of reliability of alternative routing, SDN (capability redundancy) is functionally similar to the plurality of redundant wiring (resources redundancy).

VI. CONCLUSION

For the wireless implementation of software-defined networks, the software of the universal Wi-Fi controller Chandelle has already been created [8]. This controller allows you to manage software-defined networks with a large number of WiFi access points and helps to solve two problems of centralized management of access points in such networks. Firstly, the cost of both access points and controllers developed

by manufacturers that solved the problem of centralized access point management (Cisco, Aruba, Motorola, Juniper, Zyxel) is significantly reduced, and, secondly, the equipment of access points of one manufacturer is compatible with the controllers of another. During the pilot testing of the prototype device, 63 UAP-PRO Ubiquiti access points were operating under its control, providing simultaneous access to more than 9 thousand users (25 thousand connections in peak mode). From the point of view of WOSN security, Chandelle development is interesting with the following features:

- efficient algorithm for dynamic management of frequency-power resources of access points;
- integration with software-defined networks;
- ability to actively oppose external threats;
- detect and localize third-party radiation sources;
- change of access rights of communication participants depending on their location.

Therefore, the SDR & SDN network can be implemented for WOSN.

VII. REFERENCES

[1] Wireless communications for safetyrelated avionics // W. A. Intra-Communications. 2012. URL: http://waic.avsi.aero/.

[2] Technical characteristics and operational objectives for Wireless avionics intra-communications (WAIC): Report M.2197 (ITU-R Report), approved Nov. 2015.

[3] Technical conditions for the use of the aeronautical mobile (R) service in the frequency band 4200–4400 MHz to support wireless avionics intra-communication systems: Report ITU-R M.2283, approved July 2015.

[4] Technical characteristics and protection criteria for Wireless Avionics Intra-Communication systems: Recommendation ITU-R M.2067, approved Nov. 2014.

[5] Technical characteristics and spectrum requirements of Wireless Avionics Intra-Communications systems to support their safe operation: Report ITU-R M.2283, approved Dec. 2013

[6] DO-178C, Software Considerations in Airborne Systems and Equipment Certification. FAA. 2012.

[7] Technical standard for future airbornecapability environment (FACE™) edition 1.0. Open Group, Jan 2012.

[8] Monin S., Shalimov A., Smeliansky R. Chandelle: Smooth and Fast WiFi Roaming with SDN/OpenFlow // Open Network Summit 2014. Santa Clara, 2014.

Harmonization of Critical Information Infrastructure Objects Threats

Andrey N. Petukhov[1], Pavel L. Pilyugin[2]

Critical Information Infrastructures Security Division
Moscow Technical University of Communications and Informatics
Moscow, Russia
[1]anpetukhov@yandex.ru, [2] ppl@mail.ru

Karina N. Pilyugina

National Research University of Electronic Technology,

Moscow, Russia

karinakit@mail.ru

Abstract. **This paper discusses the issues of security threat modeling for critical information infrastructure facilities. An analysis is made of the applicability of various approaches to the construction of threat models highlighting the need for harmonization (ensuring internal compliance) of the main aspects of security, including threats. It justifies the expediency of using the "Common Criteria" methodology and SDL procedures for building harmonized threat models. The experience of using such an approach to create models of threats for software-defined networks and the protection profile of an SDN controller is discussed.**

Keywords: **information security, threat model, general criteria, protection profile, software configurable networks.**

I. INTRODUCTION

Critical objects countries may be different in different areas of activity, a general definition for them is not formulated; however, a number of common signs of such objects can be noted.

Critical objects of information infrastructure have a certain specificity from the point of view of security. In fact, it is impossible to justify the level of permissible residual risk, indeed, there is practically no such concept. In addition, there is no mechanism for calculating the real damage from such an information security incident. At the same time, the contingent of "recipients" of damage in this case is much wider than the set of security management entities (and, as a rule, it goes beyond the category of owners of compromised information assets).

In these circumstances, the goal of security management is not so much to achieve a certain level of security (to balance risks with costs), as to exhaust the potential of protection (to do all that is possible).

II. MODELING THREATS TO CRITICAL INFORMATION INFRASTRUCTURE OBJECTS

Despite this, it is necessary to have an adequate understanding of the threatening danger. In any particular case, there are features of the structure and form of such representation and a common element is always the typology of the manifestation of danger, the nomenclature of identified and qualified types of such manifestation, external (aggressiveness of the environment) and internal (imperfection of the object) events and situations that cause damage (threat model is the so-called proactive aspect of security management [1]).

Threat modeling is an essential element of information security management, and is a cyclical process activity.

The concept of threat representation hierarchy ("threat tree") put forward in the last century [2] is based on the schemes of successive (evolutionary) solutions for threat modeling. In this case, the attacker's target is represented as the "root node", and the potential means of achieving the target (scripts and resources) are represented as "end nodes". The resulting threat models formally define how you can implement threats from some classes ("roots") in each of the nodes of the threat tree. As a result, the use of hierarchical structures allows us to systematically consider the set of attack vectors against any specific target.

Another methodology for joint functional assessment (harmonization) of threats, assets and vulnerabilities is based on the well-known Clements—Hoffman model [3]. This model comes from the postulate that the security system must have at least one means to ensure security on every possible path of an attacker's impact on information technology ("a system with full overlap").

We can say that the model is "utopian" in nature, because it is almost impossible to build a system with full overlap. The reason for this is that the search for all impacts on the object cannot always be performed. In addition, in reality, each protection barrier provides only some degree of resistance to security threats. And it is important that it is the "resistance" of barriers that determines the amount of residual risk (this fact is especially important for critical objects of information infrastructure).

Nevertheless, it is the Clements—Hoffman model and the "threat tree" model that today determine the main vectors for the development of the methodology for modeling threats in

general and critical information infrastructure facilities in particular.

III. METHODOLOGY OF THREAT MODELING

We can say that the implementation of the threat modeling methodology mainly tend to several types:

- support for risk analysis (feasibility assessment methods);
- decomposition of subject (how the data flows and process flows);
- declarative (base model method);
- harmonization of high-level entities (common criteria methods)

Methods of the first type are a step-by-step process of attack modeling and threat analysis aimed at preparing data for risk analysis. This process involves the harmonization of security objectives and technical requirements for information processing and transmission procedures at each stage. As a result, the dynamic, adaptable and extensible identification of threats, the enumeration of their nomenclature and the procedure for assessing the feasibility are carried out. Some practical implementations allow for an understanding of the risks, taking into account the properties of attackers, allow you to prepare a specification and methods of application of the infrastructure, which can help mitigate the effects of malicious impact on specific information assets.

Another area of application of threat models as a risk management tool is the use of threat modeling results as input for security audits. In this case, the range of threats and their characteristics are formed on the basis of requirements that establish a certain, pre-permissible level of residual risk for each category of information assets. Obviously, this makes this approach inapplicable for critical information infrastructure objects.

Subject decomposition technologies are distinguished not so much by fundamental methodological innovations as by good tools and instruments for visualizing decomposition processes. This ensures a correct evolutionary transition from one level of threat modeling to another. Threat modeling begins with creating a visual representation of the analyzed application or infrastructure, and then this representation highlights its component parts, thus decomposing the subject of analysis. Once the decomposition is complete, the visual representation is used to identify and list potential threats.

One of the common ways to visualize a formal threat modeling process is the use of data flow diagrams (DFD) [5]. These diagrams were created to model information systems; DFD provided only four elements: data streams, data warehouses, processes that change data, and external factors of data change (interactors). However, the description of the threatening effects on the data within such a description has proven to be very effective. And when the ideology of trust calculation ("common criteria") began to prevail in the management of information security, the DFD procedure was supplemented with a special concept of the "boundary of trust", specifically for modeling threats.

First, using these basic concepts, the information infrastructure is described, then, the input of each threat is analyzed for all known categories. After identifying the level of aggression, this allows to establish measures to reduce the level of danger.

Declarative methods of forming threat models assume the presence of regulated information and procedural resources to support the various stages of modeling (the base model). Methods of this type use speculative threat systematization based on the assumption of the internal content of the threat. The low level of constructiveness of such a scheme, and its lack of relations to the used assets and information technologies characteristics, the general nature of expert assessments limit the use of declarative methods in real conditions.

The threats harmonization idea is implemented in the concept of information security management taking into account the requirements and conditions (for the object of management and for the environment) on the basis of an assessment of confidence in the means of implementation of these conditions and ways to meet these requirements ("calculation of trust"). The basic statement of this concept is the international standard reproduced in Russia [6]. One of the key provisions of the standard establishes a set and relationship of the original concepts ("high-level entities") in the field of information security. These relationships show that the essence of the "threat" interacts with the entities "risks", "assets", "threat agents" and "vulnerabilities".

Harmonization reveals these connections and allows you to include their content in the threat model, threats are modeled not by themselves, but in the context of interacting with these entities.

The standard provides a security profile. This is an effective format for interaction information usage. There are several categories in the protection profile which, in aggregate, exhaustively reflect the perceptions of the danger in a harmonized way. In addition to the actual category of "threats", a derived category of "assumptions" is attracted, with the help of which the information assets space is limited and the "boundary" interface with the environment is specified. The "assumptions" related to intrinsic properties and characteristics partly support the idea of the current security of an object (the reactive aspect of security management [1], the ratio of barriers and vulnerabilities of the Clements—Hoffman model [3]). The main tool for representing the reactive aspect in the protection profile is the category of "policy", which corresponds to the possibilities (realized and hypothetical) to counteract the harmful factors of an aggressive environment. To ensure the mutual correctness of "threats", "assumptions" and "policies" in the protection, profile provides a special mechanism, expressed in the form of "security objectives" for the environment and the object itself.

The interaction of threats and risks is a "projection" of the threat onto the space of security criteria. This projection indicates the criteria (properties) of security that "suffer" as a result of the realization of the threat. This indication, together with the qualification of the damage itself, identifies the risk. Therefore, the harmonization methodology implies an indication of the compliance of threats with established safety criteria.

Until recently, three main criteria (aspects, properties) of the security CIA (confidentiality (confidentiality), integrity (integrity) and availability (availability)) traditionally dominated. Detect threats that are not related to the criteria the CIA, demanded the extension of this list ("hexade Parker"), which also highlighted the authenticity, management, or possession (control) and usefulness.

It is important that additional criteria arose in the process of attempts to harmonize threats that were not correlated in the old criteria space with any vulnerabilities or risks. The examples and comments in GOST R ISO/MEK 27033-3-2014, deserving particular attention, illustrate the limitations of the CIA criteria triad and the inability to describe some real-life threats using its elements [7].

However, the application of this developed standard in the case of critical information infrastructure objects causes some difficulties. An information security incident is described by dividing the set of objects states into subsets of admissible and unacceptable states and the safety criteria change abruptly when the state of an object changes from one subset to another. Moreover, in the general case, security criteria can have the same value for object states from different subsets. Finally, one of the possible states of an object can be the cessation of its existence as a result of an incident, which in some cases makes any evaluation of safety criteria meaningless.

IV. SECURITY DEVELOPMENT LIFE-CYCLE

Among the approaches that harmonize threats with other high-level entities (vulnerabilities, assets and threat agents), the Security Development Life-cycle (SDL) methodology attracts attention [8]. In the context of the topic of security of critical objects, the construction of diagrams and the enumeration of threats are of interest.

When constructing diagrams, they usually use the tools for constructing DFD data flow diagrams (including the "trust boundary" element). The element "boundary of trust" shows that the elements located on opposite sides of this boundary function at different levels of authority.

For enumeration of threats in the SDL, the "STRIDE threats to the element" method is used from some universal list:

- **S**poofing of user identity (spoofing subject);
- **T**ampering (intervention and modification);
- **R**epudiation (disclaimer);
- **I**nformation disclosure (leakage and disclosure);
- **D**enial of Service;
- **E**levation of privilege (capture and elevation of privilege) [9].

The methodology assumes that all threats can be grouped into groups from the STRIDE list, and that each type of DFD element corresponds to a specific threat class. For example, this correspondence has the following form:

- External element-SR (Substitution-Disclaimer);
- Process-STRIDE (all kinds of threats);
- Data storage-TRID (Unauthorized Access; Data Leak; Denial of Service);
- Data flow-TID (Unauthorized Access; Data Leak) [8].

Each SDL implementation includes guidance on detailing the threats of each class for each type of DFD element. Typically, these guides have two methodologically inseparable parts.

The first is relatively general and stable and reflects (harmonizes) the specifics of the pair "threat class STRIDE — type element DFD". The most authoritative sources of this part of the manuals (Microsoft, Cisco, IEEE 802, etc.) publish their position on how threats from each STRIDE class are manifested in relation to DFD-elements of different types.

The second part of the SDL threat detail guides applies to a specific DFD element. It represents the rules of attack scenario development and describes the scheme and conditions of threat actions. It also points to specific vulnerabilities that allow you to implement an attack scenario. Thus, the harmonization of threats is carried out.

The knowledge necessary for targeted management should contain, at a minimum, an inventory of information assets (resources and processes), information about the aggressive potential of the environment and the channels for its implementation, as well as an idea of the real possibilities of countermeasures (i.e. "what, from what and what we protect"). In the case of protection of critical objects, vulnerabilities can be methodically subordinate, because in critical objects the vulnerability can be only in three states: already eliminated, eliminated (short-term state) and unknown. Therefore, the harmonization of threats in their modeling for critical information infrastructure is aimed at identifying unknown threats.

In general, all possible combinations of asset-actor-action triples are built. Combinations that are not meaningful or allowed are discarded. The array of data associated with forbidden combinations is an ideal information base for security management, and a threat model.

The methodical decisions in the SDL do not directly operate with the category of damage and do not pursue the goal of minimizing residual risk. They are aimed at achieving the completeness of the considered threats, thereby creating the prerequisites for the exhaustion of the potential of protective actions. At the same time, the SDL fully preserves the "harmonizing" properties, making it possible to consider the manifestations of a threat in the context of a specific element of the information technology (asset), in specific conditions and with consideration of a specific source (actor). In addition, SDL has "evolutionary" properties that allow the decomposition of interacting entities, refining the threat (attack) implementation scenarios, without disturbing the harmonization. Thus, we can conclude about the suitability and feasibility of SDL techniques for modeling security threats to critical information infrastructure facilities.

V. CONCLUSION

The approach using the SDL procedures was applied in the analysis and development of security solutions for critical objects built on the basis of software-defined network technology (SDN) [10]. A protection profile was developed in MTUCI for SDN-controller, the core of the control layer of such a network [11].

A fundamental feature of software-defined networks is the separation of network services for management functions and packet forwarding functions. This creates new classes of attacks on the control center (controller) both from the data plane and in the control plane. Analysis of the structure of devices and the use of the STRIDE model to identify threats allowed us to formulate a general description of the threats. Threats to the controller are considered across all possible interfaces: from applications, control systems, and network devices. These assumptions actually determine and maintain the composition of key assets of SDN-controllers: account data, configuration and management data, log data, operating system, software, hardware, resources and controller interfaces.

This made it possible to prepare specifications of specific threats, assumptions and policies united by common security objectives, both for the object (SDN-controller) and for the environment of its operation. As a baseline, the mode of using the network is considered, when all network infrastructure, computer equipment, communication equipment and communication channels are under single control in the trusted zone, all applications are trusted and provided by the provider. This option is typical for deploying a software-defined network in a single center.

VI. REFERENCES

[1] Петухов А.Н. Информационная база управления кибербезопасностью критических инфраструктур // XI международная научная конференция «Технологии информационного общества (Москва, 15–16 марта 2017) . М., 2017. [Petukhov A. N. Management information base of cybersecurity // XI International Scientific Conference “Information Society Technologies” (Moscow, March 15-16, 2017). Moscow, 2017].

[2] Salter C., Sami Saydari O., Schneier B. et al. Toward A Secure System Engineering Methodology. Washington: National Security Agency, 1998..

[3] Хоффман Л.Дж. Современные меторды защиты информации / Пер. с англ. М.: Сов. радио, 1980. [Hoffman L. Modern methods for computer security and privacy. Moscow: Sovetskoye radio, 1980].

[4] Bundesamt für Sicherheit der Informationstechnik URL: http://www.bsi.de.

[5] Abi-Antoun M., Wang D., Torr P. Checking Threat Modeling Data Flow Diagrams for Implementation Conformance and Security // ASE'07. Atlanta, 2007.

[6] ГОСТ Р ИСО,МЭК15408-1—2012 Информационная техонология (ИТ). Методы и средства обеспечения безопасности. Критерии оценки безопасности. Критерии оценки безорпасности информационных технологимй. Часть 1. Введение и общая модель. М., 2012. [GOST R ISO/MEK 15408-1-2012. Information technology. Security techniques. Evaluation Criteria for IT security. Part 1. Introduction and general model. Moscow, 2012].

[7] ГОСТ Р ИСО/МЭК 27033-3-2014 Информационная технология (ИТ). Методы и средства обеспечения безопасности. Безопасность сетей. Часть 3. Эталонные сетевые сценарии. Угрозы, методы проектирования и вопросы управления. М., 2014. [GOST R ISO/MEK 27033-3-2014 Information technology. Sewcurity techniques. Network security. Part 3. Reference network scenarios. Threats, design methods and management issues. Moscow, 2014].

[8] Shostack A. Threat Modeling: Designing for Security. Indianapolis: John Wiley & Sons Inc., 2014.

[9] The STRIDE Threat Mode / Microsoft. 2016.

[10] Pilyugin P., Smelyansky R. Modern security issues in SDN // Proceedings of ITSN-2017 International Conference on Information Technologies, Systems and Networks 2017, Chisinau, Republic of Moldova. Chisinau, 2017. P. 182–187.

[11] Петухов А.Н., Пилюгин П.Л. Профили защиты для программно-конфигурируемой среды // Радиоэлектронные устройства и системы для инфокоммуникационных технологий REDS 2018. М., 2018. [Petukhov A.N., Pilyugin P.L. Security profiles for software defined network. International Conference on: The radio-electronic devices and systems for infocommunication technologies. REDS-2018. Moscow, 2018].

Asymptotic Enumeration of Binary Orthogonal Arrays*

Konstantin N. Pankov
Department of Postgraduate Studies
Moscow Technical University of Communications and Informatics
Moscow, Russia
pankov_kn@mtuci.ru

***Abstract*. Orthogonal arrays are basic combinatorial structures, which appear in various cryptographic applications designed to ensure information security in electronic financial services. Asymptotic estimates for the number of orthogonal arrays are proved in this paper.**

***Key words:* cryptography, authentication codes, orthogonal arrays, asymptotic enumeration.**

I. INTRODUCTION

It was proposed launching a large-scale system-wide program to develop an economy of a new technological generation, the so-called digital economy, in the President Putin annual address to Federal Assembly in December 2016 [1]. In compliance with the list of instructions of the President for the implementation of the Address in 2017, the Ministry of Communications and Mass Media of Russia prepared the program "Digital Economy of the Russian Federation" [2]. The program was approved by the Government of the Russian Federation in its resolution No. 1632-r dated July 28, 2017. The Sub-commission for digital economy of the governmental commission for the use of information technologies to improve the quality of life and the conditions of doing business was established by the Government of the Russian Federation in its resolution No. 969 dated August 15, 2017. Maxim Akimov, the Deputy Prime Minister of the Russian Federation, heads up the Sub-commission.

According to the roadmap of the program in the field of regulatory regulation, the task is to ensure the legal conditions for the introduction and use of innovative technologies in the financial market. To accomplish this task, it is supposed to promote the introduction and use of innovative technologies in the financial market, including improving the mechanisms for providing financial services in electronic form and ensuring their information security.

The most powerful tool for ensuring information security in the modern world is the use of cryptography. Cryptographic tools have become an integral part of the services provided in electronic form by financial and credit institutions. Cryptographic algorithms based on the use of mathematical methods for transforming protected information solve the problem of ensuring confidentiality, integrity, authentication, non-repudation and untraceability.

In ISO 27000 [3], confidentiality is the property, that information is not made available or disclosed to unauthorized individuals, entities, or processes. Confidentiality is a component of privacy that implements to protect data from unauthorized viewers. Data integrity means maintaining and assuring the accuracy and completeness of data over its entire lifecycle. This means that data cannot be modified in an unauthorized or undetected manner. Authentication is the confirmation of the authenticity of various aspects of information interaction: the content and source of transmitted messages, communication session, interaction time, etc. It is an important part of the problem of ensuring the reliability of the information received. This problem is especially acute in the case of non-trusting parties, when not only the third party (adversary-enemy) can serve as a source of threats, but also the party with whom the informational interaction takes place (adversary-violator). In law, non-repudiation implies one's intention to fulfill their obligations to a contract. In cryptographic protocol it also implies that one party of a transaction cannot deny having received a transaction, nor can the other party deny having sent a transaction. And, finally, untraceability is a property meaning that it is impossible for an adversary to obtain information about the actions of the protocol participants.

Users of electronic financial services must be sure of the authenticity of messages arriving at their address. This is part of the task of authentication, which plays an important role, for example, in the organization of machine-to-machine interaction on the Internet of things and in the implementation of a private blockchain network. The solution to this problem is not generally provided by the use of encryption systems designed to perform the task of confidentiality. It follows from paper [4]. To protect against active adversary attacks, the message is supplied with a tag or message authenticity code. Information authentication systems are not required to maintain confidentiality, although in some systems both of these tasks are performed [5, chapter 14].

II. DEFINITIONS

The mathematical model of authentication systems is an authentication code or A-code, originally proposed in [6] and discussed in detail in the works of A.Yu. Zubov (for example. [5; 7]).

A Cartesian authentication code or systematic authentication code is a four-tuple $(\mathcal{S},\mathcal{T},\mathcal{K},\{E_k : k\in\mathcal{K}\})$ where $\mathcal{S}$ is the source state space associated with a probability distribution, $\mathcal{T}$ is the tag space, $\mathcal{K}$ is the key space associated with a probability distribution, and $E_k : \mathcal{S}\to\mathcal{T}$ is called an encoding rule. A transmitter and a receiver share a key k for authentication purpose. If the transmitter wants to send a source state $s\in\mathcal{S}$ to the receiver, he computes $t = E_k(s)$, $t\in\mathcal{T}$, and sends the message $m=(s,t)$ to the receiver through a public communication channel. When receiving $m'=(s',t')$, the receiver will compute $E_k(s')$ and checks whether $t' = E_k(s')$. If it does, the receiver will accept it as authentic. Otherwise, the receiver will reject it [8].

One of the mathematical objects associated with A-codes are orthogonal arrays. In [9], the design of the authentication code using orthogonal arrays, which is resistant to impersonation and substitution attacks, is described in detail.

Now we shall give the following definition. An orthogonal array $OA_\lambda(t,k,v)$ is a $\lambda v^t \times k$ array whose entries are chosen from a set X with υ points such that in every subset of t columns of the array, every t-tuple of points of X appears in exactly λ rows. These parameters are given the following names: υ is the number of levels, k is the number of factors, t is the strength, and λ is the index. An orthogonal array is simple if it does not contain any repeated rows [10]. This combinatorial structure was first introduced by an Indian statistician C. R. Rao [11] for use in design of experiments.

For cryptography applications the most frequently used orthogonal arrays are those with all factors at two levels, which we usually refer to as 2-level or binary orthogonal arrays, i.e. let $v=2$. It is binary orthogonal arrays that were used to construct authentication codes in [9].

This is example of a binary orthogonal array – orthogonal array $OA_3(2,11,2)$ from [10]:

$$\begin{pmatrix}
0&0&0&0&0&0&0&0&0&0&0\\
1&1&1&0&1&1&0&1&0&0&0\\
0&1&1&1&0&1&1&0&1&0&0\\
0&0&1&1&1&0&1&1&0&1&0\\
0&0&0&1&1&1&0&1&1&0&1\\
1&0&0&0&1&1&1&0&1&1&0\\
0&1&0&0&0&1&1&1&0&1&1\\
1&0&1&0&0&0&1&1&1&0&1\\
1&1&0&1&0&0&0&1&1&1&0\\
0&1&1&0&1&0&0&0&1&1&1\\
1&0&1&1&0&1&0&0&0&1&1\\
1&1&0&1&1&0&1&0&0&0&1
\end{pmatrix}$$

Correlation-immune functions were first introduced by Siegenthaler [12] as a class of suitable functions for combining the outputs of several linear feedback shift registers. They lead to the construction of running-key generators for stream ciphers which resist the correlation attack. The equivalence between t-th order correlation-immune binary function $f(x_1,...,x_k)$ with weight λv^t and simple binary orthogonal arrays $OA_\lambda(t,k,v)$ was proved in [13]. It follows that the study of orthogonal arrays is closely related to the cryptosyntesis and the cryptoanalysis of encryption systems.

III. RESULTS ABOUT ENUMERATION OF ORTHOGONAL ARRAYS

Enumeration of orthogonal arrays have been intensively investigated in [14–18]).

In [14] the enumeration of binary orthogonal arrays is studied, and a closed expression for the enumeration of binary orthogonal arrays of strength 1 is given using the inclusion – exclusion principle and the edge-induced subgraph.

In [15] there was specified an algorithm to enumerate a minimum complete set of combinatorially non-isomorphic orthogonal arrays of given strength t, run-size N, and level-numbers of the factors. The algorithm was the first one handling general mixed-level and pure-level cases. Using an implementation in C, there was generated most non-trivial series for $t = 2$, $N < 29$, $t = 3$, $N < 65$, and $t = 4$, $N < 169$. The exceptions defined limiting run-sizes for which the algorithm returns complete sets in a reasonable amount of time.

In [16] there was given an application of orthogonal arrays to construct D-optimal discrete choice experiments. Also there was addressed the existence problem of orthogonal arrays and described new algorithms to enumerate all orthogonal arrays with given parameters. Key techniques include eliminating significant space requirements, using previously-computed information on substructures in an efficient way. Computational results show these algorithms to be significantly faster than the prior state-of-the-art. Also there was defined the linear independence of Latin squares and stated a basis for a linear space which contains all Latin squares. Furthermore, there was used this theory to study an asymptotic formula for the number of orthogonal arrays with three columns.

In [17] the construction and enumeration of mixed orthogonal arrays (MOA) are described to produce optimal experimental designs. A MOA is a multiset whose rows are the different combinations of factor levels, discrete values of the variable under study, having very well defined features such as symmetry and strength three (all main interactions are taken in consideration). The applied methodology blends the fields of combinatorics and group theory by applying the ideas of orbits, stabilizers and isomorphisms to array generation and enumeration. Integer linear programming was used in order to exploit the symmetry property of the arrays under study. The backtrack search algorithm was used to find suitable arrays in the underlying space of possible solutions. To test the performance of the MOAs, an engineered system was used as a case study within the stage of parameter design. The analysis showed how the MOAs were capable of meeting the fundamental engineering design axioms and principles, creating optimal experimental designs within the desired context.

In [18], a method is described for the construction of all geometrically non-isomorphic ternary orthogonal arrays in 18

runs. One representative from each geometric isomorphism class is provided in the electronic appendix of [18]. The approach that have taken in this paper can be extended in principle but the problem rapidly becomes very large. For instance, when increasing the number of levels from 3 to 4, and considering 32 runs, the number of incidence matrices for two factors that needs to be considered to determine all possible valid third columns is 282, and the number of valid third columns is 22,695, none of which are geometrically isomorphic. If we keep ternary factors but instead increase the number of runs to 27 then the number of incidence matrices is 55, giving 847 valid triples of incidence matrices, and 424 valid third columns. To extend this to approach to larger m would require the consideration of $6^9 \times 847$ potential fourth columns. For $N = 36$ there are 120 incidence matrices, 3921 valid triples corresponding to 1971 possible third columns.

In addition to these papers, there are others devoted to the topic of enumeration of orthogonal arrays.

Using the results of [13; 19], we can estimate the asymptotic number of simple binary orthogonal arrays.

Let $|OA_\lambda(t,k,2)|$ be the number of simple binary orthogonal arrays $OA_\lambda(t,k,v)$. Hereafter, we use the following notation: $\exp_2(x) = 2^x$. Using theorem 3.1 [13] and theorem 4 [19], we get the following:

Theorem 1. Suppose $0 < \varepsilon < \frac{1}{2}$ and $t(5+2\log_2 k) \le k\left(\frac{1}{2}-\varepsilon\right)$ for sufficiently large k; then

$$|OA_\lambda(t,k,2)| = \theta_1(\lambda,t)\exp\left(2^k \ln 2 - 0{,}05\cdot\exp_2\left(2^{k\varepsilon-\log_2 k}\right)\right) +$$

$$+\exp_2\left(2^k - (k-t)\binom{k}{t} - \left(\log_2\sqrt{\pi/2}\right)\sum_{i=0}^{t}\binom{k}{i}\right)\times$$

$$\left(\exp\left(-2^{1-k}\left(2^{k-1}-\lambda 2^t\right)^2\right)\times\right.$$

$$\left.\times\left(1+\theta_2(\lambda,t)k^{-2}\right)+\theta_3(\lambda,t)\exp\left(-0{,}4\cdot 2^{k\varepsilon+3t-\log_2 k}\right)\right),$$

where $|\theta_1(\lambda,t)| \le 1$; $|\theta_2(\lambda,t)| \le 3{,}62$; $|\theta_3(\lambda,t)| \le 0{,}8$.

Let $|OA_\vee(t,k,2)|$ be the number of all simple binary orthogonal arrays $OA_\lambda(t,k,v)$ with any index λ. From the proposition 3 [19], it is clear that the following theorem is true.

Theorem 2. If $0 < \varepsilon < \frac{\ln 2}{4}$ and $t < \frac{k}{\ln k}\left(\frac{\ln 2}{4}-\varepsilon\right)$ for sufficiently large k then, as $k \to \infty$ $|OA_\vee(t,k,2)|$:

$$\exp_2\left(2^k - \frac{1}{2}\left((k-t)\binom{k}{t}-k\right) - t - \left(\log_2\sqrt{\pi/2}\right)\sum_{i=1}^{t}\binom{k}{i}\right).$$

IV. CONCLUDING REMARKS

In conclusion, we can add that orthogonal arrays have found numerous application in cryptography. Partial de-randomization of randomized algorithms, for example, the Monte Carlo algorithm used in cryptanalysis [20], secret sharing schemes, universal hash functions, perfect local randomizers [21] are among their applications [22].

V. REFERENCES

[1] Ежегодное послание Федеральному собранию. 01.12.2016. [Presidential annual address to Federal Assembly. 01.12.2016].. URL: http://www.kremlin.ru/events/president/news/53379/.

[2] Об утверждении программы «Цифровая экономика Российской Федерации». 31.07.2017. [On approval of the program “Digital Economy of the Russian Federation”. 31.07.2017]. URL: http://government.ru/docs/28653/.

[3] ISO/IEC 27000:2009 Information security management systems — Overview and vocabulary. URL: http://pqm-online.com/assets/files/lib/std/iso _iec_27000-2009.pdf.

[4] Месси Дж. Л. Введение в современную криптологию // ТИИЭР. 1988. Т. 76, N 5. С. 24–42. [Messi J.L. Introduction to modern cryptology // Proceedings of the Institute of Electrical and Electronics Engineers. 1988. Vol. 76, N 5. P. 24–42].

[5] Зубов А.Ю. Математика кодов аутентификации. М.: Гелиос АРВ, 2007. [Zubov A.Yu. Mathematics of authentication codes. Moscow: Helios ARV, 2007].

[6] Gilbert E.N., MacWilliams F.J., Neil J. et al. Codes which detect deception // Bell System Technical Journal. 1974. Vol. 53, N 3. P. 405–424.

[7] Зубов А.Ю. Коды аутентификации. М.: Гелиос АРВ, 2017. [Zubov A.Yu. Authentication Codes. Moscow: Helios ARV, 2017].

[8] Ding C., Helleseth T., Klove T. et al. A Generic Construction of Cartesian Authentication Codes // IEEE Transactions on Information Theory. 2007. Vol. 53, N 6. P. 2229–2235.

[9] Тарапников Ю.В. Комбинаторные свойства дискретных структур и приложения к криптологии. М.: МССМЕ, 2011 [Tarannikov Yu.V. Combinatorial properties of discrete structures and applications to cryptology. Moscow: MCCME, 2011].

[10] Sloane N.J.A., Hedayat A.S., Stufken J. Orthogonal Arrays: Theory and Applications NY: Springer, 1999. (Springer Series in Statistics).

[11] Rao C.R. Factorial Experiments Derivable from Combinatorial Arrangements of Arrays // Supplement to the Journal of the Royal Statistical Society. 1947. Vol. 9, N 1. P. 128–139.

[12] Siegenthaler T. Correlation-immunity of nonlinear combining functions for cryptographic applications // IEEE Trans. Information Theory. 1984. Vol. 30. P. 776–780.

[13] Camion P., Carlet C., Charpin P. et al. On Correlation-immune functions // Advances in Cryptology — CRYPTO ’91. CRYPTO 1991. Lecture Notes in Computer Science.Vol. 576. Berlin; Heidelberg: Springer, 1992. P. 86–100.

[14] Zhang J.-Z., You Z.-S., Li Z.-L. Enumeration of binary orthogonal arrays of strength 1 // Discrete Mathematics. 2001. Vol. 239, N 1–3. P. 191–198.

[15] Demirkale F.Orthogonal Arrays; Enumeration and Applications: PhD Thesis / The University of Queensland: School of Mathematics and Physics. 2013. https://espace.library.uq.edu.au/view/UQ:310823/s42201049_phd_finalthesis.pdf.

[16] Schoen E.D., Eendebak P.T., Nguyen M.V. Complete enumeration of pure-level and mixed-level orthogonal arrays // Journal of Combinatorial Designs. 2010. Vol. 18, N 2. P. 123–140.

[17] Romero J. Enumeration of strength three orthogonal arrays and their application in parameter design: PhD Thesis / University of Canberra; Faculty of Education, Science, Technology & Maths. 2017.

http://www.canberra.edu.au/researchrepository/file/19e87ebf-9906-42a6-8c68-a72e4631bf28/1/full_text.pdf.

[18] Bird E.M., Street D.J. Complete enumeration of all geometrically non-isomorphic three-level orthogonal arrays in 18 runs // Australasian Journal of Combinatorics. 2018.Vol. 71, N 3.P. 336–350.

[19] Панков К.Н. Улучшенные асимптотические оценки для числа корреляуионно-иммунных и _эластичных двоимчных вектор-функций // Дискретная математика. 2018. Т. 30, № 2. С. 73–98.[Pankov K.N. Improved asymptotic estimates for numbers of correlation-immune and (n,m,k)-resilient vectorial boolean functions // Discrete Mathematics. 2018. Vol. 30, N 2. P. 73–98].

[20] Заикин О.С., Семенов А.А. Применение метода Монте-Карло к прогнозированию времени параллельного решения проблемы булевой выполнимости // Вычислительные методы и программирование. 2014. Т. 15, №1.С. 22–35. [Zaikin O.S., Semenov A.A. Application of the Monte Carlo method for estimating the total time of solving the SAT problem in parallel // Computational Methods and Programming. 2014. Vol. 15, N 1. P. 22–35].

[21] Maurer U.M., Massey J.L. Local Randomness in Pseudo-random Sequences // Journal of Cryptology. 1991. Vol. 4, N 2. P. 135–149.

[22] Gopalakrishnan K., Stinson D.R.Applications of orthogonal arrays to computer science // Proc. of ICDM. 2006. P. 149–164,

IT Students about Risks and Security of Industrial Internet of Things

Marina S. Chvanova[1], Alexey E. Popovich[2]
Department of Information Systems and Technologies
K.G. Razumovsky Moscow State University of technologies and management (the First Cossack University)
Moscow, Russia
[1]tmbtsu@gmail.com, [2]popovichae@mgutm.ru

Maria S. Anurieva[3], Anastasia V. Frolova[4], Karapet A. Alekyan[5], Dmitry A. Makarov[6]
Department of Mathematical Modeling and Information Technologies
G.R. Derzhavin Tambov State University
Tambov, Russia
[3]anuryeva@mail.ru, [4]fr.anastasia.vit@gmail.com, [5]alekyan188@gmail.com

***Abstract*. The article provides a brief overview of the state and prospects of Industrial Internet of Things (IIoT) development in Russia and all over the world. Presents the results of survey of IT student youth to understand the problems of IIoT, its risks and security. The survey showed that the awareness of the youth audience about the risks and safety of IIoT is insufficient, and the competencies are largely intuitive. The content of training in this field is not systematically formed due to the novelty of the subject matter for Russia.**

***Keywords:* IIoT, IoT, data security.**

I. INTRODUCTION

One of the most important areas in the field of information technology nowadays is the development of the "Internet of Things", a system of integrated computer networks and connected physical objects (things) with embedded sensors and software for data collection and exchange, with the possibility of remote monitoring and control in an automated mode, without human intervention [1]. The B2B direction of industrial Internet of things is of particular interest. This includes M2M solutions, big data, cloud, robotics, etc. From the point of view of its application, IIoT can be used in all sectors: agriculture, transport industry, financial sector, extractive industries, urban infrastructure, medicine, etc. [2]. This is a completely new form of work organization and business models of service delivery. According to the most progressive scenario, the implementation is fully digitized and automated production is controlled by intelligent systems in real time mode, without human intervention, going beyond the boundaries of one enterprise, with the prospect of uniting into a global industrial network of things and services [1]. With regard to dynamic development of this area, it is important to pay attention to the training of computer science students in this filed, since it has its own characteristics and risks. To this end, it seems necessary to understand the state of the situation by interviewing the students themselves for information and the formation of individual ideas. This is useful because of the need to choose a further strategy for training students.

II. RESEARCH

Industrial Internet of Things (IIoT) in Russia and all over the world. Review of the state and development prospects.

2011 is considered the starting year of the development of the Internet of things industry, it was the year when the number of connected physical objects in the world exceeded the number of connected people [2]. One of the key factors that gave impetus to the development of the Internet of things in Russia is state interest. The "Digital Economy of the Russian Federation" program, which was approved in the summer of 2017, has adjusted the industry and the public sector to digitalize.

According to Global Market Insights, the global IIoT market in 2017 reached $ 312.79 billion. During the period from 2017 to 2023 it will grow at an average annual rate of 14.36%. By 2023, its volume will be 700.38 billion dollars. According to forecasts of another agency, Machina Research, by 2025 the global market for industrial Internet of things (equipment, including sensors, software and platforms, services) will reach 484 billion euros [3].

Overall production growth will be a powerful accelerator of the IIoT market in Russia. Experts say that tendencies to this growth against the background of Western sanctions can be traced now [4]. However, there are also constraints to the development of the Internet of things: state of the economy, sanctions, lack of specialized networks for IIoT, lack of investment, lack of specialists, low level of production automation.

According to TAdviser estimates, the Russian IIoT market in 2017 amounted to 93 billion rubles and is expected to grow to 270 billion by 2020. Industrial enterprises, focused in our research, provided about 20% of the market volume (their share will grow to 25% by 2020). The degree of the IIoT penetration in Russia depends largely on the level of state support. Technological solutions replenish the market, and from both international and Russian suppliers, but customers still lack the use of technology scenarios that produce tangible results [5].

Experts call promising ways to use the IIoT in industrial enterprises: the ability to implement complex end-to-end, fully automated business processes, remote monitoring and on-time

maintenance, as well as the provision of new service business models. At the same time, according to IBM research, studying the use of IoT will help global organizations to prevent incidents and increase employee safety. Data collected from sensors is combined with innovative cognitive capabilities and indicators obtained from other external sources (for example, meteorology) [3].

The role of student youth awareness of IIoT for dynamics of its development.

It is important to organize targeted training of students in the field of IIoT. Thus, IIoT Samsung Academy held summer educational program on the Internet of Things, where students gained theoretical knowledge in the field of the Internet of things, as well as practical experience in creating prototypes of IIoT solutions. Participation was attended by 15 students who passed through competitive selection process and 8 teachers from MIPT and MIREA. The program was held in July 2018 on the basis of the Fiztech School of Applied Mathematics and Computer Science [6].

This shows that the state is interested in training of highly qualified personnel in the field of information technologies that can develop, implement and maintain a system of any scale, complexity and functionality.

Based on the above, it is reasonable to conduct a survey of users on the Web, which can help identify awareness of IIoT, risks and security of these technologies.

Student youth survey for understanding IIoT issues, risks and safety.

The general population of our research is made up of Internet users in Russia, who considered as a homogeneous population, represented on the basis of "access to the Internet and the use of Internet resources". We have previously considered questions of features and properties of youth communication in the Internet space [7]. In this study, during the process of developing the topic of communication with Internet users, we conducted an online survey, using the target probability sample. A random selection of respondents from the general population was made by sending out invitations to participate in an online survey and posting a link in social networks to online questioning. Thus, the so-called "method of self-selection" was used. As a result, the total number of respondents was 153 people. The time of the survey is October 2018.

The questionnaire is based on Google Forms cloud technologies. In this case, the object to subject field defines Internet users as the general population and the sample population is composed of this category of citizens, which allows to say about the sufficient objectivity of collected data.

The overwhelming majority of those surveyed are young people between the ages of 19–25 years (76.5%). Only 7.8% of respondents aged 26–35 years took part in the study. The survey showed that the majority of respondents live in cities (71.4%). A small proportion of respondents are from villages (12.2%) and urban-type settlements (10.2%). Based on the processed data and taking into account the goal of the study on maximum indicators, we can say that the hypothetical social portrait of the respondent in our sample is represented by a young man from 19 to 25 years old who lives in a city.

Most respondents are familiar with the term IIoT, it is clear that the respondents studied this question in educational institutions and independently using the Internet. It is also clear that 43% of respondents have never heard of IIoT (Fig. 1).

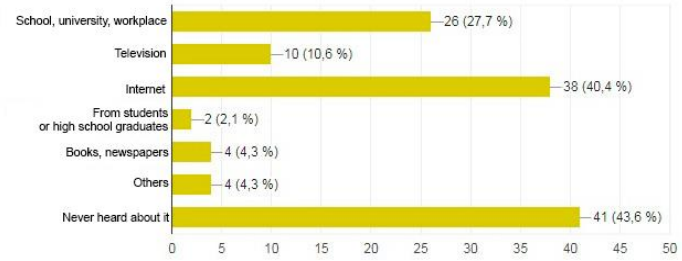


Fig. 2. Answers to the question about the knowledge of the term IIoT

Distribution of answers to the next question showed that the majority of respondents are interested in finding out how the Internet of Things can be used to improve the quality of everyday life of ordinary people (almost 60%), a quarter of the surveyed audience are interested in how IIoT can affect science or electronics (Fig. 2).

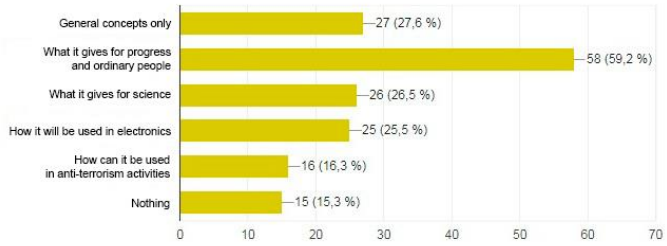


Fig. 3. Answers to the question "What would you like to know about the Internet of things?"

A significant part of respondents believe that the media reviews IIoT on the positive side, the rest of them find it difficult to answer this question (Fig. 3).

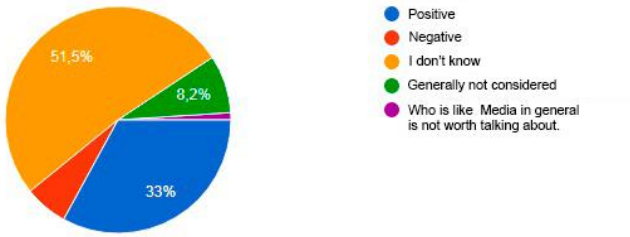


Fig. 4. Answers to the question "Opinion of respondents about the relation of mass-media to the IIoT from a negative or positive point of view"

It is necessary to consider the risks and safety of IIoT. According to the study by Kaspersky Lab [8], incidents with IoT devices are among the top three threats with the greatest financial damage to companies today. This applies to companies of all sizes: small and medium businesses, and large corporations. Realizing these risks, more and more companies

are taking steps to solve the problem of IoT security. More than 100 individuals who make decisions from companies in various industries were interviewed as part of the study "What prevents IoT managers from sleeping at night in 2018", conducted by IoT World. 72% of respondents guaranteed that their enterprises provide all the necessary measures for IoT security. At the same time, 43% of respondents do not test their IoT devices for vulnerabilities at all, which is an alarming signal, and less than half of the respondents have organized an inventory of connected devices. The reformative ability of IoT can be realized only when the safety of the entire ecosystem is priority [9]. Also, in the field of industrial cyber security, Kaspersky Lab experts identify one of their main problems - the lack of uniform standards for ensuring cyber security of industrial enterprises, including the security standards of industrial IoT devices [10].

At the same time, the tendency of Internet of things spread in everyday life of citizens carries not only material risks for society and individuals. Problems of preserving the confidentiality of personal data, reliable storage of large volumes of data using cloud technologies, which also ensure their integrity and availability, are becoming urgent. Today, IoT is one of the most vulnerable areas of information technology in security. Experts note that a safe IoT ecosystem does not exist today [11]. IoT in the context of the spread of targeted attacks is practically unprotected: attackers, interested in something, can effortlessly penetrate the private sector of life with the help of their IoT devices.

Thus, from the point of view of IIoT devices data security, each implementation in an enterprise infrastructure can provoke significant material losses; therefore, information security is crucial in this case. If we consider IoT devices, there is a high risk of interception of confidential data, which can lead to theft of money from bank accounts, as well as unauthorized access of intruders to personal information.

However, the survey showed that a little more than half of the respondents independently (without the suggested response options) found it difficult to determine the possible security risks that bear the active use of the Internet of things. At the same time, 74% of respondents fear leaks of their personal data (Fig. 4).

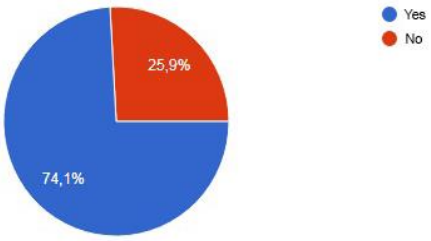


Fig. 5. Diagram showing the opinion of respondents about the risks of personal information loss in IIoT

The following diagram (Fig. 5) shows that the majority of respondents (81%) considers the risk of losing their personal information in the IIoT environment real.

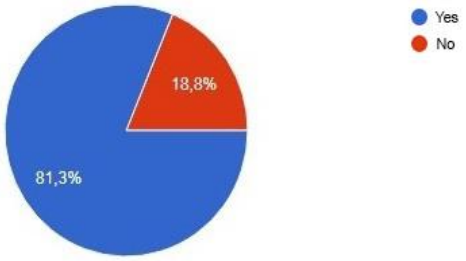


Fig. 6. Answers to the question about the risks of losing information in the Internet of Things environment

III. RESULTS OF THE STUDY

The survey showed that most of the respondents basically have an idea about IIoT (56.4%), a significant part of them learned about it from the Internet (40.4%) and in university (27.7%). The survey results revealed that the majority of respondents (59.2%) are interested in the question of how IIoT influences on progress and directly on society, 26.5% are interested in the potential of IIoT for science, 63.3% believe that IIoT can influence their daily life, but 36.7% believe the opposite.

A significant part of respondents find it difficult to answer what the safety of IIoT may be. However, 74.1% fear personal data leakage during the development of IIoT. The vast majority of respondents (81.3%) understand that there are risks of losing information in IIoT technologies.

Thus, it can be concluded that the awareness of the youth audience about the risks and safety of IIoT is insufficient, and the competencies are largely intuitive. Most likely, as the study showed, this topic is just beginning to be discussed at conferences and summer schools of students. Educational disciplines, one way or another, connected with IIoT have not yet been fully developed at universities. This topic is currently discussed fragmentary, there are practically no discussions of the problem in professional community with students. And announced sections at international conferences on this topic are often not filled with reports with students' participation. These arguments indicate the need to attract an attention of student audience to understanding of processes and problems of IIoT, as well as involvement of students in project activities on these issues.

Certain aspects investigated and the corresponding results obtained with financial support of the Russian Foundation of Basic Research for the project "Pedagogical basics of youth students' socialization in Internet space and their implementation in educational system", No. 17-36-00039, 2017-2019.

IV. REFERENCES

[1] Индустриальный (промышленный) Интернет вещей в мире и перспективы развития в Россиии. [Industrial Internet of Things in the world and prospects for development in Russia] // Analytics ICT and Digital Media. URL: http://json.tv/ict_telecom_analytics_view/mirovoy-opyt-vnedreniya-proektov-v-sfere-industrialnogo-promyshlennogo-interneta-veschey-i-perspektivy-ih-realizatsii-v-rossii--20160919061924.

[1] Индустриальный Интернет вещей в Росии и мире. Обзор состояния и перспективы развития.[Industrial Internet of Things (IIoT) in Russia and the world // Portal on modern technologies of mobile and wireless communications]. URL: http://1234g.ru/novosti/iiot-v-rossii-i-mire.

[2] Промышленный Интернет вещей в России. Исследование TAdviser и ГК «Ростех» [Industrial Internet of Things in Russia. Research TAdviser and Rostec GK]. URL: http://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:IIoT_2018:_%D0%A0%D1%8B%D0%BD%D0%BE%D0%BA_%D0%BF%D1%80%D0%BE%D0%BC%D1%8B%D1%88%D0%BB%D0%B5%D0%BD%D0%BD%D0%BE%D0%B3%D0%BE_%D0%B8%D0%BD%D1%82%D0%B5%D1%80%D0%BD%D0%B5%D1%82%D0%B0_%D0%B2%D0%B5%D1%89%D0%B5%D0%B9_%D0%B2_%D0%A0%D0%BE%D1%81%D1%81%D0%B8%D0%B8.

[3] Промышленный Интернет вещей в России [Industrial Internet of Things in Russia]. URL: http://www.tadviser.ru/index.php/.

[4] TAdviser: Российский рынок промышленного Интернета в 2017 году достиг 93 млрд рублей [TAdviser: the Russian market of industrial Internet of things in 2017 reached 93 billion rubles. URL: http://www.cnews.ru/news/line/2018-05-21_tadviser_rossijskij_rynok_promyshlennogo_interneta.

[5] Студенты представили прототипы IoT-систем в летней школе Samsung [Students presented prototypes of IoT systems at Samsung Summer School]. URL: https://mipt.ru/news/studenty_predstavili_prototipy_iot_sistem_v_letney_shkole_samsung.

[6] Chvanova M.S., Shlenov Yu.V., Molchanov A.A. et al. New forms of young students' socialization in the Internet space // Quality Management, Transport and Information Security, Information Technologies. 2017. P. 648–651.

[7] Иванов А. Разработаны рекомендации по обеспечению безопасности IoT-устройств. [Ivamov A.Developed safety guidelines for IoT devices]. URL: https://www.anti-malware.ru/news/2017-12-26-1447/25175.

[8] Вайтчерч Г. Главные проблемы для безопасности Интернета вещей [Whitechurch G. Top issues for the security of the Internet of Things]. URL: https://iot.ru/promyshlennost/glavnye-problemy-dlya-bezopasnosti-interneta-veshchey.

[9] Панков Н. Еще один шаг к защите промышленного IoT [Pankov N. Another step towards industrial IoT protection standards]. URL: https://www.kaspersky.ru/blog/enisa-recomendations/19376/.

[10] Информационная безопасность Интернета вещей (Internet of Things). [Information security of the Internet of Things]. URL: http://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:%D0%98%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D0%BE%D0%BD%D0%BD%D0%B0%D1%8F_%D0%B1%D0%B5%D0%B7%D0%BE%D0%BF%D0%B0%D1%81%D0%BD%D0%BE%D1%81%D1%82%D1%8C_%D0%B8%D0%BD%D1%82%D0%B5%D1%80%D0%BD%D0%B5%D1%82%D0%B0_%D0%B2%D0%B5%D1%89%D0%B5%D0%B9_(Internet_of_Things).

The Research of Blockchain Technology for Data Protection in IoT Devices

Mikhail G. Gorodnichev[1], Stanislav S. Makhrov [2], Elena N. Denisova[3]
[1,3]IT Department, [2] Department of Scientific Research
Moscow Technical University of Communications and Informatics
Moscow, Russia
[1]gorodnichev89@yandex.ru, [2]slavam4@yandex.ru,
[3]dummy.newmailbox@gmail.com

Ilya D. Buldin
Department of Economic Sciences
Higher School of Economics
Moscow, Russia
idbuldin@gmail.com

***Abstract*. The article presents a study aimed at determining an effective security model based on blockchain technology for IoT devices. It is completed a modification of the key features of the classical blockchain used for cryptocurrencies for adaptation to work in the network from the devices of the Internet of things, which requires minimal reaction time to the execution of control commands and knew the redemption of energy resources.**

***Key words:* blockchain, very careful registry, Internet of Things, data protection, information security.**

I. INTRODUCTION

The Internet of Things (IoT) represents one of the most significant disruptive technologies of this century. It is a natural evolution of the Internet (of computers) to embedded and cyberphysical systems, “things” that, while not obviously computers themselves, nevertheless have computers inside them. With a network of cheap sensors and interconnected things, information collection on our world and environment can be achieved at a much higher granularity [1].

IoT consists of devices that generate, process, and exchange vast amounts of security and safety-critical data as well as privacy-sensitive information, and hence are appealing targets of various cyber-attacks. Many new networkable devices, which constitute the IoT, are low energy and lightweight. These devices must devote most of their available energy and computation to executing core application functionality, making the task of affordably supporting security and privacy quite challenging. Traditional security methods tend to be expensive for IoT in terms of energy consumption and processing overhead. Moreover, many of the state-of-the-art security frameworks are highly centralized and are thus not necessarily well-suited for IoT due to the difficulty of scale, many-to-one nature of the traffic, and single point of failure [2].

Currently, IoT devices are a significant part of the total number of existing digital devices. According to Marketsandmarkets, by 2018, the global market of the Internet of things is about $200–250 billion. Expected that by 2022 it will reach $561.04 billion. Thus, the volume of the market of the Internet of things is constantly growing and an increasing number of digital devices belongs to the segment [3].

The privacy risks of IoT are exacerbated by the lack of fundamental security safeguards in many of the first generation IoT products on the market. Numerous security vulnerabilities have been identified in connected devices ranging from smart locks to vehicles [1]. IoT devices require protection of data from such basic types of vulnerabilities such as: data corruption, substitution of devices, hacking devices.

IoT is experiencing exponential growth in research and industry, but it still suffers from privacy and security vulnerabilities. Conventional security and privacy approaches tend to be inapplicable for IoT, mainly due to its decentralized topology and the resource-constraints of the majority of its devices [1].

One of the most promising modern methods of data protection today is the blockchain technology, which allows to control all transactions to perform authorization through the public key infrastructure.

Blockchain, a distributed append-only public ledger technology, was initially intended for the cryptocurrencies, e.g., Bitcoin. In 2008, Satoshi Nakamato introduced the concept of blockchain that has attracted much attention over the past years as an emerging peer-to-peer (P2P) technology for distributed computing and decentralized data sharing. Due to the adoption of cryptography technology and without a centralized control actor or a centralized data storage, the blockchain can avoid the attacks that want to take control over the system.

Combined with the blockchain technology, IoT systems benefit from the lower operational cost, decentralized resource management, robustness against threats and attacks, and so on. Therefore, the convergence of IoT and blockchain technology aims to overcome the significant challenges of realizing the IoT platform in the near future [4].

The study of the technology of the blockchain to determine an effective model of protection devices for the IoT is relevant and priority area of research.

II. PROBLEM DEFINITION

Existing methods for securing IoT devices are not always effective for protecting against the following types of threats:

- data corruption;
- device spoofing;
- breaking.

Protection against these types of threats can be provided by means of blockchain technology, which provides:

- audit logging of transactions;
- data encryption;
- public key authentication;
- execution and control over the execution of smart contracts.

Thus, the distortion of data substitution devices and hacking would be impossible in view of the specifics of the blockchain technology mentioned above. At the same time, the classic blockchain [5] is adapted for financial transactions and remuneration of miners for solving cryptographic problems. At the same time, the solution of problems is very resource-intensive to both computing resources and temporary resources. In turn, for IoT devices, these resources are critical because devices are typically Autonomous and must respond instantly to commands with minimal waiting [1].

The purpose of this study is to determine an effective protection model based on blockchain technology for IoT devices.

III. THEORY

A. Algorithm of consensus

The technology of the blockchain is concluded the algorithm of consensus. A consensus algorithm is required to select a node whose recent transaction data (represented as the last block of transactions) will accept all other nodes and add them to its block chain. This algorithm allows you to prevent the possibility of conflicting transactions in the network by multiple nodes by selecting only one node, the transaction block of which will be considered relevant and correct. The classic blockchain [5] solves the problem of double spending of funds from one wallet.

After the consensus algorithm is executed, all nodes in the network will update their block chain to match what they receive from the node that was selected by the consensus algorithm. At the same time, when updating the block chain, the receiving nodes check the correctness of the consensus algorithm. If the consensus algorithm can take a long time to execute, the check is instantaneous.

The classical blockchain [5; 6] uses the proof-of-work algorithm as a consensus algorithm (Proof of Work, PoW). The essence of the algorithm is that the miner node should be the first in relation to all other nodes to solve some mathematical problem. As a mathematical problem can be calculated from an arbitrary number of such a hash, which at the end will be 5 zeros. The solution of such a problem takes a lot of time and computing resources.

Other variants of the consensus algorithm are also known, in particular [6]:

- Proof of Activity (PoA);
- Proof of Stake (PoS);
- Proof of Capacity (PoC);
- Proof of Importance;
- Proof of Authority (PoAuthority).

Each of these, as well as other consensus algorithms, are mainly focused on operations with monetary transactions [6] and are not suitable for adaptation to work in the network from IoT devices.

To enable the use of blockchain technology in IoT devices, it is proposed to use such a proof-of-work algorithm to determine the node with the correct and relevant data instantly. In view of the absence of the need to reward miners for their work, and accordingly, the need to perform this work by the miners themselves, the consensus algorithm can be determined by the simplest generation of a random number:

$$proof = random\left(list, H_{PREV\_BLOCK}\right) \quad (1)$$

According to the formula (1), the value of the proof consensus algorithm is calculated as a random number from the $list = \{miner_1, \ldots, miner_z\}$ miners list and the calculated hash value from the previous H_{PREV_BLOCK}. A random number will be in the list of miner node numbers range of the *list.* Therefore, without spending a lot of time and computing resources, the node whose transaction block will be accepted as correct and up-to-date will be determined. Introduction in the formula (1) of the consensus algorithm hash of the previous block H_{PREV_BLOCK} will allow the nodes to upgrade chain to see if it made a substitution blocks node miner.

B. Blockchain block contents

The result of the consensus algorithm cannot be faked because the hash value of the previous block is involved in calculating the hash of each new block:

$$H_{BLOCK} = SHA256\begin{pmatrix} index + H_{PREV\_BLOCK} + \\ H_{MPROG} + timestamp \\ + proof + \sum_{k=1}^{p} transaction_k \end{pmatrix} \quad (2)$$

According to the formula (2) hash each transaction block is calculated as the hash-function *SHA256* of the following arguments: index block *index*, *hash* H_{PREV_BLOCK} from the previous block, hash H_{MPROG} code from the firmware executable on the node, timestamp of creation timestamp of the block, the result of the consensus algorithm proof and the amount of the transaction.

Introduction in the formula (2) H_{MPROG} hash code of the firmware is another difference from the classical blockchain [5] and allows to protect the network nodes of the Internet of things from breaking by controlling the integrity of the firmware. If the

firmware hashes are different, this will signal that the node is not trusted and communication with it should be terminated.

C. Data transmission and addressing of nodes

Data transfer and addressing of IoT nodes can be done through the public key infrastructure. Two keys are created – *PrivateKey* and *PublicKey*, which are used for different purposes..

The *PrivateKey* is used for the following tasks:

- creating a *PublicKey*;
- the digital signature of transactions;
- decryption of received data.

The *PublicKey* is used for the following tasks:

- addressing nodes by the *PublicKey*;
- for generating the host address: a *PublicKey* hash is generated by hash function SHA256. The resulting hash value is then hashed again using the RIPEMD160 function. The final hash of RIPEMD160 is Base64 encoded;
- validation of the digital signature of the received transaction;
- encryption of transmitted data.

D. Registration and transfer of data between network nodes

Let we have a node $node_T$, that is required to register in a network consisting of n standard $node_{1,\dots,}node_n$ and z miner nodes $miner_{1,\dots,}miner_z$. While $n>z$. Node $node_T$ generated 2 key $PrivateKey[node_T]$ and $PublicKey[node_T]$. The last key (public) is sent to all nodes of the $node_{1,\dots,}node_n$ and $miner_{1,\dots,}miner_z$, which by means of this key will be able to verify the authenticity of the digital signature of transactions received from the $node_T$, and will also be able to transmit data to the $node_T$. node in encrypted form. Digital signature of transaction by far identified the sender. In response to the public key received from $node_T$ the nodes send them their public keys $PublicKey[node_1],\dots,PublicKey[node_n]$, $PublicKey[miner_1],\dots,PublicKey[miner_z]$ to perform similar functions.

If you want to control the issuance of digital signature certificates, this function can be assigned to the miner nodes $miner_{1,\dots,}miner_z$.

IV. Experimental Prototype And Results

The model described in the theoretical part is implemented in practice in the Python programming language on the hardware basis of the IoT module Onion Omega2+ [7] (Fig. 1).

Fig. 1. IoT module Onion Omega2+

The Onion Omega2+ module has the following features:

- processor: MIPS32, 580 MHz;
- memory: 128 MB;
- flash memory: 32 MB;
- USB: one 2.0 port;
- SD-slot for 1 MicroSD card;
- support WiFi: b/g/n;
- number of GPIO: 15;
- support PWM, UART, I2C, SPI, I2S.

As part of the experiment, the code of the blockchain model was recorded on 7 devices, each device was a miner at the same time. The following types of attacks on the network from IoT devices were used:

- node substitution;
- listen to traffic;
- change the firmware code to create a backdoor in it;
- units change the chain on the nodes.

As a result of implementation of all listed attacks, the network revealed each attempt of influence and successfully resolved all conflict situations.

V. Conclusion

The developed model of data protection and its practical implementation using IoT-modules Onion Omega2+, showed that through the blockchain technology, Internet of things devices can be protected from the following vulnerabilities:

- Control over the distortion of information, hacking devices is provided by storing transactions on all nodes (transaction audit) and checking them by Consensus algorithm, for distortion of the previous information and the introduction of malicious code in the original firmware code.
- Protection against spoofing and introduction of phantom devices into the network is provided by applying a digital signature for each transaction received from the node.
- Data privacy is guaranteed by encrypting traffic according to the public key infrastructure. The information is protected and available only to the parties involved in a particular transaction.

The model provides the ability to authenticate users through a decentralized or centralized public key infrastructure. Also, through this model can be implemented proof of warranty, if the blockchain network nodes will be entered data on the manufacturer of electronic components.

The disadvantages of the model include the inability to update the software, because when you add a new block to the blockchain, the hash from the H_{MPROG} firmware is compared with its value in the previous blocks of the blockchain.

VI. REFERENCES

[1] Dorri A., Kanhere S.S., Jurdak R. Blockchain in internet of things: Challenges and Solutions. URL: https://arxiv.org/abs/1608.05187.

[2] Dorri A., Kanhere S. S., Jurdak R. et al. Blockchain for IoT security and privacy: The case study of a smart home // 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). Kona, 2017. P. 618-623. doi: 10.1109/PERCOMW.2017.7917634.

[3] Marketsandmarkets.com: Internet of Things (IoT) Market by Software Solution (Real-Time Streaming Analytics, Security Solution, Data Management, Remote Monitoring, and Network Bandwidth Management), Service, Platform, Application Area, and Region - Global Forecast to 2022. URL: https://www.marketsandmarkets.com/Market-Reports/internet-of-things-market-573.html.

[4] Ferrag M.A., Derdour M., Mukherjee M. et al. Blockchain Technologies for the Internet of Things: Research Issues and Challenges. URL: https://arxiv.org/abs/1806.09099.

[5] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system. URL: https://bitcoin.org/bitcoin.pdf.

[6] Nguyen T., Kim K. A survey about consensus algorithms used in Blockchain // Journal of Information Processing Systems. 2018. Vol. 14. P. 101–128. 10.3745/JIPS.01.0024.

[7] Omega 2 - The Invention Platform for the Internet of Things. URL: https://onion.io/omega2/.

Classification of encrypted Applications of Traffic Mobile Devices using the Data Mining

Oleg I. Sheluhin[1], Viacheslav V. Barkov[2], Mikhail V. Polkovnikov[3]
Department of Information Security
Moscow Technical University of Communications and Informatics
Moscow, Russia
[1]sheluhin@mail.ru, [2]viacheslav.barkov@gmail.com, [3]mnxamoto@mail.ru

Abstract. The problem of determining by a cellular operator what applications a particular network user has used is needed to compile statistics of the most frequently used applications. Such a definition of application statistics helps to not only monitor network status, detect failures, but also, if necessary, restrict access to network resources that, from the point of view of information security, can harm the user. The introduction of methods of data mining and machine learning allows to perform automatic classification, analysis and filtering of malicious and unwanted mobile network traffic applications. Malicious mobile applications can be a threat to the integrity or availability of data, and unwanted ones are a threat to confidentiality. The paper considers classification of network encrypted traffic by application types: email of Mail.ru, Sberbank, Skype, Pikabu, Instagram, Hearthstone and other methods of machine learning using algorithms Naive Bayes, C4.5, SVM, AdaBoost and Random Forest. For the analysis, more than two million network packets were collected from four applications that transmitted encrypted traffic, after which training and test samples were generated. To assess the quality of the qualifier, such criteria as Accuracy, Precision, Recall, F-Measure and Area Under Curve were used.

The use of the InfoGain algorithm showed that to ensure the high quality of classification of traffic of applications that use encryption, it is enough to limit thirteen attributes. Classifier Random Forest is the slowest, but has the best indicators of assessing the quality of classification. The size of the learning sample of the Random Fores algorithm to achieve a sufficiently high quality of classification of mobile applications cannot exceed 300 threads. To ensure high quality thread classification, it is enough to analyze from 16 to 58 packets in a stream depending on the application. Further increase in the number of packets in the stream does not lead to a noticeable improvement in the quality of classification.

Keywords: **classification, machine learning, algorithms, network traffic, application, packet, flow, protocol, network, mobile applications efficiency.**

I. INTRODUCTION

The problem of determining by the cellular operator which applications the network user used is needed to compile statistics of the most frequently used applications. Such application statistics help not only to monitor the network status, detect failures, but also, if necessary, to restrict access to network resources that, from the point of view of information security, can harm the user.

The introduction of machine learning methods allows automatic classification, analysis and filtering of malicious and unwanted mobile applications of network traffic [1–3].

Malicious mobile applications can be a threat to the integrity or availability of data, and unwanted ones are a threat to confidentiality. Classification of traffic of mobile applications that use encryption-using encryption does not imply its decryption. The data inside the packets remains confidential and is only accessible to the user and the remote node.

Mobile applications that use traffic encryption can be divided into three groups. The first group includes applications that use the SSL / TLS transport layer encryption protocol [4] in conjunction with the HTTPS application layer protocol. Examples of such applications are Google, Facebook, Sberbank, etc. The second group includes applications that use the P2P protocol [6] with encryption (BitTorrent, MuTorrent, Vuze, etc.). The third group includes applications that use, in addition to transport-level encryption protocols, their own encryption protocols. Examples of such applications are Skype, WhatsApp, Telegram, etc.

In a situation of complex definition of the type of traffic encryption, it is advisable to use machine learning methods to classify the traffic of mobile applications using encryption.

II. TECHNOLOGY FOR COLLECTING TRAFFIC OF MOBILE APPLICATIONS

To create a database of mobile application traffic, the "Traffic Analysis System" software package was developed, which includes a database server, an application server, a Web application and client software for mobile devices running the Android operating system (mobile client).

The process of traffic collection using the "Traffic Analysis System" software package, as well as the interaction of the components of the software complex with each other and with external mobile applications is shown in Figure 1.

A mobile client of the "Traffic Analysis System" software package is installed on a smartphone or tablet running the Android operating system. This client intercepts network traffic packets of the specified applications that are also installed on this device.

Intercepted packets of network traffic are sent to the application server of the “Traffic Analysis System” software package installed on a server computer controlled by the Windows Server operating system 2016.

The application server of the “Traffic Analysis System” software package group’s network traffic packets into flows and, using the database server, saves data to the database.

Data exchange between the components of the “Traffic Analysis System” software is carried out via the global Internet using the HTTP protocol in JSON format. The application server includes a Web service that provides REST API clients, with which you can access the collection functions of network traffic packets, manage datasets, create and train classifiers, classify and other functions.

With the use of the software complex, traffic of mobile applications of three categories was collected: “With traffic encryption”, “Without traffic encryption”, “With partial encryption of traffic”.

During the collection of traffic of mobile applications using encryption, the network traffic flows of 6 applications were collected: Instagram, Mail Mail.ru, Pikabu, Sberbank-Online, Hearthstone, Skype. Table 1 shows the numerical characteristics of the collected network packets and flows for training and test samples.

To conduct the experiment and generate the initial data on the mobile device, specialized software “Traffic Analyzer” was installed under the management of the Android operating system [7] version 4.4. Figure 2 shows the process of connecting a mobile client to the application server of the “Traffic Analysis System” software package.

Fig. 1. Scheme of collecting mobile traffic

Table 1. Characteristics of the application dataset by application type when analyzing network packets and flows

Application	Training sample		Test sample	
	packets	flows	packets	flows
email Mail.ru	162517	3356	79612	1644
Sberbank	156648	3303	80482	1697
Skype	146315	3329	73443	1671
Pikabu	167182	3325	84220	1675
Instagram	1286600	3357	629695	1643
Hearthstone	151298	3330	75876	1670
Total	2070562	20000	1023328	10000

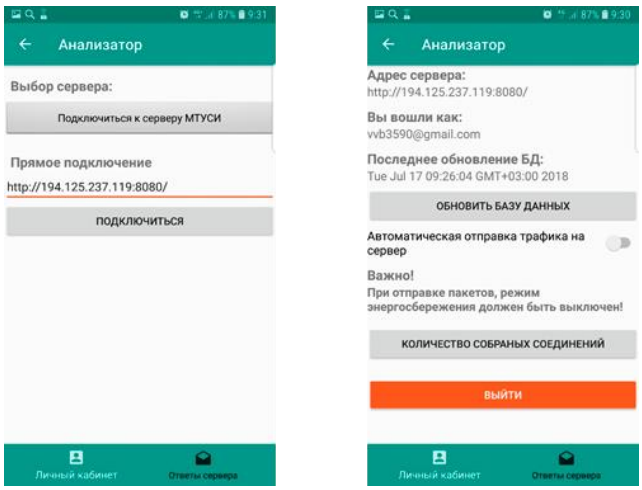


Fig. 2. Graphical user interface of the mobile client of the “Traffic Analysis System” software package: connection to the server and information about the server

Figure 3 shows the graphical user interface of the mobile client of the “Traffic Analysis System” software package during the configuration process to intercept the traffic of the specified applications and sends it to the server in the process of interception of traffic, in the process of viewing the number of collected streams.

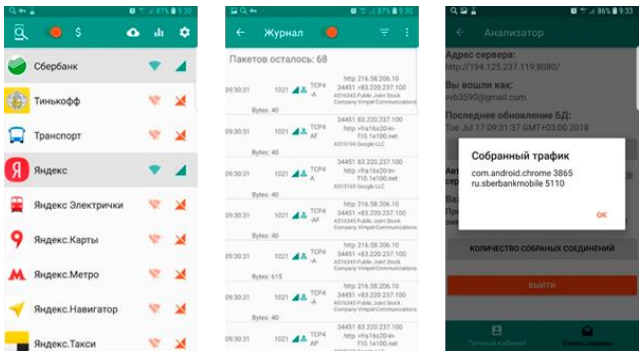


Fig. 3. Graphical user interface of the mobile client of the “Traffic Analysis System” software package

III. RESULTS OF THE CLASSIFICATION OF MOBILE APPLICATION TRAFFIC USING ENCRYPTION

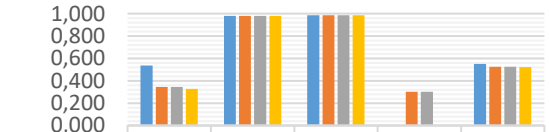
Using the algorithm for selecting the attributes InfoGain [http://www.cs.waikato.ac.nz/ml/weka] of the 23 original attributes, 13 were allocated:

- average size of the data portion from the server side (AverageSizeDataOnTransportLayerFromServer);
- average packet size on the server side (AverageSizeOnTransportLayerFromServer);
- the server's efficiency is the amount of applied application load transferred divided by the total number of

transferred application and transport layer loads (EfficiencyOfServer);

- customer address (FirstIP);
- size of network layer payload on the server side (NetworkLayerPayloadSizeFromServer);
- payload ratio is how many times the client transmitted more bytes of information than the server (RatioOfData);
- the server address (SecondIP);
- standard deviation of the size of the data portion from the client side (StandardDeviationOfDataOnTransportLayerFromClient);
- standard deviation of the data portion size from the server side (StandardDeviationOfDataOnTransportLayerFromServer);
- standard deviation of the packet size from the client side (StandardDeviationOfPacketSizeFromClient);
- standard deviation of the packet size from the server side (StandardDeviationOfPacketSizeFromServer);
- size of the payload of the transport layer from the client side (TransportLayerPayloadSizeFromClient);
- transport server payload size from the server side (TransportLayerPayloadSizeFromServer);

To assess the effectiveness of classification algorithms, the following information search metrics [1] were used: Precision, Recall, F-Measure, ROC curves (Receiver Operating Characteristic Curve), AUC (Area Under Curve) is the area under the ROC curve. Because of the processing of the experimental data, quantitative results were obtained, represented in the form of averaged histograms in Figure 4.



	Naive Bayes	C4.5	Random Forest	AdaBoost	SVM
Precision	0,535	0,980	0,985	0,000	0,552
Recall	0,344	0,980	0,985	0,302	0,524
Accuracy	0,344	0,980	0,985	0,303	0,524
F-мера	0,327	0,980	0,985	0,000	0,521

Fig. 4. Scheme of collecting mobile traffic

An analysis of the results shows that the algorithms C4.5 and Random Forest show the best results of classification. Figure 5 shows the time intervals in milliseconds that the classifiers required for training and testing required. As you can see, the fastest ones at the training stage were Naive Bayes, C4.5 and AdaBoost, and at the testing stage: C4.5, Random Forest, AdaBoost and SVM. The fastest classifiers in both phases are C4.5 and AdaBoost. However, although the AdaBoost classifier is the "fastest", however, it has the worst results of assessing the quality of classification. Classifier Random Forest is the most "slow", but has the best classification rating.

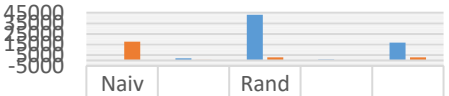
Table 2 shows the AUC values for the Random Forest algorithm, which show the high reliability of the classification of the applications considered.

The conducted researches allow to draw a conclusion that to ensure high quality of the flow classification it is enough to analyze from 16 to 58 packets in the stream depending on the application. The further increase in the number of packets in the stream does not lead to a marked improvement in the quality of the classification.

IV. CONCLUSIONS

Based on the use of the InfoGain algorithm, it is shown that in order to provide a high-quality classification of the considered applications using encryption for data transmission, it is sufficient to limit thirteen attributes. Classifier Random Forest is the slowest, but has the best indicators of assessing the quality of classification.

The size of the training sample algorithm for Random Forest sufficiently high classification quality (accuracy 90%) may not exceed 300 threads. To ensure high-quality classification of streams, it is enough to analyze from 16 to 58 packets of a flow depending on the application. Further increase in the number of packets in the flow more than not leads to a noticeable improvement in the quality of the classification.



	Naive Bayes	C4.5	Random Forest	AdaBoost	SVM
Время обучения	218	1953	42923	968	16813
Время тестирования	17673	563	2782	329	2969

Fig. 5. Time ranges for training and testing classifiers

TABLE I. AUC VALUES FOR RANDOM FOREST ALGORITHM

Class	Instagram	email Mail.ru	Skype	Sberbank	Hearthstone	Pikabu
ROC-AUC	0,9904	0,9885	0,9809	0,9920	0,9888	0,9797

V. REFERENCES

[1] Шелухин О.И., Ерохин С.Д., Ванюшина А.В. Классификация IP-трафика методами машинного обучения. М.: Горячая линия – телеком, 2018. [Sheluhin O.I, Erokhin S.D., Vanyushina A.V. IP traffic classification by machine learning methods. Moscow: Hotline — Telecom, 2018].

[2] Костин Д.В., Шелухин О.И. Сравнительный анализ алгоритмов машинного обучения для проведения классификации сетевого зашифрованного трафика // T-Comm: Телекоммуникации и транспорт. 2016. № 9. С. 46–52. [Kostin D.V., Sheluhin O.I. Comparison of nachine kearning algorithms for encrypted ckassification// T-Comm. 2016. Vol. 10, N 9. P. 43–52].

[3] Шелухин О.И., Смычек М.А., Симонян А.Г. Фильтрация нежелательных приложений трафика подвижной радиосвязи для обнаружения угроз информационной безопасности //

Радиотехнические и телекоммуникационные системы. 2018. № 1. С. 87–98. [Sheluhin O.I., Smychek M.A., Simonyan A.G. Filtering unwanbted mobile ratio traffic applications to detect imnfornatiobn security threats // Radio and Telecommunication Systems. 2018. N 1. P. 87–98].

[4] Шелухин О.И., Ванюшина А.В., Габисова М.Е. Фильтрация нежелательных приложений интернет-трафика с использованием алгоритма классификации Random Forest // Вопросы кибербезопасности. 2018. №2. С. 44–51. [Sheluhin O. ,Vanyushina A., Gabisova M. The filtering of unwanted applications in internet traffic using random forest classification algorithm // Cybersecurity issues. 2018. N 2. P. 44–51]..

[5] Rescorla E. SSL and TLS: Designing and Building Secure Systems. Reading: Addison-Wesley Professional, 2000. T. 1.

[6] Callegati F., Cerroni W., Ramilli M. Man-in-the-Middle Attack to the HTTPS Protocol // IEEE Security Privacy. 2009. T. 7, вып. 1. P. 78–81. DOI:10.1109/MSP.2009.12.

[7] Pouwelse J.A., Garbacki P., Epema D.H.J. et al. The BitTorrent P2P File-sharing system: Measurements and analysis// IPTPS 2005. LNCS. Vol. 3640. Heidelberg: Springer, 2005.

[8] Коматинэни С., Маклин Д., Хэшими С. GoogleAndroid: программирование для мобильных устройств ProAndroid 2. СПб.: Питер, 2011. [Komatinani S., McLean D., Hashimi S. Google Android: programming for mobile devices Pro Android 2. St Petersburg: Piter, 2011].

The Blockchain Technology in State Institutions

Yakov M. Dalinger[1], Evgeny A. Saksonov[2]
St. Petersburg State University of Civil Aviation
St. Petersburg, Russia
Moscow Technical University of Communications and Informatics
Moscow, Russia
[1]iakovdalinger@gmail.com, [2]saksmiem@mail.ru

Abstract. **The problems of building a distributed registry (blockchain) for state institutions are considered. The main tasks, the solution of which is necessary when creating the registry, are highlighted. The model of formation of registry blocks is given.**

Results can be useful to designers and administrators of the state institutions distributed registry .

Key words: **blockchain, distributed registry, mathematical model, state institution.**

I. INTRODUCTION

The blockchain technology is becoming increasingly popular in creating information systems of state institutions at the federal, regional and municipal levels due to its features [1]:

- the impossibility of making changes to the entries (copies) created in the registry;
- fixation of all changes in previously made records in the form of new records;
- high functional reliability, associated with the presence of copies;
- a full match of all available copies of the registry;
- the possibility of unlimited expansion of the amount of the registry.

These features allow the use of blockchain technology in building archives and specialized information systems of various levels and purpose (for example, systems of personal data processing, systems of financial organizations, ministries and departments) where it is required to ensure high reliability of information storage, control of access to data and be able to increase the amounts of stored information.

Below we will use the synonyms: blockchain, the distributed registry or the registry [5].

II. DESCRIPTION OF THE PROBLEM

The processes of applying of the systems based on blockchain technology are significantly different from the known methods of creating and operating cryptocurrency related systems, and those differents must be considered when developing.

The main differences are as follows:

- the lack of the need to compete for the right to enter data into the register and receive remuneration;
- territorial localization of copies of register in the locations of representative offices of state institutions (for example, within the Russian Federation, region, municipal district, etc.);
- compliance of standards governing the processing and access to information in each case;
- ensuring the protection of information in the registry from specific threats related to the functioning of the registry, the properties of stored information and the requirements of state institutions;
- ensuring access to the registry of various groups (types) of users (employees of state institutions, representatives of external organizations, citizens of the Russian Federation) in compliance with the established access rights;
- the availability of special means of checking data recorded in the register, depending on the purpose of the data, the specific tasks of the state institution;
- compliance with the standards for the response time to requests, recovery from failures.

In addition, copies of the registry contain a large amount of various information, structured as blocks often presented in an encrypted form, which can complicate the search for data in the registry and requires the creation of means of its presentation in a convenient form for users of the registry.

III. REGISTRY CREATION TASKS

When creating a systems on base blockchain technology for a particular state institution, are require solution the following tasks:

- Development of methods for creating and provided operating a distributed registry. This includes determining the structure of the registry, the composition of the server groups for the registry (it is possible to store each copy on several servers), the procedure for forming data blocks, and writing blocks to the registry.
- Development of the structure of records and blocks of the distributed registry. This will allow you to more easily obtain a representation of registry entries in the required form for processing.
- The selection of different types of registry users and the assignment for users each type of access rights. Users can be employees of state institutions and have rights write and read or it can to be particular persons with only read rights.
- Development of methods for ensuring the coherence of data in accessible copies of the registry. After the blocks

are entered in the register, obtaining coherent copies is possible using various known methods, depending on the requirements of the state institution.

- Creation of mathematical models for evaluating the characteristics of a distributed registry depending on its organization, parameters and purpose (reliability, length of the procedure for forming and inserting blocks and ensuring the coherence of data in copies, response time to user requests, recovery time after of failurs, protection from threats attacks of various kinds).
- Simulation and development of recommendations for the registry operation depending on its parameters (the number of copies and their location, the number of users, the number of users of various types, etc.).
- Creation of a prototype for carrying out field tests, matching models with actual conditions.
- Creation of basic software (development of requirements and composition of a software, development of a prototype of basic software modules, etc.).
- Development of guides for the creation and operation of a distributed registry, taking into account the purpose and conditions of operation.

There are possible solutions to the above registry creation tasks. A possible solution for creating a registry and organizing its work is shown in Figure 1.

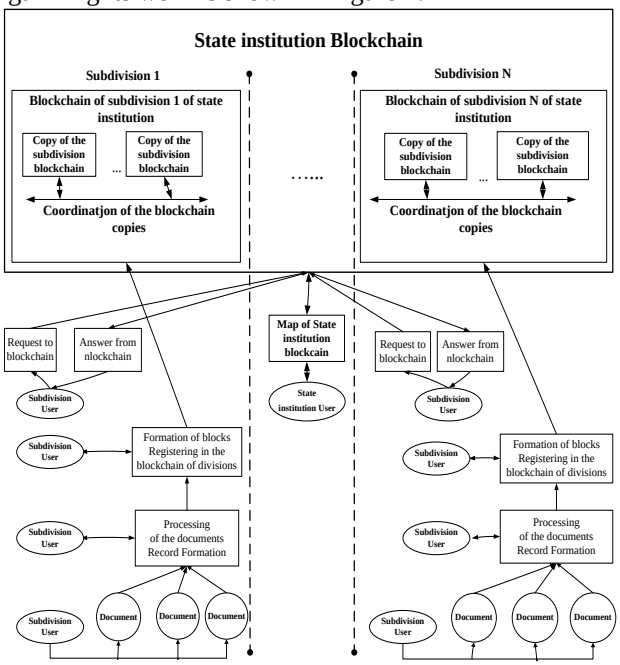


Fig. 1. Schema of the system based on the blockchain technology

Here, for writing of blocks, it is not necessary to solve the problem of finding a number with a given hash value. However, the hash of the block is also calculated and the hash of the created chain is generated. Since all users of the same subdivision (server) form a common queue, there will be no conflicts between them. Due to the specifics of the work of most state institutions, groups of users of each subdivision work with information of this subdivision, so other subdivisions create their own registers.

Here it is proposed for each subdivision of a public institution to create its own distributed registry. This solution allows to split data from different subdivisions, which often meets security requirements, reduce the amount of data in each registry, reduce the number of users with the right to enter data into the registry. For the access the registry uses known methods of control.

Consolidation of private registers into a general register of state institution is possible using a data allocation card. It is also possible to create a structure of a general registry in accordance with the structure of a state institution, for example, a network or hierarchical structure.

The number of copies of the registers of each subdivision is determined in accordance with the requirements for functional safety.

The structure of records and blocks in the registry should contain the creation date, the author and subdivision identifier.

The users of the registry of state institution and registries of divisions can be employees of state institution or private individuals (not employees). Employees are divided into those who are given the right to enter data in the registry and those who can only read the registry data. The private individuals are access rights only for reading data with restrictions on the amount and composition of data.

IV. MATHEMATICAL MODELING OF THE REGISTRY

When blocks are formed for entry in the registry, each block contains several documents. Thus, the intensity of the flow of blocks is less than the intensity of the flow of documents. This is the effect of information absorption in the formation of blocks.

A mathematical models of the block formation process are given in [2, 4, 6].

Below we consider the individual elements of the model with absorption [2].

System description

In general, we consider that the node for created of blocks receives N types document streams ($\infty > N \geq 1$), from which M types of blocks are created. The absorption algorithm is given by a matrix $\mathbf{M} = \|m_{ij}\|$, where $m_{ij} \geq 0$ the number of documents of the flow j, which are part of a block of the type i ($i = 1,2,...,M; j = 1,2,...,N$), the i-th row of the matrix is represented as a vector $\mathbf{m}_i = (m_{i1}, m_{i2}, ..., m_{iN})$.

Processing of a document consists in its inclusion in a block. Processing blocks are maintenance in the processing node (server of register) as a single (unified) message. Two types of queues can be formed in a node: a queue of documents and a queue of blocks waiting for service.

We investigate the case when the number of places for waiting for documents is limited. The number of places to wait for documents of each thread that are part of a block of type i is set by the vector $\mathbf{h}_i = (h_{i1}, h_{i2}, ..., h_{iN})$, where $\infty > h_{in} \geq m_{in}$ is the number of places to wait for documents of the flow number n.

As a model of the node, investigate the queueing system of a single service device, the input of which receives

N independent Poisson streams of documents, is studied. The rate of the flow of documents *i* – λ_i ($\infty > \lambda_i \geq 0$, *i* = 1, 2,…, *N*). The service device produces the formation and maintenance of blocks according to predefined rules.

Duration of processing of block of the type *i* (*i* = 1, 2, ..., *M*) is a random variable β_i with a distribution function $B_i(t)$, with first and second moments: $0 < b_{1i} = \int_0^\infty t dB_i(t) < \infty$ and $0 < b_{2i} = \int_0^\infty t^2 dB_i(t) < \infty$.

Here, to simplify the formulas, it is assumed that the duration of processing a block does not depend on its composition.

To analyze the operation of such a system, it is necessary to determine the characteristics of flows of blocks and probabilities of the documents loss.

Analysis of the system

The state of the system at time *t* is given by a vector $\mathbf{g}_i(t) = (g_{i1}(t), g_{i2}(t), ..., g_{iN}(t))$, where $g_{in}(t)$ is the number of documents of the flow number *n* ($m_{in} \neq 0$), which are in the system, and not included in the block type *i* ($g_{in}(t) = 0$, if $m_{in} = 0$).

If we consider the system at the time of arrive of documents that make up a block of type *i*, the set of states forms a nested finite Markov chain with the number of states - K_i.

Denote $\overline{\lambda_i} = \sum_{\substack{n=1 \\ (m_{in} \neq 0)}}^{N} \lambda_n$, *i* = 1 , 2,..., *M*. The probability that an arrived document will be a message of flow *n*: $q_{in} = \lambda_n / \overline{\lambda_i}$, *i* = 1 , 2,..., *M*; *n* = 1, 2,..., *N*.

For this Markov chain, a matrix of transient probabilities is constructed: $\mathbf{P}_i = \left\| p_{\mathbf{g}_i, g_i^*} \right\|$, where $\mathbf{g}_i$ and $\mathbf{g}_i^*$ are vectors of the states of chain; $p_{\mathbf{g}_i, g_i^*}$ is the probability of transition from state $\mathbf{g}_i$ to state $\mathbf{g}_i^*$.

The following formulas are obtained to calculate the values of transition probabilities:

$p_{\mathbf{g}_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^*, g_{i2}^*, ..., g_{iN}^*)) = q_{in}$, *i* = 1, 2, ..., *M*;

$p_{\mathbf{g}_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^*, g_{i2}^*, ..., g_{iN}^*)) = q_{in}$, *i* = 1, 2, ..., *M*;

$p_{\mathbf{g}_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^*, g_{i2}^*, ..., g_{iN}^*)) = q_{in}$, *i* = 1, 2,..., *M*;

$p_{\mathbf{g}_i, g_i^*} = \Pr((g_{i1}, g_{i2}, ..., g_{iN}) \to (g_{i1}^*, g_{i2}^*, ..., g_{iN}^*)) = \sum_{r=1}^{N} q_{ir}$;

This finite Markov chain is ergodic, and there is a limiting matrix: $\mathbf{A}_i = \lim_{r \to \infty} (\mathbf{P}_i)^r = \left\| a_{imn} \right\|, (m, n = 1, 2, ..., K_i)$, where $a_{imn} = a_{in}$ for all *n* = 1, 2, ..., K_i [3]. Here a_{in} is the limiting probability of hitting of the system at the next step in the state number *n*.

Among the set of States of the system, we select a subset of those for which a block is formed upon arrive of some document. We denote this subset - $\mathbf{U}_i$. The vector of the state of this subset will be denoted - $\mathbf{g}_{i\mathbf{U}_i}$. Let the limiting probability of such a state be $a(\mathbf{g}_{i\mathbf{U}_i}) = a_{ik}$ (the state number is *k*, $k \in \mathbf{K}(\mathbf{U}_i)$, where $\mathbf{K}(\mathbf{U}_i)$ is the set of state numbers included in the subset) and for the formation of a block from this state, the receipt of an document of the flow *n* is required. Then, the probability of occurrence of a block from this state is equal $a_{ik} q_{in}$. In this case, the probability of formation of a block when the next document of the total flow of documents is arrived is calculated as: $z_i = \sum_{\substack{j=1 \\ (j \in \mathbf{K}(\mathbf{U}_i)}}^{N} a_{ij} q_{ij}$, *i* = 1, 2,..., *M*.

Among the set of states of the system, we distinguish a subset of those for which, when a document of the flow *j* arrives, it will be lost - $\mathbf{H}_{ij}$. This subset of is denoted - $\mathbf{K}(\mathbf{H}_{ij})$. The probability of losing of document of the flow *j* of the block type *i*: $P_{ij} = \sum_{\substack{k=1 \\ k \in \mathbf{K}(\mathbf{H}_{ij})}}^{K_i} a_{ik} q_{ij}$, *i* = 1, 2,..., *M*; *j* = 1, 2,..., *N*.

Here is the limiting probability of getting the system to the state number *k*.

The probability of loss of any document constituting a block of type *i*: $P_i = \sum_{\substack{j=1 \\ m_{ij} \neq 0}}^{N} P_{ij}$, *i* = 1, 2,..., *M*.

After carrying out calculations on the constructed mathematical model, we obtain a set of characteristics of the of document flows that make up blocks of different types: sets of matrices: { $\mathbf{P}_1, \mathbf{P}_2, ..., \mathbf{P}_M$ } and { $\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_M$ }; a set of parameters of blocks flows: { $\Lambda_1, \Lambda_2, ..., \Lambda_M$ }; a set of values of the probability of loss of documents: { P_{ij} }, { P_i }, where (*i* = 1, 2, ..., *M*; *j* = 1, 2,.., *N*).

On practice, there are quite often cases when a block includes one document or when a block consists of a group of documents of only one flow. Applying the developed model for these cases, we obtain, for example: *M* = 1, *N* = 2, *h* = (1,1). The matrix of transient probabilities has the form:

$$\mathbf{P}_1 = \begin{pmatrix} 0 & q_{12} & q_{11} \\ q_{11} & q_{12} & 0 \\ q_{12} & 0 & q_{11} \end{pmatrix}.$$ The row of limiting matrix:

$a_{11} = (1 - q_{11})(1 - q_{12}) / (1 - q_{11} q_{12})$;

$a_{12} = q_{12}(1 - q_{11}) / (1 - q_{11} q_{12})$; $a_{13} = q_{11}(1 - q_{12}) / (1 - q_{11} q_{12})$.

The rate of the flow of blocks:

$$\Lambda_1 = \frac{\lambda_1 \lambda_2 (\lambda_1 + \lambda_2)}{\lambda_1^2 + \lambda_2^2 + \lambda_1 \lambda_2}.$$

The probability of losing documents of the first and second flows: $P_{11} = a_{13}$, $P_{12} = a_{12}$.

Blocks of different types form a common queue, and are served one at a time in the order of receipt. The rate of the total flow of blocks:

$$\Lambda^* = \sum_{i=1}^{M} \Lambda_i \cdot$$

This system with the total flow of blocks can be considered as a queueing system type **M**/**G**/**1**/∞ [6].

The obtained results allow to calculate the loss probabilities of the documents, the parameters the flows of blocks and to explore the work of processing node.

V. CONCLUSION

The results can be considered as a method of preliminary design of the distributed registri of state institutions.

Mathematical model can be used to optimize the parameters of the registry.

REFERENCES

[1] Беларев И.А., Обаева А.С. О ра пределенном реестре и возможности его применения // Финансы: теория и практика. 2017. № 2(21). С 94–99. [Belarev I.A., Obaeva A.S. Distributed Ledger and its potential application. // Finance // Theory and practice. 2017. №2(21). P. 94–99].

[2] Dalinger J. M. Analysis of data flows in the systems with absorption of messages // Informatics and control systems: Publishing house of the Amur state University. 2012. №3 (33). P. 25–34.

[3] Kemeny J., Snell J. A finite Markov chains. Princeton; New York, 1967.

[4] Миролюбов А.Л., Саксонов Е.А. Система с комплексированием сообщений // Современные информационные и компьютерные технологии: Сб. науч. ст.: В 2 ч. / Гродненныский гос. университет им. Я. Купалы. Гродно, 2009. Ч. 2. С. 128–131.. [Mirolubov A.L., Saksonov E. A. The System with integration of the messages // The Modern information and computer technology: collection of scientific articles in 2 parts / The Ministry of education of the Republic of Belarus, Grodno State University Yankee Kupala. Grodno, 2009. Part 2. P. 128–131.

[5] Tanenbaum E., van Steen M. Distributed systems. Principles and paradigms. Upper Saddle River: Pearson Prentice Hall, 2007.

[6] Вишневский В.М. Теоретические основы построения компьютерных систем. М.: Техносфера, 2003. [Vishnevsky V. M. Theoretical bases of computer networks design. Moscow: Technosphere, 2003].

The Application of Generational Theory to Digital Content Management

Irina V. Katunina
Faculty of Economics
Dostoevsky Omsk State University
Omsk, Russia
i.v.katunina@gmail.com

Valeria E. Kashtanova
K2 Marketing
Omsk, Russia
valeria.e.kashtanova@gmail.com

***Abstract*. Despite the spread of the generational theory concepts in marketing and management, there is very little empirical research in the field. The study is aimed to investigate whether it is possible to apply the generational perspective to manage digital content. The authors revealed on the base of two experiments that the differences in values, preferences and behavior of Gen X and Gen Y reflect in their reaction on digital content.**

***Keywords:* digital content; content management; generational theory; experiment; digital marketing.**

I. INTRODUCTION

It must be highly improbable to find such a professional in business and management who is not familiar with the terms "Millennials" or "Generation X". Collectively, these labels are known as generational theory that distinguishes different generations in terms of systematic differences in values, preferences and behavior that are stable over time [1].

Since W. Strauss and N. Howe published their book, the generational theory has been influential in the fields of generational studies, marketing, and business management literature. As this seems to be of interest to business leaders who are seeking the way to enter a market or increase the productivity of a workforce, the generational perspective appears to be attractive in managing multigenerational workforce and explaining consumer behavior.

What do millennials want? What defines the generations? How can we best work together?

From this point of view, well-structured digital content management aimed to facilitate content intelligence and deliver frictionless customer experience should be built on the understanding of a customers' generational structure.

But at the same time, there is very little empirical research which supports generational theory. In fact, most of the generalized understanding of different generations (millennials or other groups) is based on assumptions. Additionally, there is some criticism that much of the literature on generational theory issues oversimplifies and overstates the depth of generational gap. Moreover, some scientists concluded that meaningful differences among generations probably do not exist in the workplace [2]..

Taking into consideration these controversial views, the purpose of the study was to consider whether it is possible to apply the generational perspective to manage digital content.

The paper is structured as follows. Firstly, the brief review of the generational perspective and its application is given. Secondly, the authors present research methodology and empirical findings. The paper ends with conclusions and several directions for future research.

II. LITERATURE REVIEW

Despite the criticism, a number of recent studies have revealed the use of the generational perspective in explaining such issues as:

- leadership, ethics and human recourse management for different generations [3–5];
- consumer behavior [6, 7];
- configuring digital products [8–10].

Nevertheless, one of the problems in the field is connected with the definition of generations.

Strauss and Howe define a social generation as the aggregate of all people born over a span of roughly twenty years or about the length of one phase of life [1].

Deloitte experts [2] shown in their review stress that generational differences in any society are shaped by the political, socioeconomic and cultural events and illustrate how this variability plays out across such countries as India, South Korea, Japan, Russia, Bulgaria, Czech Republic, South Africa, Brazil, and USA.

In line with the statement that while the idea of generational differences is to some extent universal, the generations are defined specifically to a given society, the authors define Russia's Gen X and Gen Y as follows:

- Gen X (1965–1983) experienced the period of an apparent freefall of society and economy tends to gravitate to roles of being their own boss as well as to become self-taught. Gen X demonstrates informality of views, opinion expression, individualism, search for emotions, self-reliance, independent decision-making and practice orientation.
- Gen Y (1983–2000) being the first truly post-Soviet generation demonstrates a willingness to sacrifice work-life balance in exchange for quick benefits. Gen Y tends to be naive, able to obey, opportunistic, focused on entertainment, they are more theorists and prefer group decision-making.

Keeping these statements in mind, a set of research questions guided the empirical research were formulated. Are there any differences in digital content perception among Gen X and Gen Y? Whether it is possible to manage digital content in accordance with the generational structure of consumers?

III. RESEARCH METHODOLOGY

The chosen research design was an experiment (design pre-test – post-test experimental and control groups design).

Hypotheses put forward by the authors are as follows:

- Hypothesis 1.1: Gen X is more susceptible to call-to-action digital content than Gen Y.
- Hypothesis 1.2: Gen Y is more susceptible to the influence of viral digital marketing creating a "majority effect" and positioning product as a trend than Gen X.
- Hypothesis 1.3: Gen X is more receptive to media advertising with the banner than Gen Y.

In order to test Hypotheses 1.1 and 1.2, the authors conducted two experiments on the base of two online schools for learning programming languages (Schools A and B).

Both schools have official webpages and public groups in a popular social network offering similar information and educational digital content to their subscribers.

Table 1 demonstrates the generational structure of the public groups.

TABLE II. Generational Structure of the Schools' Public Groups

Public Group	**Proportion of Subscribers, %**		
	Gen X	***Gen Y***	***Others***
School A	27	58	15
School B	30	55	15

Both public groups have the standard content plan including 3 post per day, 90 posts per month:

- entertaining content 30% (27 posts per month);
- educational content 40% (36 posts per month);
- user generated content 20% (18 posts per month);
- call-to-action content 10% (9 posts per month).

The first experiment (experiment 1) was conducted in the schools' public groups. Subscribers of School A public group were considered as the experimental group, subscribers of School B public group – as the control group. Table 2 illustrates the experimental design.

TABLE III. Experimental Design

Group	**Pre-test**	**Treatment**	**Post-test**
Experimental	O_{e1}[a]	yes	O_{e2}[b]
Control	O_{c1}[c]	no	O_{c2}[d]

[a.] The outcome of pre-test for the experimental group

[b.] The outcome of post-test for the experimental group

[c.] The outcome of pre-test for the control group

[d.] The outcome of post-test for the control group

According to the experimental design, the result of the treatment can be calculated using the formula (1)

$$\text{Net outcome} = (O_{e2} - O_{e1}) - (O_{c2} - O_{c1}). \quad (1)$$

$$\alpha + \beta = \chi. \quad (1)$$

Both groups were pretested in terms of their monthly activity (the amount of comments, likes and reposts). The experimental group received the treatment (increasing call-to-action digital content from 9 to 25 posts per month) and both groups are post-tested to examine the effects of this manipulation.

The second experiment (experiment 2) was conducted on the basis of an application for the photo and video content exchange through the official webpages of both schools.

The design of the experiment was the same. Both groups were pretested in terms of their monthly purchase activity. The experimental group received the treatment (live user digital content) and both groups are post-tested.

In order to test Hypothesis 1.3, the authors collected data on user activity in terms of impressions, referral traffic and click-through rate (CTR) placing a banner advertisement of school A. Click-through rate can be calculated using the formula (2):

$$\text{CTR} = \text{Number of Click-Through} / \text{Number of impressions}, \quad (2)$$

$$\alpha + \beta = \chi. \quad (1)$$

Placement channels chosen were a popular social network, a partner application, and a partner website. The format of the posted advertisement included the title, the address, the website, the clickable banner and the description.

IV. RESULTS

For experiment 1, the authors collected data on user activity before and after the introduction of call-to-action content. This kind of content has an "incentive" message, urging the user to perform a specific action. In the case of an experiment, calls-to-action were as follows: leaving a comment, putting a like, making a repost.

Experiment 1 has shown that Gen X is more susceptible to call-to-action digital content.

According to the results of the pre-test, in the experimental group during the previous month user activity was 109 actions, 65 of which were committed by Gen X and 44 actions by Gen Y. The total activity of the control group was 105 actions, 64 and 41 of which were made by representatives of the generations X and Y respectively.

During the month of intensive call-to-action digital content for the experimental group, the total number of actions increased by 52 and amounted to 161, of which 101 actions were done by Gen X and 60 actions by Gen Y. Activity in the control group was observed at the level of 60 actions of Gen X and 45 actions of Gen Y.

Tables 3 and 4 summarize the outcomes of pre- and post-tests.

TABLE IV. Outcomes for Gen X (experiment 1)

Group	**Pre-test**	**Treatment**	**Post-test**
Experimental	65[a](6 + 57 + 2)[b]	yes	101(25 + 71 + 5)
Control	64(3 + 60 + 1)	no	60(5 + 55 + 0)

[a.] The total outcome

[b.] Comments + Likes + Reposts

TABLE V. **Outcomes for Gen Y (Experiment 1)**

Group	**Pre-test**	**Treatment**	**Post-test**
Experimental	44 (1+43+0)	X	60 (3+57+0)
Control	41 (0+41+0)		45 (1+44+0)

According to the formula (1), the net outcome for Gen X is 40, while for Gen Y is just 12:

$$\text{Net outcome}_{e1} = (101 - 65) - (60 - 64) = 40;$$

$$\text{Net outcome}_{c1} = (60 - 44) - (45 - 41) = 12.$$

Thus, Hypothesis 1.1 has been confirmed.

Moreover, calculating the impact of the treatment on the comments, likes and reposts separately (Fig. 1), it was revealed that 11 out of 12 of Gen Y net outcome were "like-actions" while for comments and reposts the impact was 1 and 0 respectively.

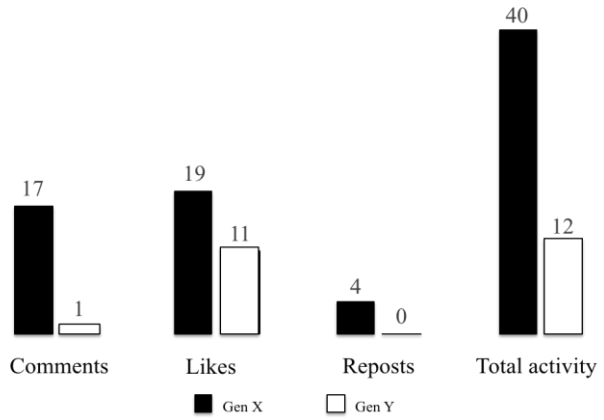


Fig. 1. The impact of the treatment on the amount of comments, likes and reposts

This result supports Hypophysis 1.1 as "like-actions" seem to be to some extent impersonal, while comments and reposts are an open and detailed demonstration of an individual point of view or attitude.

In Experiment 2, the authors collected data from the payment systems on the amount of subscriptions sold out through the schools websites. The price of one course subscription is the same for both schools (500 rubles). The pre-test shows that school A had sold out 11 subscriptions during the month before the treatment (4 to Xers and 7 to Yers). During the same period, school B had sold out 43 subscriptions (19 to Xers and 24 to Yers).

The treatment was the launch of live user content such as demonstration of the training process or demonstration of the school's brand in the context of "life style". This live user content was distributed by using geolocation to create the effect of mass use and "trend."

The post-test data collection began two weeks after the launching the viral marketing in school A in order to obtain sufficient coverage and generate the required amount of content.

Thus, for the month during which the viral marketing was launched in school A, it sold out 7 subscriptions to Xers and 16 to Yers. School B sold 23 and 21 subscriptions respectively. Tables 5 and 6 summarize the outcomes of pre- and post-tests.

TABLE VI. **Outcomes for Gen X (Experiment 2)**

Group	**Pre-test**	**Treatment**	**Post-test**
Experimental	4	X	7
Control	19		23

TABLE VII. **Outcomes for Gen Y (Experiment 2)**

Group	**Pre-test**	**Treatment**	**Post-test**
Experimental	7	X	16
Control	24		21

Experiment 2 has confirmed that Gen Y is more susceptible to viral digital marketing. According to the formula (1), the net outcome for Gen X is –1, while for Gen Y this indicator is +12:

$$\text{Net outcome}_{e2} = (7 - 4) - (23 - 19) = -1;$$

$$\text{Net outcome}_{c2} = (16 - 7) - (21 - 24) = 12.$$

Thus, Hypothesis 1.2 has also been confirmed.

Finally, the authors examined the data obtained from media advertising on a popular social network, a partner application, and a partner website. To determine how susceptible the representatives of different generations to media advertising, from the data obtained, CTRs for all channels for each generation group were calculated.

The results are presented in Table 7.

TABLE VIII. **Media Advertising on Various Sites**

Indicator	Channel	Generation			Total
		Gen X	Gen Y	Others[a]	
Impressions	Partner website	9275	16526	16001	41802
	Social network	117458	192114	1901	311473
	Partner application	6715	9536	4032	20283
Referral traffic	Partner website	423	396	277	1296
	Social network	857	1091	8	1956
	Partner application	107	81	46	234
Click-through rate, %		2.3	1.3	1.1	

[a] Users under the age of 18 and older than 53 as well as those whose age was not identified (unauthorized users).

The advertisements placed on the site and in the partner application were not targeted and were shown to all users. The advertisements placed on the social network were configured to demonstrate to users of 20–50 years old, which explains the relatively low indicator of the "other". The average CTR showing the percentage of users who interacted with the advertisements as well as the total number of users to whom it was demonstrated for Gen X is 2.3%, while for Gen Y it is 1.3% and "other" is 1.1%. Thus, it can be concluded that Gen X is more responsive to media advertising than Gen Y. Hypothesis 1.3 has also been confirmed.

V. CONCLUSIONS AND FUTURE RESEARCH

The findings contribute to understanding of consumer perception of digital marketing tools in order to form effective programs for managing digital content on the Internet.

It was investigated how the system of values formed in the historical context influences the perception of digital marketing tools. The results confirm the relationship between the key behavioral characteristics of users belonging to different generations and the peculiarities of their behavior in interaction with digital marketing tools.

Experiment 1 has shown that Internet users considered as representatives of generation X are more susceptible to content such as call-to-action. These people are characterized by a high need for expressing their individuality, they like to show their attitude and express an opinion, while generation Y, on the contrary, strives for unification and imitation, "be with everyone or be alone."

Based on the results of experiment 2, it can be concluded that Gen Y is more susceptible to the tools of viral marketing than representatives of Gen X, which is dictated by the desire of Gen Y to comply to, to be a part of a specific group ("be in trend") and make decisions based on the opinion of the majority.

Finally, it was concluded that Gen X is more responsive to media advertising than Gen Y. Gen X is focused on searching information, cognition and action that is also manifested in their consumer behavior within the framework of traditional non-digital marketing.

It should be also noticed that Gen Y has been accumulating its experience in the era of digital and ICT technologies. Growing up in the information garbage, they are used to "filter" incoming information, actively use ad blockers and more often explore the Internet from mobile devices with an adaptive design eliminating side columns where up to 40% of media advertising is placed.

This allows to conclude that the application of generational theory can indeed underlie the formation of maximally client-oriented promotion programs by identifying the predisposition of specific groups of consumers to various online marketing tools and their effective combination.

The findings of the study provide new challenges for companies operating in a digital world. Further research seems to be focused on developing descriptive patterns of generational consumer behavior for studying (1) advertising placement, (2) user experience for the formation of user interface, and (3) creation of customer journey map to develop an effective integrated digital strategy.

The paper has offered a preliminary results that should be enriched with the longitudinal study within the more prolonged period and within other online businesses. Moreover, in-depth case study analysis would contribute to the understanding how digital content management can be organized from the generational perspective.

VI. REFERENCES

[1] Strauss W., Howe N. The fourth turning : an American prophecy. New York: Broadway Books, 1997.

[2] Hole D., Zhong L., Schwartz J. Talking aboutwhose generation? Why western generational models can't account for a global workforce // Deloitte Review. 2010. N 6. P. 83–97.

[3] Anderson H. J., Baur J. E., Griffith J. A. et al. What works for you may not work for (Gen)Me: Limitations of present leadership theories for the new generation // The Leadership Quarterly. 2017. N 28(1). P 245–260. doi:10.1016/J.LEAQUA.2016.08.001.

[4] Welsh D. H. B., Memili E., Rosplock K., J. Roure, and J. L. Segurado, "Perceptions of entrepreneurship across generations in family offices: A stewardship theory perspective // Journal of Family Business Strategy, 2013. Vol. 4, N 3. P. 213–226.

[5] Çelikdemir D. Z., Tukel I. Incorporating Ethics into Strategic Management with Regards to Generation Y's view of Ethics // Procedia - Social and Behavioral Sciences. 2015. Vol. 207. P. 528–535.

[6] Li X., Li X., Hudson S. The application of generational theory to tourism consumer behavior: An American perspective // Tourism Management. 2013. Vol. 37. P. 147–164. doi: 10.1016/J.TOURMAN.2013.01.015

[7] Huang Q., Lu Y. Generational perspective on consumer behavior: China's potential outbound tourist market // Tourism Management Perspectives. 2017. Vol. 24. P. 7–15.

[8] Boonsiritomachai W., Pitchayadejanant K. Determinants affecting mobile banking adoption by generation Y based on the Unified Theory of Acceptance and Use of Technology Model modified by the Technology Acceptance Model concept // Kasetsart Journal of Social Sciences. In press. doi:10.1016/J.KJSS.2017.10.005.

[9] Freestone O., Mitchell V. Generation Y attitudes towards E-ethics and Internet-related misbehaviours // Journal of Business Ethics. 2004. Vol. 54, N 2. P. 121–128. doi:10.1007/s10551-004-1571-0.

[10] Kumar A., Lim H. Age differences in mobile service perceptions: comparison of Generation Y and baby boomers // Journal of Services Marketing. 2008. Vol. 22, N 7. P. 568–577. doi:10.1108/08876040810909695.

The Role of Professional Labor Migration in Developing an Effective Knowledge Management System in High-Technological Organization

Karina V. Sayapina
Department of Management
Financial University under the Government of the Russian Federation
Moscow, Russia
k.v.sayapina@gmail.com

***Abstract*. The paper is devoted to search of instrument, based on labor migration aspect and consequent cross-cultural peculiarities, for building an effective knowledge management system in high-tech company. The instrument is approved by case study method in large international high-tech corporations.**

***Key words:* knowledge management; cross-cultural factors, high-tech organization; labor migration .**

I. INTRODUCTION

Modern problems and tendencies in labor migration are the object of important large-scale research during recent years. Studies in labor migration sphere are especially actual in European continent. It is because of multinational economic area of the "Big Europe", including Russia, outside of the EU, and of intensifying labor migration processes in conditions of developing economic, social and cultural contacts in a global world. Research of labor migration processes from the CIS countries are also of great meaning, so far as with education and general liberalization in sphere of labor migration there, this aspect nowadays is of mass proportions with serious social, economic and political consequences.

Nowadays society goes through active process of globalization: there is constant exchange of information, technologies and other resources, including human ones, in different organizations' activity. Development and acceleration of such exchange are realized so fast, that a company being without international status, however, could consist of representatives from different nations and cultures all over the world. On the other part, in high-technological companies management aims to create multinational multicultural teams in order to form special unique atmosphere for interaction of different ideas and scientific schools to increase effectiveness in elaborating new strategy course, and thus to have a resultant knowledge management system. That is why for managers, responsible for decision-making process, significance of taking into account cross-cultural peculiarities in organizational strategy is risen. This is also of significant importance, because for number of actively developing "knowledge economy" sectors forming, necessary team of specialists is realized by mean of labor migration, i.e.attraction of high-qualified personnel from different states all over the world.

II. LITERATURE REVIEW

In evolutionary development of strategic management theories, increasing role of accounting cross-cultural aspect in forming knowledge management system has already been traced by the end of XX century [6].

First works in sphere of strategic management appeared in 60-s of XX century. Here, the main accent was absolutely in industrial and financial aspects. One of the main founders of this course is I. Ansoff (martix "product-market", "Scheme of strategic planning", hierarchy of aims and strategies of organization) [1]. In 70-s of XX century results of his work were continued in further development by Harvard Business School: SWOT-analysis was elaborated (which is strengths, weaknesses, opportunities and threats) in frames of accounting internal and external factors of organization's activity [4]. In the same period, R. Normann designed a concept of building strategy on basis of business-idea (target segment identification, product formation under needs in this segment, creation of internal organization due to its actives, structure and management procedures) [13]. This concept was improved in the end of 70-s by Ch. Hofer and D. Schendel: knowledge effectiveness strategy formation is realized in terms of special algorithm: market identification and searching company's main competences, forming competitive advantage, search of synergy inside the company [9]. In 80-s M. Porter, taking into account sharpened competition, developed a concept of industry's competitive analysis, competitive advantage (with formation of further appropriate strategy) and value chain [14]. In 90-s, in conditions of business environment changes, theories of knowledge management were elaborated on basis of internal and external organizational environment (which is necessity to react on these changes and search for new decisions in context of macroeconomic factors) [12]. However, these factors also had rationally technical character, and problems of cross-cultural aspect hasn't still been taken into account. It was H. Minzberg theory about forming knowledge management system in terms of organization's internal resources development, where for the first time cross-cultural aspect has been reminded; it is realized through special personnel 'school of training' improving its professional skills and methods of mutual interaction. The main task of interaction control was to solve potential cross-cultural problems in organization [11].

In Russian exploratory society, publications of S. Ryazantsev play significant role, where impact of labor migration in knowledge management system are estimated, mainly in the Commonwealth of Independent States (CIS) and Baltic countries, and in the European Union (EU), in the USA, in the BRIC countries (Brazil, Russia, India and China) [15].

Specified review of existing theoretical works towards building an effective knowledge management system shows frames and limitations in existed approaches to investigation of this problem, because cross-cultural difficulties as a natural consequence of global labor migration are insufficiently considered in these approaches.

III. THEORETICAL GROUNDING & METHODOLOGY

The research of influencing professional labor migration aspect and appropriate cross-cultural factors in building knowledge management system is held in two stages. On the first stage analysis of common knowledge management system, used mainly in high-technological organizations where culture is considered as one of its resources, is held (Fig. 1).

According to the model, solving of definite tasks is realized by cultural factors, which work as internal stimuli influenced on knowledge management system of a company. Unique knowledge, created on basis of different types of resources, and common knowledge, cultural knowledge and cross-cultural technologies, is a competitive advantage of a company. Due to analysis of market reaction towards this competitive advantage, necessary improvements and changes in its internal structure are realized, and these changes are faster and more qualitative respond to market needs of definite region. That's why a company could be equal here to "portfolio of abilities, not types of business" [8, p. 29].

It is necessary to underline that simple recruitment process on international level is not so effective in context of potential cross-cultural problems. Thus, in frames of preventing possible negative consequences in this process, and due to importance of cross-cultural technologies aspect, the author suggests algorithm of attracting foreign specialists, and immediate solution of cross-cultural problems (Fig. 2).

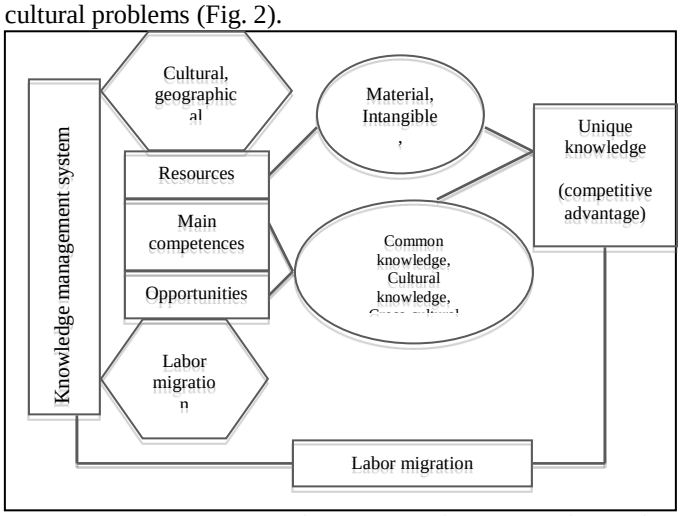


Fig. 1. Model of organizational functioning in frames of culture factor (made by the author on basis of [5; 7; 8; 10])

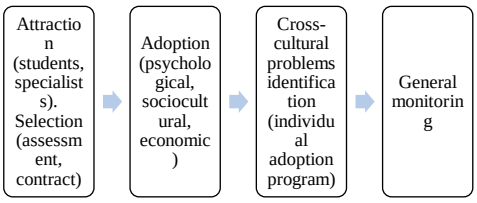


Fig. 2. Algorithm of acctracting foreign specialists in organization

Attraction stage assumes involvement of the most talented, successful and smart students from leading higher educational institutions all over the world in appropriate high-tech industry, and similar specialists in country of origin and abroad. It can be realized by competitive intelligence method, active cooperation with international organizations in appropriate sphere, and through participation in international conferences, seminars and forums. Selection consists of assessment process of specialists attracted (interviewing, identifying professional and communicational skills), and signing contract with selected foreign specialist under mutual agreement.

Adoption stage is really important and supposes psychological adoption (achievement of psychological satisfaction inside new environment depending on personal abilities of foreign specialist and his expectations towards new place of life and work, his internal family atmosphere), socio-cultural adoption in country of origin (on basis of acquiring social skills, solving socio-cultural problems in everyday life, and active involvement of foreign specialist in cultural life of the country), and economic adoption through special educational programs, studying on working place, trainings and motivation program for personnel.

Cross-cultural problems identification stage is built on principals of assigning definite code of rules to simplify communication and overcome interactional barriers inside personnel, analysis of specialist's communication among staff, and concrete difficulties, appearing in this process; analysis of cultural similarities and differences, leading to misunderstanding among staff; and analysis of appearing conflict situations inside organization. Then special individual adoption program is elaborated, which main goal is to hold the fastest and painless homogeneous adoption of foreign specialist on basis of the following factors:

- cultural differences of foreign specialist in comparison with culture of the country of origin;
- spectrum of duties for specialist in context of his level of professional language, which is official working language in an organization;
- expectations of foreign specialist towards new work;

- life perception by foreign specialist: personal perception of cultural environment in the country of origin, perception of business atmosphere inside organization;
- personal attitude of foreign specialist to moral norms and traditions of the country of origin;

TABLE I. CROSS-CULTURAL PARAMETERS SYSTEM

Factors	Parameter	Meaning
Cultural	Foreign specialists, %	30
	Specialists with higher education diploma, %: total including foreigners	 70 25
	Specialists with Ph.D. degree, %: total including foreigners	 10 50
Branding	SME grants, % (in total financing)	2
	Membership in international associations, (minimum level of participation)	3–5
Economic	Venture investment projects, % (in total financing)	10
Institutional	Number of foreign specialists attraction programs	5–7
	Number of exchange education programs (including qualification and experience exchange)	3–5

Finally, general monitoring stage is held by interviewing, questionnaire poll, and planned meetings with HR-manager (or psychologist) with discussion of potential problems during interaction process among staff, in society during free time. It is also important to estimate turnover control, volume of hiring and dismissal, expenditures on personnel by different reasons, and to make plan of estimation for more operative and strategic decisions in terms of human resources management and knowledge management system in organization.

The second stage of the research is devoted to assessment of cross-cultural factors by case-study method for two international high-technological companies ("Arkema" and "BIOCAD") on basis of the specially elaborated cross-cultural parameters system as a result of cross-cultural analysis of main Silicon Valley high-tech companies (Tabl. 1).

After cross-cultural factors assessment elaboration of practical recommendations for using the instrument of attracting high-qualified foreign personnel in high-tech industry is held.

IV. EMPIRICAL ANALYSIS

Two large high-technological corporations (French organization "Arkema", specialised in industrial and special chemistry, and Russian biopharmaceutical organization "BIOCAD") are good examples of using cross-cultural factors in frames of buidling effective knowledge management system.

In 2017 sales volume in Arkema corporation was 8,3 bln euro (with almost 3% of which was dedicated to R&D direction) [2]. In the end of 2017 total number of specialists was 19800 (where 47% are from France, 13% — other EU states, 18% - the USA, 20% — Asian countries) [2]. In frames of common knowledge management system, maximisation of personnel abilities is achieved through internal experience exchange in the corporation's subsidiaries (France, Belgium, Denmark, Germany, Hungary, Italy, Poland, Spain, Russia, Switzerland, Great Britain, USA, Canada, Australia, Brazil Algeria, UAE, Egypt, Turkey, Japan, China, India, Malaysia, South Korea, Thailand).

TABLE II. CROSS-CULTURAL PARAMETERS SYSTEM: ARKEMA & BIOCAD

Factors	Parameter	Meaning	Arkema	BIOCAD
Cultural	Foreign specialists, %	30	25 (rank 0,83)	12 (rank 0,4)
	Specislists with higher education diploma, %: total including foreigners	 70 25	 25 (rank 0,83) 40 (rank 1,6)	 12 (rank 0,4) 20% (rank 0,8)
	Specialists with PhD degree, %^ total including foreigners	 10 50	 8 (rank 0,8) 40 (rank 0,8)	 6 (rank 0,6) 15 (rank 0,3)
Branding	SME grants, % (in total financing)	2	0,5 (rank 0,25)	0 (rank 0)
	Membership in international associations (minimum level of participation)	3-5	More than 5 (rank 1)	More than 5 (rank 1)
Economic	Venture investment projects, % (in total financing)	10	7 (rank 0,7)	5 (rank 0,5)
Institu-tional	Number of foreign specialists attraction programs	5–7	2 (rank 0,4)	5 (rank 1)
	Number of exchange education programs (including qualification and experience exchange)	3–5	More than 5 (rank 1)	More than 5 (rank 1)
Sum of ranks			8,16	5,57

Unique knowledge of "Arkema" corporation is created through the following key programs of foreign specialists' adoption:

- "International experts" program, where leading specialists are sent in different foreign subsidiaries for special project management;
- "Development and qualification" program: smart and talented specialists have an opportunity to work abroad for 2–3 years and to share experience after finishing the program;

- "Influence and education" program: it is short-term internship to get successful experience and share the knowledge among colleagues.

Cross-cultural problems are prevented through holding special interviews of specialists by cross-cultural committees inside the organization. According to algorithm of attracting foreign specialists, main stages are adoption and general monitoring ones. Adoption of foreign specialists is held in the "Arkema" subsidiaries. Representatives of cross-cultural committees are responsible for general monitoring of atmosphere and satisfaction among specialists.

Russian "BIOCAD" company is much smaller than "Arkema", but it is also innovation-oriented high-technological company and it is among leaders of Russian pharmaceutical companies [3]. Level of turnover was 14 bln rubles in 2016 [3]. Subsidiaries of the company are located in Belarus, Ukraine, the USA, Brazil, India and China. Number of personnel is 1300, including 450 high-qualified specialists, many of them are members of Russian and international professional associations [3].

Great role of building effective knowledge management system in the company belongs to students and young specialists attraction, which is realized through the following directions:

- trainings in R&D centers with appropriate externship;
- holding scientific research works on basis of R&D centers;
- bioinformatics projects.

Key stages of the algorithm for attracting foreign specialists, mentioned above, are students attraction, preliminary long-term labor contract, adoption of new specialists (with acquaintance of cross-cultural features inside the company), and general monitoring analysis (which is held more often in American subsidiary, than in Russian one).

Summarizing, it is necessary to underline that the most important stage of foreign specialists attraction algorithm is the second stage of realizing correct adoption, because it is possible to overcome potential cross-cultural problems and conflicts in a whole organization in frames of achieving necessary aims on this stage.

Cross-cultural factors estimation is more effective to be held through special group of parameters. Periodical analysis of the meanings can help to build an effective knowledge management system of an organization.

V. REFERENCES

[1] Ансофф И. Новая корпоративная стратения СПб.: Питер, 1999. [Ansoff I. New corporate strategy. St Petersburg: Piter, 1999].

[2] Reference document 2017 including the Annual Financial report // Arkema. URL: https://www.arkema.com/export/sites/global/.content/medias/downloads/investorrelations/en/finance/arkema-2017-reference-document.pdf.

[3] BIOCAD company. URL: https://biocad.ru/uploads/ru/files/BIOCAD_15.04.pdf.

[4] Fine L.G. The SWOT analysis: using your strength to overcome weaknesses, using opportunities to overcome threats. 2009. URL: http://lawrencefine.com/downloads/SWOT%20-%20PDF.pdf.

[5] Финкельштейн С. Ошибки топ-менеджеров ведущих корпораций. Анализ и практические результаты. М.: Альпина Бизнес Букс, 2006 [Finkelshtein S. Top-managers' mistakes of leading corporations. Analysis and practical results. Moscow: Alpina Business Books, 2006].

[6] Friedman T.L. The World is Flat: A Brief History of the Twenty First Century. [S.l.:] Findaway World, 2006.

[7] Гусева Н.И. Теоретические и методологические основы кросс-культурных исследований // Известия Иркутской государственной экономической академии. 2003. № 3–4. С. 82–86. [Guseva N. I. Theoretical and methodological principles of cross-cultural research // Izvestiya of Irkutsk State Economic Academy. 2003. №3–4. P. 82–86].

[8] Hannan M.T., Freeman J. Organizational ecology. Cambridge: Harvard University Press, 1993.

[9] Hofer C. W., Schendel D. Strategy formulation: analytical concepts (the West series in business policy and planning). St Paul: West Publishing Company, 1978.

[10] Холден Н.Дж. Кросс-культурный менеджмент. М.: Юнити-Дана, 2005. [Holden N.J. Cross-cultural management. Concept of cognitive management. Moscow: Yuniti-Dana, 2005].

[11] Mintzberg H. The structuring of organizations. Prentice Hall, 1979.

[12] Moore F. Transnational business cultures: life and work in a multinational corporation (cross-cultural management). Hants: Ashgate Pub Ltd, 2005.

[13] Normann R., Ramirez R. Designing interactive strategy: from value chain to value constellation. Chichester: Wiley & Sons, 1998.

[14] Porter M. E. Competitive strategy: techniques for analyzing industries and competitors. New York: Free Press, 1998.

[15] Рязанцев С.В. Глобальный рынок труда и международная миграция. М.: Экономика, 2010. [Ryazantsev S.V. The global labor market and international migration. Moscow: Economics Publishing House, 2010].

Web-Environment Supervision as a New Form of Internet Socialization

Marina S. Chvanova[1], Ekaterina D. Saushkina[2]
Department of Information Systems and Technologies
K.G. Razumovsky Moscow State University of Technologies and Management (the First Cossack University)
Moscow, Russia
[1]tmbtsu@gmail.com, [2]tmbtsu@gmail.com
Maria S. Anurieva[3], Alexey V. Samokhvalov[4], Anatoly A. Molchanov[5]
Department of Mathematical Modeling and Information Technologies
G.R. Derzhavin Tambov State University
Tambov, Russia
[3]anuryeva@mail.ru, [4]samohvalov@gmail.com, [5]ykdosto@gmail.com

***Abstract*. The article considers the issues of Internet-socialization of computer science students as a part of their interaction with professional community on current problems of the subject field that are in constant transformation. Results of innovation course "Information Reengineering" design, based on informational and educational system modeling, are presented in the article.**

***Key words:* information reengineering, Internet socialization.**

I. INTRODUCTION

One of the main problems associated with the training of modern personnel is the creation of conditions conducive to the development of specialists capable of solving professional tasks that are continuously being modified, and, as a result, the methods, means and technologies for solving them are changing. The need for constant interaction of educational process participants in professional community is increasing. This is expressed not only in the need to attract specialists from among the managers and employees of organizations whose activities are related to the specialization of the educational program being implemented, as faculty members, but also in their active participation in the formation of basic educational programs, the creation of new academic disciplines and practice programs. This problem is particularly acute in the training of IT professionals, when the subject field of studied disciplines is in constant transformation: not only the content of previously taught disciplines changes continuously, but completely new ones appear.

For example, due to the adoption of the Concept of Information Society Development in Russia for 2017–2030, a number of directions for the development of Russian information and communication technologies have been formed and identified: Internet of Things and the Industrial Internet, Big Data processing, cloud and fog computing, trusted technologies of electronic identification and authentication including financial sphere and others.

At the same time, educational disciplines and their components are not formed due to the emergence of new knowledge. The understanding of a new subject field often occurs during the process of analyzing and integrating professional practice. At the same time, discussion of topical issues by students and representatives of professional community is very important for the professional socialization of young people.

II. RESEARCH

Communication with professional community in the process of learning through project activities.

One of the objectives of the national priority project “Science” is creation of world-class scientific and educational centers based on universities and research organizations, their cooperation with the professional community operating in the real sector of the economy. Training of modern personnel should be carried out in conditions conducive to the development of specialists capable of solving professional tasks that are continuously transforming, and their solution requires the use of new tools, which leads to an improvement in quality (the emergence of new methods and technologies) and quantitative (increase in execution speed, volume of tasks) results. Professional tasks can both be transformed from old ones, adapting to the new conditions of scientific and technical progress, and be fundamentally new, not previously solved, and requiring fundamentally new means, methods and solutions [1–4]. This forces the process of training specialists to constantly make changes in both the content and the organization.

Knowledge of the fundamentals of project, research, innovation activity, introduction of entrepreneurial ideas into the content of existing courses are becoming especially necessary [5].

Various pedagogical technologies (project method, problem-based training, etc.) make it possible to solve the tasks in some way. At the same time, in the conditions of the information society, their use without reliance on modern management technologies and information tools will not significantly improve the educational process. Project activities of students, which are integrated into the needs of the professional community, are coming to the fore. At the same time, the former management methods do not allow solving promising tasks, and the methods and technologies used today have become

insufficient to integrate educational projects into the professional field.

The above makes it necessary to consider the logic and algorithm of the organization of project activities in Web-environment for training high school students.

The general idea of the organization of project activities in high school is as follows. First of all, it is the solution of an urgent task, significant for the student, preparing him for future activities in the modern information society and in the real sector of the economy. Students acquire not only professional competence in design and technological activities, but also experience in teamwork, experience of interaction with colleagues and customers. Organizational and administrative competencies are formed; future IT specialist is socialized in professional community. Experience of project work increases the competitiveness of future specialists on the job market, helps to take the first step in building a successful career.

The system of communications provides various components of interaction of participants of educational process in web environment [6]: computer "imitation" of communication in traditional learning systems at the university. Since, we are talking about the training of computer science students, it is important to build additional elements to solve professional problems and carry out specific types of communication in the system. These elements are:

- Communication between student and professionally oriented community communication (search for problems in professional field, search for potential partners for a project).
- Communication between student and world information space communication (interaction in the global intellectual community of professionals, consultations are sought and carried out to find solutions to professional problems and potential project participants).
- Communication between student and innovative infrastructures communication (characteristic feature of the information industry is research activity, while relevant employees must have the skills, forecasting techniques, readiness to work in new areas and activities). A specialist must be able "to sell" a relevant product, economically and efficiently explaining the benefits and corresponding risks in development of high-tech products. This forces to include in to the developing model not only "classical" elements of pedagogical system as a teacher and a student, but also an innovation infrastructure, e.g. technology transfer centers, centers of marketing research, business incubators. "Open" system involves interaction with external system components, e.g. venture funds, representatives of academic community, digital libraries).

Organization of communication in the Internet and automation of project management.

Management of innovative educational projects is one of the most difficult and labor-intensive areas of management. Project's success is determined by the efficiency of realization of project's idea, where the interests of teams and organizations working on project's implementation are concentrated. Effective implementation of the project intent is possible only with a coordinated purposeful development of all processes that ensure this implementation, with effective management of the project implementation process.

The project management process is presented in the form of a sequence of the following steps:

- defining of goals, sub-goals and objectives;
- identifying the resources needed to complete tasks;
- drawing up a project implementation plan;
- initiating the process of starting implementation;
- project development tracking [7].

When developing a web-based environment for communication of subjects of the educational process with the professional community, the possibility of passing each step should be realized.

Web- environment for socialization of students in professional community.

The high dynamics of development of the subject field and courses studied by computer science students necessitates the constant transformation not only of content, but also of educational technologies. Actual educational disciplines are often set forth fragmentary and not formed due to the formation of a new knowledge. Understanding of the new subject field occurs in the process of theoretical analysis and professional practice. It happened, for example, with the development of the concept of "Information reengineering". Components of this process have long been used in various areas of computer science, but a holistic view is not yet sufficiently meaningful. The above dictates the need to create a new course "Information Reengineering" based on modeling a web environment to support student's training. In this case, the term "information and education system" means a complex that includes computing and communication equipment, software and personnel, providing support for a constantly changing information system.

The set of elements included in the structure is determined by the purpose of the system, in our case we mean the improvement of training and virtual socialization of computer science students in a particular educational institution or as a model for a group of universities that train specialists of the same profile.

The content of ILE was formed according to knowledge and skills necessary to perform labor functions in information reengineering, identified in professional standards for informatics (№ 616N "Information Technology Manager" [8], № 645N "Head of software development" [9], № 893N "Project Manager in the field of information technologies" [10]).

Professional standards determine not only the profession and generalized work functions, but also the qualification level of a specialist, as well as the scope of application and specific work activities, the necessary knowledge and skills in performing these functions. The range of duties have a clearly defined information orientation, while the need for knowledge of

information reengineering processes is not reflected in the standards, which indicates a lack of attention to the on-time elimination of problems in the information processes of enterprises.

The educational course we have created contains theoretical material, laboratory work, practical tasks, and test tasks on the topics of “Information Reengineering”, which are necessary to solve the tasks set for a future specialist.

The material is divided into 3 units of “Information reengineering” topics:

- The first one includes means of identification organization’s IT resources needs, ways to identify IT service needs and change management.
- The second unit includes technical specification of software, team information exchange.
- The third block includes quality management in IT projects, requests for changes in IT projects, improvement of the efficiency of project management system and risk management in IT projects.

“Information Reengineering” web-system was posted on the Internet (Fig. 1); Information Reengineering group was created in Vkontakte social network (Fig. 2). Communication in the group contributes to the online socialization of students in professional community when discussing the problems of new knowledge development.

Fig. 3. Page of informational educational system “Information reengineering”

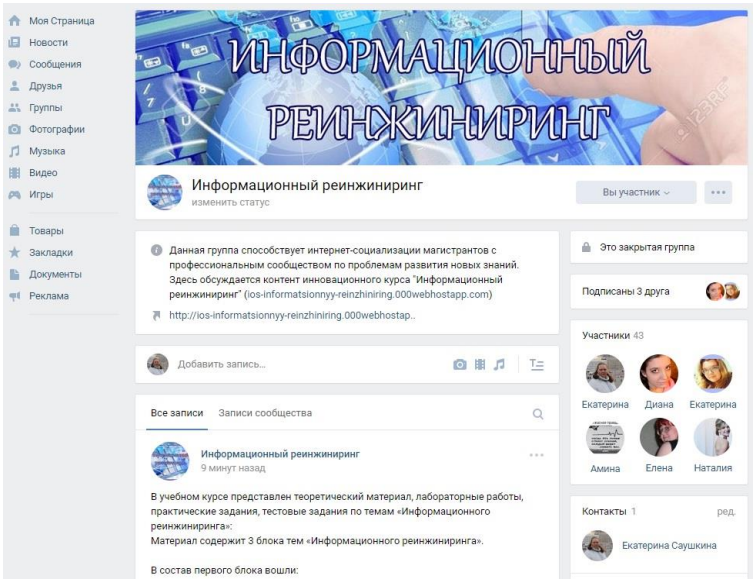


Fig. 4. Group in the social network “Vkontakte”

III. RESULTS OF THE STUDY

Educational process of training computer science students is currently carried out in high dynamics conditions of studied disciplines’ subject field. New disciplines are being introduced; traditional disciplines are being modified by including new topics: the Internet of Things, Industrial Internet, block chain, smart technologies, information reengineering, Big Data. A joint discussion of new problems of computer science by students and representatives of professional community actualizes the

creation of specialized web environment that contributes to professional socialization of young people.

Web environment provides various components for interaction of educational process participants, which include the following elements: communication between student and professionally-oriented community, between student and world information space. between student and innovative infrastructures.

Results of our research are implemented in practice of quality management of educational process in K.G. Razumovsky Moscow State University of Technologies and Management (the First Cossack University) and G.R. Derzhavin Tambov State University. Approbation of the “Information reengineering” course for computer science master degree students proved the desirability of its use. Reengineering of information processes is used both in production and banking, as well as in education, medicine, telecommunications, etc.

The study does not exhaust the full extent of the problem of socialization of computer science students in professional community, but it also provides an effective solution to one of the important problems. Obtained results can be used to further study the patterns, principles and features of Internet socialization of students in professional community. Developed web environment (for supervision of training computer science students in new course “Information Reengineering”) has a great educational potential and contributes to development of Internet socialization in professional community.

The results obtained with financial support of the Russian Foundation of Basic Research for the project “Sociological analysis of the Internet space as an indicator of the socio-cultural dynamics of the development of open education”, No 16-06-00176, 2016–2018.

Certain aspects investigated and the corresponding results obtained with financial support of the Russian Foundation of Basic Research for the project “Pedagogical basics of youth students’ socialization in Internet space and their implementation in educational system”, No 17-36-00039, 2017–2019.

IV. REFERENCES

[1] Быков А.Н. Инновационная политика в условиях глобализации // Актуальные проблемы развития мировой экономики: Сб. ст. М., 2007. С. 32–50. [Bykov A.N. Innovation policy in the context of globalization // Actual problems of the global economy: a collection of articles. Moscow, 2007. P. 32–50].

[2] Добрынин А.И., Ивлева Е.С. К вопросу о модернизации модели выцсшего образования в России // Экономика и управление.. 2011. № 5. С. 128–134. [Dobrynin A.I., Ivleva E.S. On the issue of modernization of the higher education model in Russia // Economics and Management. 2011. N 5. P. 128-134ъ.

[3] Чванова М.С., Лыскова В.Ю. Кластерный подход: профессионально-ориентированнная информационно-учебная вики-среда для студентов-информатиков // Вестник Тамбовского университета. Сер. Гуманитарные науки. 2012. № 10. С. 130–134. [Chvanova M.S., Lyskova V.Yu. Cluster approach: professionally-oriented informational and educational wiki environment for computer science students // Bulletin of Tambov University. Series: The humanities. 2012. N 10. P. 130–134].

[4] Юрьев В.М., Чванова М.С, Теоретические основы подготовки специалистов наукоемких специальностей: становление унимверситета как центра инновационно-образовательного кластера // Вестник Тамбовского университета. Сер. Гуманитарные науки. 2007. № 10. С. 7–10. [Yuriev V.M., Chvanova M.S. Theoretical bases of training specialists of science intensive specialties: University development as a centre of innovation education cluster // Bulletin of Tambov University. Series: The humanities. 2007. N 10. P. 7–13].

[5] Чванова М.С., Малышева Н.В., Киселева И.А. и др. Проектная деятельность студентов и школьников на основе кластерного подхода // Вестник Тамбовского университета. Сер. Гуманитарные науки. 2009. № 9. С. 240–253. [Chvanova M.S., Malysheva N.V., Kiseleva I.A. et al. Project activity of students and schoolchildren based on the cluster approach // Bulletin of Tambov University. Series: The humanities. 2009. N 9. P. 240–253].

[6] Юрьев В.М., Чванова М.С. Кластерный подход в подготовке специалистов наукоемких специальностей // Вестник Тамбовского университета. Сер. Гуманитарные науки. 2009. № 5. С. 872–876. [Yuryev V.M., Chvanova M.S. Cluster approach to the training of high-tech specialists // Bulletin of the Tambov University. Series: Natural and Technical Sciences. 2009. N 5. P. 872–876].

[7] Чванова М.С., Храмова М.В. Проблемы организации коммуникаций студентов наукоемких специальностей в системе открытого образования // Образовательные технологии и общество. 2011. № 2. С. 482–501. [Chvanova M.S., Khramova M.V. Problems of organizing communications of high-tech students in an open education system // Educational Technologies and Society. 2011. N 2. P. 482–501].

[8] Приказ Минтруда России от 13.10.2014 № 716н «Об утверждении профессионального стандарта “Менеджер по информационным технологиям”». [Order of the Ministry of Labor of Russia, October 13, 2014 No. 716N "On approval of the professional standard" Information Technology Manager"]. URL: http://fgosvo.ru/uploadfiles/profstandart/06.014.pdf.

[9] Приказ Минтруда России от 17.09.2014 № 645н «Об утверждении профессионального стандарта “Руководитель разработки программного обеспечения”». [Order of the Ministry of Labor of Russia, September 17, 2014 No. 645N "On approval of the professional standard" Head of software development"]. URL: http://fgosvo.ru/uploadfiles/profstandart/06.017.pdf.

[10] Приказ Минтруда России от 18.11.2014 № 893н «Об утверждении профессионального стандарта “Руководитель проектов в области информационных технологий”» [Order of the Ministry of Labor of Russia, November 18, 2014 No 893N “On the approval of professional standard "Project Manager in the field of information technologies”]. URL: http://fgosvo.ru/uploadfiles/profstandart/06.016.pdf.

The Ecosystem for Online Entrepreneurship Teaching

Natalya F. Altukhova[1], Elena V. Vasilieva[2]
Department of Business Informatics
Financial University under the Government of the Russian Federation
Moscow, Russia
[1]nfaltuhova@fa.ru, [2]evvasileva@fa.ru

***Abstract*. The authors voice the problems of teaching students technological entrepreneurship. The authors also highlight the basic requirements for the competencies of the head, that today must have knowledge in the professional and technical field, be able to think globally and have a wide managerial outlook and experience to have the gift of foresight and work ahead of the curve, translating the strategy of the organization into a new quality. The key participants in the process of training specialists whose competencies can be in demand in the digital world are identified. Recommendations for improving the effectiveness of the learning process are given. An important advantage of the approach to teaching the course "Internet entrepreneurship" is the integration of Lean Startup methodology and Design thinking. The authors developed cases and practical recommendations for the creation of digital services based on the Human-Centered Design.**

***Key words:* IT education, personnel training, technological entrepreneurship, digital markets, design thinking.**

I. INTRODUCTION

Digitalization of the economy and society will revolutionize the work of organizations and will force managers to reconsider the strategy of their functioning, because new ways of management, approaches to recruitment will be needed. In addition, it is important for managers themselves to strengthen and improve their competence not only in their subject areas but also in the management of work and people in order to achieve success. The speakers of the digital conferences usually highlight the importance of team skills development and personal development of a person throughout his life. Soft skills are necessary for large companies, who find it most difficult to adapt the control mechanisms and technological platform to the requirements of the digital future. The labor market today is in demand for ambition and entrepreneurship.

II. THE DIGITAL MENTALITY OF THE HEAD AND STAFF OF THE STATE ORGANIZATION

The company introduced the position of Director of digital technology, because they understand the importance of management of digital processes and digital transformation activity. At present, the issues of updating the structure of the organization in state corporations and organizations are being discussed. There are recommendations to introduce in large state corporations and structures responsible for the implementation of the program "Digital economy of the Russian Federation", the position of Director or Top Manager for digital transformation, Chief Digital Officer (CDO) or Head of digital transformation. The duties of the Director of digital technologies include the strategic development of the organization in the context of its transition to digital platforms, the transfer of its activities in the Internet space.

The Ministry of economic development of the Russian Federation, together with Rosatom, Rostech, digital economy, the center for strategic development, the Agency for strategic initiatives (ASI), Russian Railways, MTS and other representatives of scientific communities and business structures, developed and published on the website the documents "Plan for the formation of state corporations and companies with state participation of structural units for digital transformation and the organization of their work for 2018–2019".

The tasks of the top Manager for digital transformation will include the introduction of "end-to-end" technologies provided by the program "Digital economy", including the collection, processing and analysis of big data (Big Data), neural networks, blockchain, artificial intelligence (AI), robotics and sensor technology, simulation system of the full life cycle of the product, the Internet of Things (IoT), augmented and virtual reality (AR/VR), quantum and cloud computing. Based on this list of tasks, we conclude that the head of the new formation, which will determine the digital strategy of the company and be responsible for its transition to a new quality, must have extensive knowledge in the technical field and business, be well versed in information and communication technologies (ICT), be able to think globally, be competent in finances, marketing, own sales tools, have a wide managerial horizons and experience, have the gift of foresight (proactivity) to be able to work ahead of the curve, and enterprise, "to be able to manage change and do it quickly".

As a digital leader, a Manager must have the skills of team building and management, interpersonal communication, be able to engage and motivate his subordinates to new tasks and self-improvement. A new type of manager should develop a culture of innovation in the organization. Important questions that the manager should ask himself today: "Do you have specialists for the introduction of new technologies? Are your employees ready for changes?" Employees must be ready for

changes, it understand the essence of changes so that the digital transformation of their activities can be successfully implemented.

The lack of digital competence and the absence of positive expectations for the introduction of ICT led to the refusal to make a decision on the beginning of digital changes in the organization. The low digital literacy or lack of ICT experience of staff in organizations, as well as their possible neutral or negative attitude to changes related to digital technologies, are called "fear factors" of global corporate digital transformations [1].

But how to improve the effectiveness of training of future leaders of the digital world?

III. THE ECOSYSTEM OF THE COURSE "INTERNET-ENTREPRENEURSHIP"

In 2015, the Financial University under the Government of the Russian Federation signed an agreement with the Foundation for the development of Internet initiatives (IIDF) on the inclusion of the course "Internet entrepreneurship" in the educational programs of the bachelor's and master's degree in the field of "Business Informatics". Over the past two years, IIDF has been actively promoting the idea of introducing this course in more than 100 universities in Russia, as well as in Kazakhstan and Belarus.

However, we are faced with a problem that is likely to be relevant for many economic universities. Ideas offered by students in the classroom on the course "Internet entrepreneurship", and do not reach the stage of implementation in a real product, are not implemented in the Internet service or mobile application. Our students, who have good training in business engineering, management, are not confident enough in application development. As a result, for 2 years of teaching students the basics of Internet entrepreneurship, we have accumulated a sufficient base of unrealized useful ideas. This would not have happened if our university or other universities had a developed ecosystem of training support for Internet entrepreneurs.

Participants in such ecosystem (Fig. 1) should be not only teachers and students of one university, but also some support of the university management. The course itself and the new forms that we use in training and for which the IIDF highly appreciated us, giving twice the title of laureate, could not be realized without the participation of the Vice-Rector for educational and methodical work, the head of the educational program. Now we are promoting the idea of Internet entrepreneurship, organizing various sessions of design research, hackathons, business games within the framework of scientific conferences and events at the Financial University, as well as holding elective classes, including courses for teachers and administrative staff. Now the course "Internet entrepreneurship" has become mandatory in the direction of 38.04.05 "Business Informatics", master's program "Strategic Management of Information Technologies in Business".

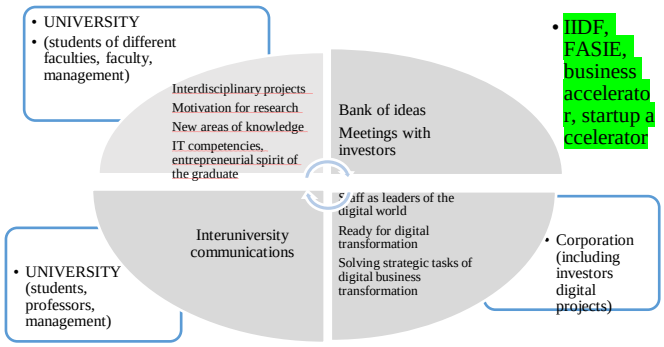


Fig. 1. The ecosystem of the course "Internet Entrepreneurship"

We also believe that it is important to develop inter-university relations when students of different faculties or areas of training of different universities can be involved in the project, which will ensure the feasibility of the principle of an interdisciplinary approach to the project. Communication and organizational activities, as well as the interaction of the authors of the best projects with investors, can take on the IIDF, as the main inspirer of the course.

An important role is given to IT companies, which will receive not only interesting design solutions, but also their potential employees prepared to work in the digital markets. It is not for nothing, for example, the introduction of Design Thinking principles among key partners is one of the key areas of the SAP technology leader. For the development of new customer-centered services in the context of digitalization of society and the economy, it is important for advanced companies to be able to determine the deep needs of customers, so they use innovative techniques and tools. Design Thinking is now implemented in IBM, General Electric, Procter & Gamble, Philips Electronics, Airbnb. In Russian, Design sessions are held in Sberbank, Raiffeisen bank. In the ecosystem of training of future leaders of the digital world also become important participants and entrant of it areas of universities. Career guidance work should be carried out with students to the profession came aimed at the result of people.

IV. THE APPROACH OF DESIGN THINKING IN THE ECOSYSTEM OF THE COURSE "INTERNET ENTREPRENEURSHIP"

Note that an important advantage of our approach to the teaching of the course "Internet Business" is to integrate the Lean Startup methodology [2–4] with the stages of Design Thinking, according to "d.school", as follows: Empathize, Define (the problem), Ideate (Generation and Voice), Prototype, and Test. This allowed us to solve the problem arising in practice at many universities, testing this course in the educational process, with the formation of new ideas, customer research at the stage of customer development, testing hypotheses with the participation of representatives of the target audience. Graduates of schools who are passionate about information technology and decide to go to university to continue IT education are also becoming important participants in the ecosystem of training future leaders of the digital world. Therefore, we believe that universities should actively cooperate with schools so that schoolchildren come to the profession, who are focused on results. This approach should now become the basis for the production of goods and services. The importance of studying the value of any innovation for people before its implementation in the product is stated in the works of Blank [2], Kotler [5] and other scientists. Creating the soul of a brand requires knowledge

not only of the laws of marketing and business, but also psychology, anthropology, cultural studies. And empathy, which is the basis of most Design Thinking techniques is, first of all, understanding the experience of the consumer, his feelings and feelings, which is aimed at further filling the developed innovative product understandable to the consumer emotions [6–8].

How to implement Design Thinking in the processes of the company that has chosen the strategy of digital transformation? The answer to this question depends on the purpose of the company. This can be a solution to the problems of studying changes in customer requests or problem areas in customer service, finding ways to transform outdated approaches in doing business, training staff tools to make decisions in unusual situations. Session duration is possible from 8 hours to 3 days. Tasks step-by-step can be performed in a team from 5 to 7 people.

During the two years of teaching the course "Internet Entrepreneurship," we have tested the inclusion of a variety of Design Thinking tools in the Lean Startup methodology. Work on projects of technological entrepreneurship was both team and individual. The team came up with the idea, tested hypotheses in customer research. The analysis of startup metrics and calculation of economic indicators of the project solution, description of business processes of startup activity and definition of ways of holding clients were carried out individually. The results were made not only in the presentation of the idea but also in the writing of research works in which it was necessary to show the features of its creation, to give proposals for promotion. The course "Internet Entrepreneurship" was held in the Financial University within the faculty classes (36 hours, every two weeks), training practice (two weeks, from 9.00 to 13.00, four times a week). However, in our deep conviction, for the success of learning the skills to create a startup is suitable for a rich program, as it was in the organization of training practice [8; 9] when nothing distracted from the project, when each day began with a warm-up thinking, learning new techniques of design thinking, adjustments and mentoring by teachers, teamwork and research.

We have seen the main purpose of the course is to give students an idea about the features of creating a startup, promotion, in practice, tried out all the steps from the development of innovative ideas to testing and promotion of the project solution feel like young entrepreneurs. For two years we have been actively implementing various forms of research, communication and design activities. As part of the faculty stage of the International student scientific conference, it has become a tradition for our students to participate in the hackathon, in which they study the principles of design thinking. Student teams develop business ideas, create prototypes of products or services that solve the problems of the city, society, etc.

Design thinking and the Lean Startup methodology, if included in the training course for IT specialists, develop students' qualities such as enterprise, teamwork, trust in success, not fear failure, which will help them solve difficult non-standard situations in their future profession in the face of uncertainty. Our approbation in the educational process of new forms of education has given good results in final qualifying works. In their master's theses, our graduates included the results of design research and offered innovative digital development strategies for companies such as Russian Post, Sberbank. We believe that these are important steps to the transformation of the educational process. The skills of creative thinking, teamwork and personal development, which we are developing with such new methods of training our graduates will make them competitive in the labor market, will allow them to competently shape their strategies for professional growth.

V. CONCLUSION

The digital transformation of organizations is becoming a reality, the management of which requires the formation of new skills from the head. It is he who has a special role in the processes of digital transformation as the main carrier of possible technological changes, the competent implementation of which should provide quality, positive changes in the organization.

The trend of customer orientation and reactive, sometimes unexpected changes in their needs, tastes, preferences make business change, look for new business moments, which, in Gartner's terminology, means a constant search for short-term opportunities used dynamically: "…a business moment can come from nowhere, and yet they are becoming more and more" (Gartner Symposium/Itxpo 2013). Until recently, the human resource was perceived as the main participant of business processes in enterprises, as an integral part of the organizational capital of the company. However, automation and the introduction of industrial technologies 4.0 replace human programs, robots, bots, drones, etc. This does not mean that in the new era of human functions in the process will be minimized.

VI. REFERENCES

[1] Алтухова Н.Ф. Условия реализации цифровой трансформации в организации // Экономика. Налоги. Право. 2018. Т. 2. С. 70–75. [Altuhova N.F. Conditions for the implementation of digital transformation in the organization" // Economics, Taxes & Law. 2018. Vol. 2. P. 70–75]. DOI: 10.26794/1999-849X-2018-11-2-70-74.

[2] Blank S. The Four Steps to the Epiphany: Successful Strategies for Products that Win. Sussex: Quad/Graphics, 2014.

[3] Blank S., Dorf B. The Startup Owner's Manual: The Step-By-Step Guide for Building a Great Company. Pescadero K&S Ranch Press, 2012.

[4] Ries E. The Lean Start-up: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses. NY: Crown Business, 2013.

[5] Kotler P., Kartajaya H., Setiawan I. Marketing 3.0: From Products to Customers to the Human Spirit. NY: John Wiley & Sons, Inc., 2010.

[6] Kelley T., Kelley D. Creative Confidence Unleashing the Creative Potential Within Us All. New York: Barnes & Noble, 2013.

[7] Liedtka J., Ogilvie T. Designing for Growth: A Design Thinking Toolkit for Managers. New York: Columbia University Press, 2011.

[8] Vasilieva E. Design thinking: a little bit about the approach and a lot about the tools of creative thinking, learning client requests and creating ideas: monograph. Moscow: RU-SCIENCE, 2018. (In Russ.)

[9] Алтухова Н.Ф., Васильева Е.В., Громова А.А. Опыт примене- ния техники дизайн-мышления в курсе «Интернет-пред-принимательство» // CEUR Workshop Proceedings. 2016. Vol. 1761. P. 219–225 [Altuchova N., Vasilieva E., Gromova A. Teaching experience of Design Thinking in the course of "Internet-business" // A. Conference: CEUR Workshop Proceedings. 2016. Vol. 1761. P. 219–225].

Russian Digital Economy: State and Development Prospects

Natalya L. Udaltsova
Department of Management
Financial University under the Government of the Russian Federation
Moscow, Russian Federation
udaltsova.nl@yandex.ru

***Abstract*. This scientific article covers the conceptual framework of digital economy and the influence thereof on the development of world economic processes, and investigates the efficiency of digitalization process development in Russia and in some other countries. According to estimates of consulting companies, the digitalization of the Russian economy has a potential effect of 4.1–8.9 trillion roubles for GDP, which will amount to approximately 19–34% of the total increase of GDP by 2025. Russia has all necessary prerequisites for the achievement of the digital potential. In order to achieve this, first of all, it is necessary to determine the technological focus and to concentrate investment opportunities on the goals set, to develop the innovation culture in all possible ways on the model of the country's digital companies. Another important task that requires a timely solution will be the education area and training of specialists. It is obvious that the digitalization will lead to emergence of new professions and obsolescence of some existing professions.**

***Key words:* digital economy, Industry 4.0, information and telecommunication technologies.**

I. INTRODUCTION

The modern development of information and telecommunication sector and its accompanying technologies significantly change the relations in the society. Therefore, one may observe the formation and further development of the information society which is called the "digital economy". It is obvious that digital technologies have become an integral part of social and economic life of the society and a key direction for the development of the state policy at the modern stage. As a matter of fact, the economy digitalization process is a process of change of the technological scheme and production process. To all intents, we talk about another industrial revolution "Industry 4.0" presenting a serious challenge for the international community, economies and companies all over the world that should be ready for it. The concept of the fourth industrial revolution (Industry 4.0) was stated in 2011 at the Hannover Exhibition. The participants of the event defined it as the implementation of "cyber-physical systems" into industrial processes. While the concept "Industry 3.0" was aimed mostly at the automatization of certain processes and machines, Industry 4.0 provides for the end-to-end digitalization of all physical assets of the enterprise and the integration thereof into the unified ecosystem together with the partners that participate in the added value chain. It is assumed that the computer-aided manufacturing will be controlled by intelligence systems, going beyond the boundaries of one enterprise, on a real-time basis at continuous interaction with external environment, with a prospect of unification into a global industrial network of things and services. Industry 4.0 is related to a number of key technologies, such as Industrial Internet of Things, additive manufacturing, Big Data analysis, cyber security and augmented reality [1]. It is expected that these particular technologies will blur the lines between physical, digital and biological spheres.

It might seem that a new paradigm means new opportunities for everybody, but as the history shows, the change of the formation may become a beginning of the end for those enterprises that were not ready for it [2]. In the nearest future, the level of competitiveness will be defined by the ability of the companies to use digital technologies in their manufacturing process, and it is obvious that the transfer to digital economy will lead to unrecognizability of many industries and sectors of economy.

It is evident that the digital economy breaks the usual models of growth for the companies, industries and national economies in general. Understanding this, well-known companies try to implement digital solutions in their operation. Digital platforms constitute a valuable asset and a core of the "new" economy. Modern digital platforms are developed by such famous IT companies as Amazon, eBay, Facebook, iTunes, Airbnb, Tencent, Vkontakte, Yandex, Avito, OZON. Digital platforms are used as the main channels and tools for communicating with clients, making financial transactions, and constitute a kind of a mechanism for creation of innovation business models. Even today we can see how the digital technologies developed by these companies changed the appearance of certain industries and sectors of economy, for example private transportation sector (taxi services), tourism industry and printing industry.

II. WORLD TRENDS OF DEVELOPMENT OF DIGITAL ECONOMY

Each country has its own achievements and approaches to creation of digital economy. So far, we cannot observe an absolute leader in the development of all aspects of digital economy, however, Germany is considered to be the leader of industrial technologies and the home of the term "Industry 4.0". Nearly 10% of population in Germany are engaged in high technology industries. The purposeful state policy and financing of promising digital projects play a key role in the formation of the digital economy in Germany. The main tasks of the state policy in Germany include

ensuring institutional conditions for the promotion of research activities and developing the fundamental education.

The features of digital economy in Japan and South Korea are formed around famous large corporations such as Toyota, SoftBank, Samsung, LG, etc., where innovation products appear and digital solutions are developed. The startup ecosystem here is less developed, although certain projects in the internet commerce appear (Rakuten) and internet messengers are created (Line, Kakao).

The USA achieved certain success in development of digital technologies in the country. The volume of digital economy in the structure of GDP is fairly high and amounts to 10.9% (Table 1), which is related, first of all, to extensive investment of private and public sectors to digital technologies. The USA managed to produce innovations on a large scale in many areas of economy.

The economy of China is developing towards the digitalization particularly rapidly and efficiently. Although the economy is still referred to developing countries, its share of digital economy in GDP is comparable to indicators of the USA. The experience of China in commercialization of promising digital projects targeted at the external market, for example, Alibaba, Huawei, proved to be very successful. The countywide digitalization in China resulted in the extensive use by the population of online retail services, digital banking and various online possibilities of ecosystem. The physical retail trading was replaced by electronic commerce resulting in the development of financial online services [4].

When studying the experience of different countries in the development of digital economy and its accompanying technologies, it is necessary to mention that every country has its own way of development but at the same time they all have two common features as favourable conditions for the commercialization of innovation projects and significant volumes of investment capital contributed to the area of information and communication technologies (ICT). Moreover, at the first stages of digitalization, new entrepreneurial opportunities are opened for the companies due to emergence of new market niches and trends (online medical services, online retail services, etc.). Therefore, Russia has a chance of making a technological leap forward in many spheres. For example, we can observe something of the kind in the banking sector where Russian financial institutions managed to escape some remains that were an obstacle to banking systems of developed countries in ’80s–’90s, and proceed to the development of digital functional right away.

It is expected that the transition to the digital economy of Russia will be one of the main factors of GDP growth. According to forecasts of experts of McKinsey consulting company [3], the potential effect from the digitalization of the Russian economy for GDP by 2025 is estimated to be 4.1–8.9 trillion roubles, which will amount to 19–34% of the total increase of GDP. The priority directions of Russian digitalization include digital ecosystems, digital platforms, 3D-printing, advanced analytics of big data arrays, internet of things, etc. The experience of China and the USA forms certain expectations for the development of digital economy, since these particular countries managed to increase the performance and qualification of labour force, reduce prices by means of competition growth and facilitate the access to obtaining the information.

According to expert investigations, in Russia the share of digital economy in the GDP structure amounts to 3.9% (Table 1), which is 2–3 times less as compared to the countries being the leaders [3]. On the one hand, such percentage proves the insufficient level of economy digitalization as compared to the world standards, on the other hand, a positive growth dynamics of this indicator may be observed over a period of 5 years. For example, Russia’s GDP increased by 7% from 2011 to 2015, and the volume of digital economy for the same period increased by 59% [3]. It is fair to say that the digital economy grows 8.5 times faster than other sectors of economy. The world leader is United Kingdom with the share of digital economy in GDP amounting to 12.4% following the results of 2016 [4].

TABLE I. Contribution of digital economy into Russia’s GDP and its components as compared to other countries, in % of GDP [5]

	USA	China	Germany, Italy, France, Sweden, United	India	Brazil	Russia
Expenses of households in digital area	5.3	4.8	3.7	3.2	2.7	2.6
Investments of companies into digitalization	5	1.8	3.9	2.7	3.6	2.2
Government expenses on digitalization	1.3	0.4	1.0	0.6	0.8	0.5
Export of ICT	1.4	5.8	2.5	2.9	0.1	0.5
Import of ICT	–2.1	–2.7	–2.9	–6.1	–1.0	-1.8
Total amount of digital economy	10.9	10	8.2	6.3	6.2	3.9

Expense of Russian households in digital area amount to 2.6% of GDP, that is approximately 2 times less than in the USA and China. They consist of the costs related to the use of electronic commerce possibilities and purchase of digital devices, as well as other digital expenses, such as payment for the

Internet, payment for the services made by means of mobile applications, etc. The electronic commerce that was growing at fast pace during the past years accounts for the half digital expenses of Russian households: the annual growth of electronic commerce in Russia during the last several years amounted to 20%, while certain categories, for example, online sales in tourism area grew by 30%.

As regards to the indicator "investments of private companies" in digitalization, so far Russia falls behind the countries being the leaders. In Russia, this indicator amounts to 2.2% of GDP, in the USA — 5%, in countries of Western Europe — 3.9%, in Brazil — 3.6% (see Table 1). The share of export of digital goods and services is an important indicator for the evaluation of digital economy efficiency. The share of export in general reflects the competitiveness of national products on the world market and the degree of integration of the economy into world economic processes. As it follows from schedule 1, the percent of export of digital solutions amounts to 0.5% in the structure of GDP. In China and India the same indicator reaches the level of 5.8 and 2.9 respectively. It is obvious that a relatively low volume of export does not correspond to the digital economy standards.

However, despite of fairly low indicators as compared to other countries, Russia still has certain results in the area of digitalization. Famous large enterprises Rosatom, Rostec, Sberbank became pioneers in the implementation of digital solutions. Many of them not only deploy the technologies, but are also engaged in the development of their own solutions. The experience of Gazprom Neft, that for several years has been reforming its production facilities according to a new model, may serve as an interesting example of putting Industry 4.0 concept in operation. Vnesheconombank defined blockchain technologies as the priority direction of development. In 2017 blockchain projects were launched with the Pension Fund, Rospatent. Digital Economy ANO was established, which has among its members such companies as Mail.ru, Yandex, Rambler, Megafon, VimpelCom, MTS, etc. Also, a number of private companies that were established virtually from scratch achieved some success: for example, Tinkoff Bank online bank that does not have physical branches and that has become the largest independent bank in the world just in 10 years; Kaspersky company specialising in production of digital solutions in security area; digital portals and ecosystems of Yandex and Mail.ru services; Avito electronic announcement platform.

Due to a globally emerging trend, the government policy targeted at digitalization significantly improved the quality of public services and made them more accessible for citizens and for the business. The level of digitalization of public services in Russia is slightly lower the indicators of countries being the leaders. The level of government expenses on information and communication technologies in the structure of GDP is comparable to China, India, but is two times less than in the USA and Western Europe. For the last several years, Russia created government digital platforms that are comparable with similar systems in the leading countries, and sometimes even surpassing their services: obtaining a passport or driver's license, vehicle registration, school enrollment, payment of taxes and fines and many others. Following the results of 2016, nearly half of all public and municipal services may be obtained in electronic form, and more than 50% of people applying for those services prefer to obtain them in electronic form.

III. PROBLEMS OF RUSSIAN DIGITALISATION

The development of digital economy in Russia was mentioned for the first time in the message of the President to the Federal Council in December 2016. The President of the Russian Federation pointed out the need to form a new web-economy in order to increase the industry efficiency by means of information technologies. Within the policy outlined by the President the Government prepared a "Digital Economy" program that was approved in July 2017. It is obvious that the approved program not only showed the intent of the government to be in line with the global trends, but also set a goal to integrate into the world reality using the technologies of Industry 4.0 project. The main task of digital transformation of Russian economy is to improve the quality of life of the population by making high quality products produced using modern digital technologies.

For the implementation of the Digital Economy program, first of all, it is necessary to define the technological focus and only then to concentrate the investment capital for the goals set. The correct goal setting is a guarantee of success for any action. All the more, the promising directions are already globally formulated and clear. Probably, in order to reach the maximum effect the accent should be made on the development of those technologies where Russia has strong competencies, for example, IT industry. Investments in those technologies where the competitive advantage already exists will bring better results as compared to the development from scratch.

Secondly, most probably, the Russian economy will have to meet two challenges at the same time: modernization and technological leap. Both challenges imply two scenarios: catching-up and outstripping. In case of catching-up scenario, investments in downstream industries, infrastructure facilities and export shall be made. The catching-up scenario does not imply technological advances. The outstripping development centers around innovation projects related to blockchain technologies, quantum technologies, scaling, market place and other B2B.

The difficulty of digital transformation is connected not only with the choice of proper technologies and priorities. The problem is rooted in the insufficient maturity of digital culture in the society and at enterprises. According to the information of Digital IQ-2017 research, the investments in breakthrough technologies have a

fundamental meaning; however, eventually the success and efficiency of digital transformation will depend on a broad range of factors related to people, and not on sensors and analytics tools. In keeping with this approach, the task will be to form the digital culture and the conditions for the personal interest in the digital transformation on the part of companies' personnel and management. The digital economy is the environment that will be forming for several years and will need the engagement of competencies of many skilled specialists.

However, the growing economy digitalization sets to the industries and the government new tasks that they have not encountered before. As and when digital automation and robot systems are introduced with the resulting increase of labor productivity, the physical services will be replaced by digital services. As consequence, a number of specialties and professions may be at the risk of extinction. According to the research of McKinsey institute, by 3036 up to 50% of jobs will be automated. First of all, this will apply to the jobs that require average qualifications, because it is easier to automate the types of jobs that are connected with performance of predictable repeated physical operations and jobs related to information gathering and analysis. In keeping with this approach, the digitalization of the Russian economy requires from employees and employers to adjust to new conditions.

At the same time digital technologies and platforms may have a significant positive influence on the labor market: they will facilitate the personnel hiring process, reduce the period of searching for a job, increase the productivity of employees, improve the situation with engagement of personnel in the economy with the help of remote jobs and provide access to the quality education. The task of the government in this case will include the timely coordination of actions related to training and employment of released personnel.

IV. CONCLUSION

The creation and improvement of new technologies takes place so quickly that it seems to be impossible to use old technologies within the digitalization of global economic processes. This is such a moment when it is expedient to enter to maximum extent possible into the general information and technological stream of global digital changes and try to use them efficiently in the national economic environment.

So far, Russia has been able to hold the leading positions in terms of hydrocarbons and grain products export. The fact that Russia missed the third industrial revolution and as a result became a consumer of innovation solutions and not a producer of high technologies worsens the situation. There exist unresolved problems related to commercialization of breakthrough inventions and infrastructure that is highly needed for innovations. We may ascertain only specific innovation results of particular companies. Corporate structures that have a margin of financial safety and potential are often not interested in innovation solutions. As a result, we face a global problem: the economy is able to generate knowledge and make discoveries but is not able to commercialize the obtained results and incorporate into the global market so far.

The spread of digital technologies at "natural rates" in Russia will not bring the intended results, and the absence of well-directed efforts will only worsen falling behind the countries that will manage to catch a wave of digital revolution. Against the background of the world events, it is necessary to develop a smoothly running mechanism for the achievement of the goals set that will be based on close integration and efficient cooperation of the science, government and business.

V. REFERENCES

[1] Тарасов И.В. Технологии Индустрии 4.0: влияние на повышение производительности промышленных компаний // Стретегические решения и риск-менеджмент. 2018. №2. С. 62–69. [Tarasov V. Technologies of 4.0 Industry: Influence on the Improvement of Performance of Industrial Companies // Strategical Solutions and Risk-Management. 2018. N 2. P. 62–69].

[2] Трачук А.В., Линдер Н.В. (2017а). Инновации и производительность российских промышленных компаний // Инновации. №4 (222). С. 53–65. [Trachuk A.V., Linder N.V. Innovations and Performance of Russian Industrial Companies // Innovations. 2017. N 4(222). P. 53–65].

[3] Цифровая Россия: новая реальность. 2017. [Digital Russia: New Reality. 2017]. URL: https://www.mckinsey.com/~/media/McKinsey/Locations/Europe%20and%20Middle%20East/Russia/Our%20Insights/Digital%20Russia/Digital-Russia-report.ashx.

[4] Цифровая экономика:глобальные тренды и практика российского бизнеса. [Report "Digital Economy: Global Trends and Experience of Russian Business"]. URL: https://imi.hse.ru/data/2017/10/06/1159517769/!Цифровая%20экономика%20-%20глобальные%20тренды%20и%20практика%20российского%20бизнеса.pdf.

Research on the Application of Neuromarketing Technologies on Financial Markets

Svetlana V. Karpova[1], Ilya V. Rozhkov[2]
Department of Management
Financial University under the Government of the Russian Federation
[1] SVKarpova@fa.ru, [2]irojkov@fa.ru

Abstract. **The change in marketing paradigms, which occurs under the influence of the rapid development of information technologies and new forms of communication, requires the search for more objective methods of consumer research. Neuroeconomics and neuromarketing as its applied branch allow to study psychophysical states of a person in the process of consumer choice. The article attempts to identify the most promising areas of application of neuromarketing technologies. The review of the most significant directions in Russia and abroad is carried out. Possible prospective directions for the development of neuromarketing research on the financial market are identified: machine learning, social scoring, artificial neural networks.**

Key words: **neuroeconomics, neuromarketing, machine learning, social scoring, artificial intelligence.**

I. PROBLEM STATEMENT

For more than a century of studying marketing as a methodology and a philosophy of market activity went through several stages of development. The change in the marketing paradigms, characteristic of certain periods of socio-economic relations development throughout the 20th century is a consequence of the fundamental changes that took place in society during this period. Each subsequent stage moved the vector in the marketing activities of organizations from the intra-organizational environment to consumers, their needs and necessities, as well as the sustainable development of society as a whole [3].

Within the framework of modern marketing and marketing concepts 3.0 and 4.0, consumers are perceived as "full-fledged human beings endowed with thoughts, emotions, and soul" [4]. P. Kotler notes that a person in the modern world not only meets his needs and desires, but seeks to make our modern globalized world better or, at least, not to harm the society and nature. Thus, consumers not only get satisfaction from their functional, but also from emotional needs. The basis for such self-realization of consumers is formed by the possibilities of modern information technologies and digital-marketing as social networks, mobile devices for transferring information, new forms of collective interaction of people on the basis of information services.

With the growth of technical capabilities of communication and interaction of people, the question of improving existing and searching for new technologies for consumer research is being updated. Such technologies can provide deep and qualitative study of behavioral patterns, mental insights, psychological and emotional state of a person in order to improve the efficiency of market activities. Particularly important is the study of factors of increasing consumers' involvement in interaction with the company, communication with its brand and products, participation in the joint development of new products (co-marketing) and their promotion on the market (viral marketing). The most promising methodology for this kind of research is neuromarketing, it is a section of a more general scientific area of neuroeconomics, which object of research is consumer behavior.

At present, neuroeconomics is increasingly attracting the attention of Russian and foreign scientists and is becoming a long-term priority of scientific research. According to the estimates of the NeuroNet branch association, engaged in the development and promotion of neurotechnologies in Russia, the market for neuroscientific research will reach $100 billion by 2020 and 1.8 trillion by 2035 [6]. Prospects for the use of neurotechnologies are also taken into account at the state level in the long-term planning of the development of the digital economy in the Russian Federation. So, according to the program "Digital Economy of the Russian Federation" No. 1632-r dated July 28, 2017, the main end-to-end digital technology is "neurotechnology and artificial intelligence" [1]. In this connection, it is of particular interest to analyze the current state and prospects for the development of Russian and international neuromarketing research markets. Its results can serve as the basis for the formation of foresight of applied and fundamental research, not only to solve the problems of promoting the products of domestic producers, but also to increase its competitiveness by taking into account the needs of consumers more deeply.

II. BACKGROUND REVIEW

In general, the concept of neuromarketing was first developed by Harvard University psychologists in the 1990s. The technology is based on the model according to which the bulk (more than 90%) of human thought activity, including emotions, occurs in the subconscious area, that is, below the levels of controlled awareness.

Neuroeconomics or neurobiology of decision-making is an interdisciplinary field at the intersection of psychology, biology

and behavioral economics. Currently, neurotechnologies are used in various applied fields: medicine (neurophysiology and neurotechnology), transport and logistics, industrial sphere, information technologies, etc. The most significant studies in the field of neuromarketing are presented in Table 1.

Table 1

Analysis of the most significant foreign studies in the field of neuromarketing

Country	Scientists	Research Center	Directions of research
United Kingdom	Roman Borisyuk	University of Plymouth	Biological realistic models of the formation of neuro-architectures and functions
Netherlands	Ale Schmits	University of Erasmus of Rotterdam	Neuromarketing
USA	Gregory Burns	Emory University	Consumer behavior
USA	Brian Knutson, Alexander Geneva	Stanford University	Neuroprediction
France	Hilke Plassmann	INSEAD	Neuroprediction
USA	Vernon Smith	Center for Neuroeconomics Studies	Experimental Economics
Germany	Berndt Weber	Bonn University	Visual communication
USA, Israel	Daniel Kahneman	Princeton University	Behavioral finance
USA	Jerry Saltman	Harvard Business School	Neuromarketing

Without the use of neurotechnologies, the development of the following technologies couldn't be completed:

- artificial intelligence;
- virtual and augmented reality;
- search engines (neurocompliments);
- communications;
- assessment of physical, emotional and intellectual states of a person;
- Neuro-linguistic programming.

In the framework of neuroeconomical researches, scientists distinguish neuromarketing, which is a complex of methods for studying the behavior of consumers, the impact on it, as well as emotional and behavioral reactions to this impact, using the latest developments in the fields of marketing, cognitive psychology and neurophysiology [2].

Unlike traditional marketing methods, neuromarketing does not use subjective data to obtain information about consumer preferences. Neuromarketing allows you to influence the thoughts of a person before he has realized them and developed his position. The data obtained as a result of the analysis are most often used for:

- testing concepts and ideas for new products;
- testing of adverts and communications;
- testing of package and brand identification;
- studying the image and reputation of the organization;
- testing of sites' usability and interfaces;
- study of merchandising, work of trade and office space, efficiency of POS.

In Russia, a lot of scientific centers are engaged in neuroscience issues (Table 2).

Table 2

Russian research centers in the field of neuroscientific research

Research Center	Directions of research
I. Kant Baltic Federal University, Institute of Living Systems	Genetically engineered technologies, nanomaterials, applied engineering technologies, etc.
Lomonosov Moscow State University, laboratory of neurophysiology and neurocomputer interfaces of the Faculty of Biology	Development of human-machine interface technologies
Kurchatov Institute SRC	Neurotechnology of intelligent information processing in real time
Pavlov Institute of Physiology	Processing of image and sound signals in the context of neurotechnology
University ITMO, branch union "NeuroNet"	Technologies related to the integration of the human brain and computer

III. THEORETICAL GROUNDING AND METHODOLOGY

The most famous neuromarketing technology was developed by Harvard professor Jerry Saltman, who patented it as the Zaltman metaphor extraction method (ZMET).

The model is tested by the largest brands: Coca-Cola, Toyota, PNC, Proctor & Gambel, General Motors, etc. Its main goal is to bring to the realized level the hidden thoughts of customers that continuously occur in their brains. They are able to evoke an emotional positive response and activate hidden requests that stimulate purchase.

The instrumental approach in neuromarketing is associated with the use of special equipment:

- Mobile Eye-tracking (mobile tracker);
- Polygraph / lie detector (for instrumental psychophysiological studies for synchronous recording of parameters of breathing, cardiovascular activity, electrical resistance of the skin, etc.);
- Stationary Eye-tracker (eye tracking technology);
- eye-tracker + audit of a site (the Webvisor in Yandex.Metric, the technology of recording and analizing of actions of visitors on a site);
- Face-reading (mobile laboratory testing the effectiveness of advertising);
- EEG and neurocomputer interface based on electroencephalography (systems for registration electroencephalograms helping to determine the brain's response to certain stimuli);
- MRI scanners (functional magnetic resonance tomography).

IV. EMPIRICAL ANALYSIS

Currently, among the practical areas of application of neuromarketing researches in the financial market, we can distinguish the following:

- artificial intelligence and machine learning;
- social scoring;
- artificial neural networks.

Artificial intelligence and machine learning

The prospects for neuro forecasting in the stock market are quite promising. Thus, the Cindicator service is a decentralized, community-driven technology infrastructure that enables the use of the values of hybrid artificial intelligence on the securities market, investment and crypto-currency markets. This synergy allows to make more precised decisions in the face of increasing uncertainty.

Cindicator is a mobile application in which users answer questions about how much companies' stocks or other financial instruments will cost in the nearest future. After that, the answers are aggregated and processed with the help of artificial intelligence, which forms a single forecast. The robot commits transactions based on the forecast on the stock exchange, and part of the profit from the trade is distributed among users. As a result, all participants become not only users, but also contributors, and their internal mental work becomes capital that can be effectively used. Over the past six months, the value of the company's portfolio has increased by 168%, and the users of the mobile application have received dividends worth more than 24 thousands of dollars.

The Russian service of predictive analytics for the financial and cryptocurrencies market Cindicator has collected $15 million in the cryptocurrency based on the initial placement of tokens (ICO).

To promote this startup a number of marketing tools was used. For example, most Russians learned about the project from the video on YouTube and from articles in magazines. Also, various conferences, meetings and other public events were held.

Marketing tools based on social scoring

One of the directions for implementing neurotechnologies based on Big Data is the integration of information about consumers from different sources. This technology is called “social scoring”, which implies the collection of information about a client of a financial institution judging by his actions in social networks and data located in his accounts. Currently, there are several services using the technology of social scoring on the financial market. For example, the Double Data service is a solution for retail banks to improve the efficiency on all stages of working with individuals: from attracting customers and evaluating them to the collection of overdue debts due to the use of Big Data technologies, advanced machine learning techniques and data mining. Double Data collects information from open sources and includes them in existing models for evaluating potential customers, which, according to the product developers, allows to increase the efficiency of models by 10–15% and reduces the number of non-performing loans. This service has already received more than 200 million rubles. investments from the funds of LETA Capital and Simile Venture Partners and is preparing to enter the markets of Europe, Asia and South America.

Social networks have also joined the development of social scoring technologies. Thus, Facebook has officially patented the program for sharing information with banks. From now on, credit institutions will be able to assess its creditworthiness with the help of a potential borrower profile in a social network. A new system (and in fact scoring on the Facebook social network) will tell banks about the borrower's friends in the social network. The more the borrower has friends with a good credit history, the higher the chances of getting a loan.

Marketing tools based on artificial neural networks

Neural network technologies are a set of information technologies based on the use of artificial neural networks. Artificial neural networks are software or hardware implemented systems built on the principle of organization and functioning of their biological analogue (the human nervous system).

FscoreLab is a tool developed by the Russian entrepreneur from St. Petersburg, Nikita Valeev, in 2017. It allows us to assess the creditworthiness of potential customers using neural network technologies on their photos and other parameters taken from social networks VKontakte, Odnoklassniki, Facebook. The service uses a neural network, trained on 600 thousand real cases of loan payment taken from the bases of the bureau of credit histories. The program examined photographs of people with arrears on loans and without it and independently identified the features intrinsic to the persons with debts. The service is planned to be monetized with the help of turnkey scoring projects for the largest banks (box solutions according to banks).

Currently, only 3.8 million rubles have been invested in the project. However, the startup has implemented five successful projects, which allowed to recoup investments and ten projects are in progress now. FscoreLab is present in Russia, Greece and the USA.

The cases examined indicate the prospects for the development of Russian marketing services based on neurotechnologies that will allow to design more sophisticated financial products that better meet the needs of the target audience and strengthen the Russian financial system.

V. REFERENCES

[1] Распоряжение Правительства РФ от 28.07.2017 N 1632-р «Об утверждении программы "Цифровая экономика Российской Федерации"» [The order of the Government of the Russian Federation of July 28, 2017 No. 1632-r “The Digital Economy of the Russian Federation”]. URL: http://www.consultant.ru/document/cons_doc_LAW_221756/.

[2] Маркетинговые коммуникации: учебник и практикум для прикладного бакалавриата / С Азарова и др.; Под ред О.Н. Жильцовой М.: Юрайт, 2014. –[Marketing communications: textbook and practical work for applied bachelor's degree / S. Azarova et al.; ed O. N. Zhiltsova. Moscow: Yurajt, 2014].

[3] Инновационный маркетинг / Под ред. Е.В. Карповой. М.: Юрайт, 2017. [Innovative marketing: a textbook for undergraduate and graduate programs / Ed. S.V. Karpova. Moscow: Yurajt, 2017.
[4] Kotler P., Kartajaya H., Setiawan I. Marketing 3.0: From Products to Customers to the Human Spirit. Hoboken: John Wiley and Sons, 2011.
[5] Рожков И.В. Информационные системы и технологии в маркетинге. М.: Русайнс, 2014. [Rozhkov I.V. Information systems and technologies in marketing. Moscow: Rusains, 2014].
[6] NeuroNet. URL: http://rusneuro.net.

Features of Joint Value Creation in the Context of Digitalization

Elena A. Davydenko
Department of Business Economics
Saint-Petersburg State University
Saint-Petersburg, Russian Federation
E-mail: elenadavidenko1@yandex.ru

Abstract. **In the context of digitalization, the role of the consumer in value creation is also changing. While companies have previously minimised the involvement of consumers in the creation of products, it is now common for consumers to be involved in the value creation process. At the same time, the process of involvement can be both direct and indirect, for example, through the interaction of consumers with each other and brands in online communities, which gives valuable information to the business for building and adjusting advertising strategies and offering value to customers.**
Key words: **digitalization, digital economy, value, brand, consumer.**

The proliferation of sensors and the Internet of Things (IoT) has enabled the creation of digital models based on the vast amount of data collected. These digital models have led to the emergence of new services and business models by connecting multiple devices to high-performance networks, which erases territorial and time constraints to date. The role of the consumer is also changing in the new economic environment. Actual information technologies have led to the fact that the consumer has become active partner for the manufacturer taking part in the process of designing the product or service. The transformation of existing business processes in the digital economy leads to an active dialogue between the producer and the consumer, which leads to a new understanding of joint value creation in the digital economy [1]. Active interaction of consumers in the network leads to the emergence of so-called collective intelligence, which generates such a phenomenon as the effects of social networks, or eWOM (electronic word of mouth). It is worth noting that the effects of social networks have an impact in almost all spheres of activities, where purchases of goods and services are made. At the same time, such effects have the strongest impact in business, based on the impressions of consumers and their personal experience of interaction with the company.

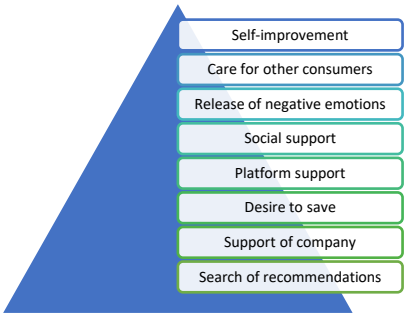
The process of value creation in its traditional sense involves a clear distinction between the roles of the producer and the consumer. The market mechanism allows to move the created value from the producer to the consumer. The value creation process postulates the need to create value together with the consumer, rather than embedding it in the product of the company [2].

It is interesting to see the difference between the terms *co-creation of value* and *co-production.* So, P. Chathoth and his co-authors R.J. Harrington, E.S.W. Chan and others [3] show the differences between these terms in a number of criteria. As shown in Table 1, in the framework of joint value creation, the role of the consumer is not limited to advice and opinion expressed by the company, the consumer becomes an equal participant in the value creation process, a participant involved in this process.

Table 1. Comparative analysis of the terms "joint value creation" and "joint production" [3]

Criteria	Joint value creation	Joint production
The role of the consumer	Consumer-partner and source of information	Considered as a resource
Consumer participation	Repeated interactions at all stages of creation and value consumption. There is a joint creation of products and services with the company.	In the final stages of the value chain. Adapts to existing offerings of the company.
Key players in value creation	Clients, managers and employees	Managers and employees
Focus	Customer focus and experience of his participation in the creation of value	The focus on the company and production
Innovations	Initiated in cooperation with the consumer	Initiated by the company
Nature of customer interaction	Interaction is based on an open dialogue with the client.	Representatives of the company listen to the opinion of customers.

To date, R. J. Brodie, L. D. Hollebeek and several others [4] consider the term *involving the consumer* from the point of view of its participation in virtual communities. Today, digital platforms allow consumers to take an increasingly active part in the activities of virtual communities. And here there are natural questions-why the consumer to participate in such communities? What does he get from this participation? K. P. Gwinner, G. Walsh and others [5] have identified motivations that encourage consumers to participate in these communities through research.



Motivations for consumers to participate in online communities [5]

In turn, the interaction of consumers in online communities with each other and brands is extremely beneficial for companies. What are the advantages of such interactions for business?

- The company can learn more about its active customers. For example, if the demographics of customers in online communities are different from what the ad campaign was originally set up for, the company will need to reconfigure its ad campaign and may even affect how new products are developed or offered.
- The company can understand what content is interesting for its customers. Numerous repost images of the goods indicate that the company has attracted the attention of consumers. The more attention a company gets, the more brand awareness spreads.
- Increased site traffic. Interaction in social networks is a way to organically increase the flow of traffic. The increase in activity in social networks, whether “likes”, reposts or comments, gives an increase in visits to the business site. Well, every new visit is another chance for conversion.
- Strengthening ties with the customer base. Social media advertising enables the company to reach out to consumers in a way that doesn't make them feel pressured. You can use social media to answer consumer questions, solve problems, and even make consumers feel special by making them special offers.
- Using online communities, you can perfect your brand identity and gain the trust of your customers. Trust and branding are two strongly intertwined factors. Online presence helps the brand to communicate its values and views that may well resonate with customers. Many consumers argue that “shared values” are the main reason for their brand loyalty.

Social media encourages customers to provide marketing materials for the company. Creative consumers use their pages on social networks and online communities to publish information about the products they use, the places they visit, the companies that mean something to them. Thus, companies receive virtually free advertising. Social media allows you to communicate with customers directly-and allow them to communicate with the company.

At the same time, today many companies focus exclusively on getting subscribers and clicks. Instead, a more effective strategy is to coordinate a PPC (pay per click) campaign with social media marketing. Cross-channel events allow you to fine-tune your PPC campaign with all the information that social networks provide. Companies can adapt their PPC strategy to deliver optimized ads directly to customers who are likely to convert.

In turn, paid clicks can actually lead to the creation of custom content, which ultimately will lead to an organic increase in site rankings in search engines. If a customer clicks on an ad, and then places a link to the site or talks about a product on his page, he raises the company's rating in the search results.For example, on Facebook, the custom audiences feature can only be used to promote your ad to users who show the most organic engagement. In Google AdWords, you can use the "customer email Targeting" feature for the same purpose.

Thus, the following conclusions can be drawn:

- The transformation of existing business processes in the digital economy leads to an active dialogue between the producer and the consumer, which leads to a new understanding of joint value creation in the digital economy.
- Digital platforms allow consumers to become increasingly involved in virtual communities.
- Among the motives that encourage consumers to participate in virtual communities, we can call such motives as finding recommendations, social support, care for other consumers, support for the company, self-improvement.
- Interaction of consumers in online communities with each other and with the brand turns out to be beneficial for business, allowing organizations to optimize their advertising campaigns for the actual needs and interests of the target audience.

REFERENCES

[1] Трачук А.В., Линдер Н.В. Трансформация бизнес-моделей электронного бизнеса в условиях нестабильной внешней среды // Эффективное антикризисное управление. 2015. № 2. С. 58–71. [Trachuk A. V., Linder N. V. Transformation of business models of electronic business in conditions of unstable external environment // Effective Crisis Management. 2015. № 2. P. 58-71].

[2] Lusch R.F., Vargo S.L. Service-dominant logic: reactions, reflections and refinements // Journal of Marketing Theory. 2006. № 6 (3). P. 281–288.

[3] Chathoth P., Altinay L., Harrington R.J. et al. Co-production versus co-creation: A process based continuum in the hotel service context // International Journal of Hospitality Management. 2013. № 32. P. 11–20.

[4] Brodie, R.J., Hollebeek L.D., Ilic A. et al. Consumer engagement in a virtual brand community: An exploratory analysis // Journal of Business Research. 2011. № 66(1). P. 105–114.

[5] Hennig-Thurau T., Gwinner K.P., Walsh G. et al. Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? // Journal of Interactive Marketing. 2004. № 18 (1). P. 38–52.

Artificial Intelligence as a Challenge for Industries, Economy and Society

Irina V. Alyoshina
Department of Marketing
State University of Management
Moscow, Russia
ialyoshina@gmail.com

Abstract. **Artificial Intelligence changes the global landscape of technological development and wellbeing and challenges national economies and societies with hardly predictable threats.**

Key words: **artificial intelligence, globalization, e-platform, neural network, algorithm, digitization.**

I. ARTIFICIAL INTELLIGENCE IN INDUSTRIES' AND MARKETS' TRANSFORMATION

In a broad sense, artificial Intelligence (AI) can be regarded as " software technologies that make a computer or robot perform equal to or better than normal human computational ability in accuracy, capacity, and speed" [1]. AI is viewed as a software that implements the tasks of finding solutions using an algorithm with the ability of self-learning. It is used for speech recognition (translation of texts, Google translator), image and face recognition (Facebook), for consulting (business and financial advice, medical services, legal advice), automated provision of public services, recruiting, credit risk assessment, psychological targeting of Internet-advertising, including political, in social media services (Cambridge Analytica). AI is used in industrial manufacturing, in transportation, in agriculture, rapidly expanding and deepening the scope of application. According to leading technology companies, the world market of AI technologies will reach $16 billion in five years. Annually, the industry grows by 60%. Now artificial neural networks are mainly used for face recognition, automatic translation from a foreign language or online consultation [2]. Association for the Advancement of Artificial Intelligence (AAAI) gives more abstract and broad definition of AI, "the scientific understanding of the mechanisms underlying thought and intelligent behavior and their embodiment in machines" [3].

One of the first resonant evidence of the success of the AI was discovered in 1997, when the IBM computer Deep Blue for the first time beat the world chess champion Garry Kasparov. In 2017 AlphaGo, created by Google, defeated the strongest player on the planet in Go. The machine analyzes a large number of factors (scenarios, variants and results) inaccessible to the human brain, which works in a simplified scheme.

Systems of AI, Internet of things, robotization and block-chain are voiced by visionaries of IT business among the main directions of technological development of a number of industries. These technologies are interrelated and interdependent. They decentralize decision-making in society and the economy, using the principle of distributed trust in the network of participants instead of the administrative vertical approach. A large number of devices, people and organizations are connected by intensive information, material, financial and other resources' flows and processes. New technologies rely on artificial neural networks, networks of computational elements ("neurons"), or systems of algorithms for complex decision makings according to the rules "if – then", improving in the process of functioning ("learning systems").

The leading IT and Internet companies of the world are engaged in research, development and implementation of artificial intelligence technologies (Google, Amazon, Facebook, Apple, Alibaba, Microsoft, SAP, IBM). All these companies use business models of the digital cloud platform, providing a platform for interaction of suppliers and consumers of goods and services, ideas. The platform acts as an intermediary for transactions of a large number of participants in various markets, charging them a fee similar to the tax. Competitiveness of the platform depends on the number and quality of participants customers and partners involved. The Internet infrastructure is international, so the global scale of operations is provided by the extraterritoriality of the Internet-infrastructure of digital platforms.

Companies working with end-users (Amazon, Apple, Google) put the development of AI and voice assistants in the center of their business models. These technologies allow to monitor and analyze consumer behavior in real time, bringing business closer to the process of consumer decision making and so to management of buying decisions.

The large audit four (PwC, EY, KPMG and Deloitte) are actively increasing their investments in the research and use of AI technologies. These technologies shorten the processing time of customer orders, facilitate Big Data analysis and identify unusual transactions, outside interference and fraudulent schemes. Image recognition, automation of accounting and of other routine tasks optimize the work of auditors.

The large audit four –(PwC, EY, KPMG and Deloitte) are actively increasing their investments in the research and use of AI technologies. These technologies shorten the processing time of customer applications, facilitate the

analysis of large data and identify unusual transactions, outside interference and fraudulent schemes. Image recognition, automation of accounting and of other routine tasks optimize the work of auditors.

The ability to work with neural networks in the summer of 2017 was declared by G. Gref as a non-alternative criterion for hiring lawyers in Sberbank. In 2016, 450 lawyers who drafted statement of claim for the court were reduced there, and their work was transferred to neural networks. In 2017, it was planned by Sberbank to cut 3,000 lawyers, replacing the work of consultants with AI systems.

Automation and robotization of social sectors, including education and health, are able to improve the quality and efficiency of their functioning through the individualization of services, adapting them to the needs of citizens.

AI optimizes traffic flow. The systems of AI, with which cars are equipped, track the changing of the traffic situation and the actions of the driver, adapting the settings of the car. Self driving cars are already being designed without the steering wheel and pedals, threatening to leave the driver's profession in the past.

The spending of industrial manufacturing sectors on Internet of Things technologies is more than three times higher than those in the consumer sector of the economy [4]. The industrial Internet platforms of SAP, GE, Siemens transform the industries, assembling value chains on a global scale. They allow factories to link information about supplies and stocks, technical processes, repairs, safety, to analyze and optimize the operation of the chain. The growth of competition causes the integration of the platforms of individual players ("islands") into "continents", which means the potential reduction of today's dozen of the largest industrial platforms to 2–3 leading players. It is obvious that the owner of the platform not only owns almost unlimited information about the functioning of the largest companies, industries and regions, but can influence this operation.

AI is also used in the news industry, the media. Robot Heliograf wrote more than 800-t articles for the Wall Street Journal, and the world's largest news agency AP automated analysis of company reporting for the preparation of financial news.

AI technologies are supposed to change the structure of the global economy in the same way as computer, telecommunications and internet have already succeed to change the world. Developing countries and regions loose the advantages of cheap labor force being replaced by automation and robotization [5]. For example, Sewbot is much more productive than human worker and takes jobs from clothing industry in Bangladesh, Vietnam, India and Pakistan. AI helps the machine to perform human tasks, using robot camera as eyes like for self-driving car. Robotized factories do not need much workers in low paid jobs and may be moved in developed countries in order to be more close to target markets and consumers in US and Western Europe.

The global scale of automated manufacturing decreases cost per unit of product produced and makes the product available for increasing number of people around the world. AI makes mass customization possible and gives competitive advantage of personalized products and services for millions and billions of customers, as examples of Facebook, Amazon and Google show. As a result, Internet companies and social media services overtake mass media industry by the number of users and audience attention, and because of it, by advertising budgets. Traditional industries are pressed to be more digitized and also to use AI technologies to survive.

Education is one of the industries involved in global technological competition. Russian traditional offline universities are in quite uneasy situation with resources necessary for automation and digitization. They could hardly compete with the leading American and Western European universities in the world market of education basically because of financial resources available and environmental infrastructure.

Unless you are unable to digitize your business processes, you are unable to use AI technologies and real advantages or individualized learning path for each student. So, we need to think about digitization, because the problem of segmentation of student audience and the problem of individualization of learning in an age of growing uncertainty are more serious than it can be imagined.

II. ARTIFICIAL INTELLIGENCE IN ECONOMY'S AND SOCIETY'S TRANSFORMATION

Upcoming transformation of the production of goods, services, ideas based on AI systems is quite impressive and designated as the next, the fourth industrial revolution; Industry 4.0, Digital Economy. It is about radical changing the technological structure of the economy and society and, accordingly, the next serious social transformation. The development of AI assumes an increase in the automation of the work of many specialists, beginning with routine components with gradual displacement even more complex types of labor. Replacement of people with robots and intellectual decision-making systems raises public issues of employment reduction for a number of professions and rising unemployment, technological and economic inequality, unconditional income as compensation to the population for job cuts, and also robot taxation.

Leadership in the research and application of AI is a significant factor in the restructuring of industries and economies of countries and regions, as well as the geopolitical positioning of national economies. For example, China plans to become the leading world-class center in the field of AI [6] by 2030. In the summer of 2017, the Chinese government approved a three-stage plan to develop and implement AI technologies.

There is no accidental interest in the field of AI of Russian IT companies and government agencies looking for opportunities to optimize business processes. Rapid growth of AI technologies presupposes the need for specialists from different regions and social groups to participate in the study and discussion of this sphere in order to reduce the risks of social institutions' development deformation. The solution of this task on the scale of the Russian Federation supposes involvement of broad audiences of potential researchers, developers and users in the theoretical and practical issues of AI subjects. The speed and scale of AI development depend on the scale and nature of the audience's involvement, which means the need to provide broadband Internet access throughout the country. The accessibility of the Internet for the population seems to be a key task of building an Internet infrastructure as a necessary condition for the formation of a national digital economy. Like most information technologies, the development of AI requires the

development of an extended market and an environment for IT services consumption, the price and availability of which is determined by the effect of the scale of their usage. The more people in Internet, the more extended market created and that is why the scale of operations too.

AI is not only a technology, but also a methodology (system of methods) and moreover, ideology (a system of ideas, values, views of a philosophical nature). The implementation of AI systems involves answering the questions "what is good and what is bad" for a person, business and society, i.e. resolving legal, ethical problems but not only technological ones. It is quite problematic to consider AI in isolation from the digital economy as an environment for the functioning of AI, since the AI assumes the collection and transmission of large amounts of data from a large number of sources, as well as intensive information processing and use for a multitude of subjects for real time decision-making. This requires an extended and developed information technology infrastructure, specialists of appropriate qualifications, and the rules of their work. The digital economy assumes the transformation of all spheres of society's life. AI is a tool for resolving the problems of a wide range of interrelated and interdependent spheres and systems, technical, biological, socio-economic, socio-cultural, political and legal.

In terms of the degree of radical change, AI can be attributed to breakthrough, radical innovations that significantly change the life of society. And simultaneously AI is a hype, i.e. an acutely fashionable term for some circles of the society. According to IT specialists in Russia, disruptive decisions (which radically change the business model) do not take root in the traditional business environment, and therefore customers are advised not to immediately implement hype approaches, but to test them "from the side" of the main business for checking their efficiency and effectiveness.

The development of AI systems that process large data on a large number of citizens and generate the most important resource of the modern economy (information) carries the risk of a digital dictatorship to society if the information on citizens is concentrated in the hands of elites. Today, large data on citizens are collected and used by corporations for commercial purposes, including for marketing and advertising their products to people as consumers. However, manipulation of such data in the interests of elites can also change political processes, transforming democracy (distributed processing of data) into a digital dictatorship. The concentration of information about citizens in the hands of a limited circle of people (elites with the resources to collect and process large data) means the possibility of hacking (deciphering algorithms) of not just information systems (computers, banking information systems), but also human organisms with the aim of modifying them in specified purposes. Such modification carries risks for the biological and social evolution of society and that risks are far more serious than destruction of a number of professions.

III. ARTIFICIAL INTELLIGENCE: THE AGENDA FOR HIGHER EDUCATION

The use of artificial intelligence systems requires competences in both IT and in the field of application. This means changing the education system both at the content and at technological level and also at the organizational level. Thus, for example, the training of specialists in the IT medicine program is already conducted in a networked form, combining a medical college, an IT university and a technical university. Leading specialists in the field of medicine and medical education say that a modern medical university should train not just a doctor, but a specialist who can work with medical robots. Automation and robotization of medicine, the introduction of IT technologies (beginning from storing images on the server instead of printing) can improve the quality of health care and its effectiveness.

For most Russian universities, the movement in the development trend of AI systems should begin at least with the development of the IT infrastructure, the growth in the availability of Internet access for training and for all other spaces of university campuses. Without such a base, it is impossible to develop research and training programs, courses and materials on digital topics, including AI. At the same time, even the most IT-equipped universities as small islands of digital infrastructure can do little in the country, where the infrastructure necessary for Industry 4.0 is practically absent outside of large cities. Obviously, it is more difficult to cover the entire territory of the Russian Federation with the Internet than the territory of any Western European country. However, without solving the problem of building a digital infrastructure, the Russian economy will not be able to become digital for a long time, and the scale of the application of AI systems will not allow the technology to really function. So, the creation of an adequate digital infrastructure as a "net of roads" for digital information flows is the first step in the development of artificial intelligence in Russia and the construction of the digital economy as a whole.

IV. REFERENCES

[1] AI definition. Retrived from http://www.businessdictionary.com/definition/artificial-intelligence.html.

[2] Колганов Г. Язык нейронов. Чего ждать от искусственного интеллекта // Коммерсант. 2017. 1 августа. [Kolganov G. The language of Neurons. What to expect from artificial intelligence. // Kommersant. 2017. August 1] URL: https://www.kommersant.ru/doc/3372769?from=doc_vrez

[3] Association for the Advancement of Artificial Intelligence. URL: https://aitopics.org/search.

[4] Alessi C. GE, Siemens Vie to Reinvent Manufacturing by Harnessing the Cloud // WSJ. URL: https://www.wsj.com/articles/ge-siemens-vie-to-reinvent- manufacturing-by-harnessing-the-cloud-1488722402.

[5] Emont J. The Robots Are Coming for Garment Workers. That's Good for the U.S., Bad for Poor Countries // The Wall Street Journal. 2018, Feb. 16. URL: https://www.wsj.com/articles/the-robots-are-coming-for-garment-workers-thats-good-for-the-u-s-bad-for-poor-countries-1518797631?mod=e2fb.

[6] Курмукова А. Бесстрастный анализ. Может ли Россия стать лидером в сфере искусственного интеллекта // Коммерсант/ 2017. 26 сент. [Kurmukova A. Passionless analysis. Can Russia become a leader in the field of artificial intelligence? //Kommersant, 26.09.2017. September 26]. URL: https://www.kommersant.ru/doc/3421743.

System of Evaluation of Intellectual Capital and Intangible Assets Created on the Basis of Intellectual Property

Niyaz M. Abdikeev
Institute of industrial policy and institutional development
Financial University under the Government of the Russian Federation
Moscow, Russia
Nabdikeev@fa.ru

Abstract. **The article discusses the evaluation of the value of intellectual capital and intangible assets of the organization. Intellectual capital includes all the knowledge of the organization. These are different types of knowledge, innovations, know-how, knowledge possessed by the employees of the organization, the knowledge base of the organization, the electronic network and the database based on it, this cooperation with customers.All of that allows you to respond to changes in the market situation faster than competitors. That is, intellectual capital is a set of explicit and implicit knowledge. The main thing is that this is the knowledge that the organization can turn into profit. Intellectual capital is the result of the interaction of people with each other, people and information resources, as well as people and elements of physical capital in the production process. It is the result of past investments, but it is forward-looking. The assessment of intellectual capital is formed on the basis of the results of its future use.**

Keywords: **intellectual capital; intangible assets; intellectual property; valuation; innovation; new technologies.**

I. INTRODUCTION

Intangible assets have a certain value for the company, although they have a different form of presentation than physical objects. Some of the intangible assets were recognized many years ago. These are patents, copyrights and trademarks.

Intangible assets include:

- market assets (service marks, product marks, trademarks, corporate name, business cooperation, licensing and franchise agreements, etc.);
- intellectual assets (patents, software, design rights, copyrights, trade secrets, know-how, trademarks);
- human assets (education, qualifications, skills and abilities of employees);
- infrastructure assets (corporate culture, management concepts and management processes, relationships, etc.).

The approaches to determining the cost of work on the creation of new technology, such as regulatory, adaptive, program-target and program, are considered. The method of calculating the value of intangible assets created on the basis of intellectual property and aimed at improving the technologies operating in the market, and the creation of fundamentally new products or fundamentally new technology. This method shows that the result of innovative projects can be used in determining the price of the contract for research and development. In this case, its main component is the expected result of the sale of the created intangible assets in the market. Thus, the method is a tool to stimulate performers to create market-promising results of research and development work.

II. INTELLECTUAL CAPITAL AND INTANGIBLE ASSETS

Intellectual capital (IC) includes all the knowledge of the organization. These can be ideas, different types of knowledge, innovations, know-how, knowledge owned by employees of the organization, the knowledge base of the organization, the electronic network and the database on its basis, this cooperation with customers. All of that allows you to respond to changes in the market situation faster than competitors. That is, intellectual capital is a set of explicit and implicit knowledge. The main thing is that this is the knowledge that the organization can turn into profit.

Intellectual capital is the result of the interaction of people with each other, people and information resources, as well as people and elements of physical capital in the production process. It is the result of past investments, but it is forward-looking. The assessment of IC is formed on the basis of the results of its future use.

The intellectual capital, in L. Edvinson's interpretation [1], includes two components. The first is human capital, i.e. the totality of knowledge, practical skills and creative abilities. It is not the property of the company and is described in detail in the economic literature. The second is structural capital such as brands, documented business processes and everything that provides the performance of staff. The structural capital is the property of the company. It consists of consumer capital (relationship capital) and organizational capital.

Consumer capital is a set of established relationships with the entire environment of the company, primarily with customers. The organizational capital is the organizational knowledge, systematized and formalized competence of the company.

Their interpretation of IC suggested K. Sveiby [2]. He identifies the IC and intangible assets (IA). However, there are

other interpretations of these concepts, since the question arises of their identity Are all intangible assets knowledge and able to benefit from their use? The balance sheet measures and sums the value of individual resources of the company, and the market considers and evaluates the company as a whole. Moreover, there is already a synergetic effect when the sum of the properties of the constituent parts is less than the properties of the system as a whole. Therefore, the cost of the whole must exceed the sum of the costs of its elements.

IA have a value for the company, although they have a different form of presentation than physical objects. Some of the IA was recognized many years ago. These are patents, copyrights and trademarks. However, only a few companies have tried to give a monetary evaluation of IA.

IA include:

- market assets (service marks, product marks, corporate name, business cooperation, licensing and franchise agreements, etc.);
- intellectual assets (patents, software, design rights, trade secrets, know-how, trademarks);
- human assets (education, qualifications, skills and abilities of employees);
- infrastructure assets (corporate culture, management concepts and management processes, relationships, etc.).

The scheme of the IC by K. Sveiby contains three elements: the competence of the employees (human capital), internal structure, external structure (customer capital, and relationship capital).

The competence of employees is the ability of people in the organization to act in different situations. It includes teaching, skill, education, experience, social skills, etc. The competence of employees creates internal and external structures.

The internal structure consists of goals, objectives, models, technologies, information systems, etc., which are owned by the organization. It includes business strategy and concept, process knowledge, patents and all information stored and implemented in computer systems. It provides the staff of the organization with the ongoing support they need to carry out their work.

The external structure consists of links with customers and suppliers, competitors, brands, reputation. An external structure is a relationship with suppliers, stakeholders, partners, and customers.

III. METHODOLOGY OF THE RESEARCH

Approaches to valuation of intellectual capital

If IC is put in a row with other factors of production, there is a problem of its estimation. Today, traditional methods of economic valuation and measurement, based on the principles of accounting, have ceased to be adequate to the conditions. For example, traditional accounting practices treat a trademark as an intangible asset that, by analogy with a tangible asset, loses its value in the course of its use and transfers its value in parts to a manufactured product. In this regard, IA are accounted for according to the same rules as tangible assets, depreciation rates are applied to them and their write-off is made. At the same time, the trademark or brand in the process of their exploitation not only does not lose its value, but, on the contrary, it is often increased. Moreover, many elements of intangible assets are not reflected in the balance sheets, including connections with consumers, staff qualifications, knowledge base, etc.

Such researchers of knowledge-based economic intellectual capital as L. Edvinsson [1], K. Sveiby [2], M. Malone [3], T. Stewart [4], E. Broking [5] and others have developed a number of methods to assess intellectual capital.

In particular, K. Sveiby identifies 26 methods of valuation and measurement of intellectual capital, grouped into four categories:

- methods of direct measurement of intellectual capital (Direct Intellectual Capital methods, DIC). It is based on the identification of intellectual capital and the valuation of the monetary values of individual assets or components of intellectual capital, after which the integrated assessment of the intellectual capital of the company;
- methods of market capitalization (Market Capitalization Methods, MCM), when calculated the difference between the market capitalization of the company and the equity of its shareholders. The obtained value is regarded as the value of its intellectual capital or intangible assets;
- methods of return on assets (Return on Assets Methods, ROA). The ratio of the average income of the company before taxes for a certain period to the material assets of the company (ROA of the company) is compared with the same indicator for the industry as a whole. To calculate the average additional income from intellectual capital, the resulting difference is multiplied by the tangible assets of the company. Then, by directly capitalizing or discounting the resulting cash flow, you can determine the value of the company's intellectual capital;
- scoring methods (Scorecard Methods, SC). Various components of IA or IC are identified; indicators and indices are generated and recorded in the form of scoring or as graphs. The use of SC-methods does not imply a monetary valuation of intellectual capital. These methods are similar to methods of diagnostic information system.

All known methods of IC assessment are divided into four categories listed. It should be noted the relative proximity of DIC and SC methods, as well as MCM and ROA methods. In the first two cases, the movement comes from the identification of individual components of IC, and in the remaining it is from the integral effect.

None of the methods meet all objectives of the evaluation, so the choice methodically capital depends on the task and situation, for example, to study the company is usually most suitable methods of scoring and methods of direct measurement of intellectual capital.

In order to extract value from intellectual capital, organizations need to manage information flows between different types of capital that make up intellectual capital. The purpose of management of the intellectual capital of the

enterprise is to achieve the maximum result from its use while minimizing the cost of its development.

The problems of efficient use of IC and IA are usually left out of management activities, as in domestic enterprises are mainly solved only the problem of accounting for such assets. This is due, in particular, to the lack of development or absence of the methods to assess the effectiveness of the IC management system, as well as the effectiveness of its creation and use.

Evaluation of the effectiveness of the management system intellectual capital is the calculation of a set of indicators. This is because the process of intellectual capital management involves the implementation of a large number of functions, the quality of which is difficult to assess. In addition, such indicators can be used as benchmarks in the development of the company's strategy for the creation, acquisition and use of intellectual capital. Methodological problems of accounting, analysis, management and evaluation intellectual capital and intangible assets are also considered in the works of domestic researchers [6–12.

Approaches to the valuation of intangible assets

Intangible assets can bedivided into three main groups: intellectual property (IP), organizational expenses, and goodwill.

IP includes industrial property, rights to secrets (know-how) and copyright objects.

Industrial property includes inventions, utility models, industrial designs, trademarks and breeding achievements.

The objects of copyright include: computer programs, databases, topology of integrated circuits, scientific public lectures.

The organizational costs include fees for lawyers for drafting constituent documents, services for the registration of the company and others. Costs for re-registration of constituent documents, the production of new seals and stamps intangible assets are not included.

Goodwill appears in an enterprise only if it has made a purchase of another enterprise. Business reputation can be defined as the difference between the purchase price of the acquired enterprise and the value of the balance sheet of all its assets and liabilities.

The market value standard is not always applicable to IA, as market value is the most probable price at which an object can be disposed of in an open market in a competitive environment. But for some intangible assets, it is difficult to talk about the most probable price of the transaction, because the market of such objects is underdeveloped.

For the valuation of IA, the same approaches that are traditionally used for the evaluation of objects are mainly applicable: cost, revenue and comparative. However, with respect to IA, the use of these approaches has its own specificity in intellectual capital, since intangible assets are rather unusual object for evaluation.

Key in the evaluation of IA is often a profitable approach. It is based on the income that the owner of the valuation object (IA) expects to receive. When assessing IA under the income approach, a method is sought to allocate income related to the assessed IA in order to determine the value of IA using the conventional methods of income approach (capitalization or discounting) [13].

There are a number of approaches to determining the cost of work on their creation in the practical intellectual capital of program-target, planning and management of the development of new technologies. The most frequently used and well proven both in our country and abroad are normative, adaptive, program-target and program approaches [14].

The normative approach provides for the establishment of a certain standard to determine the cost of work in the creation of new technologies as a percentage of the total cost of research and development (R&D). In fact, all leading foreign countries adhere to a normative approach in one form or another. The normative approach is quite simple to implement, but it has low accuracy and can only be used to estimate the amount of work on the development of new technologies.

The adaptive approach is based on determining the share of financing for the creation of new technologies in a specific promising direction from the total cost of R&D, depending on the goals set out in the development strategy of the enterprise.

Compared with the normative approach, the adaptive approach will determine the necessary costs depending on the development strategy of the enterprise. However, its practical application requires more raw data.

The program-target approach provides for full-scale research and development overall set of directions contained in the list of promising technologies for the enterprise.

A programmatic approach to determining the value of new technology creation and development involves extending the financing of new technologies over a period of time, taking into account the overall dynamism of the intellectual capital of R&D financing.

However, the considered approaches do not take into account the emerging market price and technical characteristics of the intellectual capital of the new technical intellectual capital and technology and the relevant trends.

In this regard, a method of calculating the value of IA created on the basis of the IP and the price of the contract for their development is proposed.

IV. THE RESULTS OF RESEARCH

Calculation of the value of intangible assets created based on the IP and aimed at improving the technologies operating in the market

To calculate the value of intangible assets on the basis of the IP used to improve the technical level of the product or process, the following procedure is used, which is a modification of the capitalization method of parametric indices [15].

The following stages are implemented within the calculation procedure:

1. On the basis of the set of technical characteristics, the intellectual capital of the best analogues and the original technology, it is formed a set of indicators *i* in the form of the vector $P_{i,k0}$, the state of the basic production technology of the

considered sample technical intellectual capital, which is "ideal" at the moment.

2. The characteristics of analogs and original technology are compared with the characteristics of the "ideal" sample. Parametric index is calculated:

$$SP_{i,k} = P_{i,k} / P_{i,k0}.$$

At the same time, each technology of analogues of the original technology is assigned the corresponding numbers k = 1, 2, ..., n; the original technology has the number (n + 1); i is technical characteristics of the intellectual capital of the technical intellectual capital (no more than five main characteristics of intellectual capital are usually used in the practical intellectual capital of marketing and benchmarking), n is the number of analogues in the market. Reference model, characterized by the best performance among the analogues on the market, is assigned the number k_0.

3. The sum of the parametric indices of analog technologies and original technology is calculated:

$$\sum_i SP_i, w_i,$$

where w_i is weight coefficient P_i of the technical indicator, adopted in practice, the intellectual capital of marketing.

4. The economic effect of the original technology in relation to each k of n analogues is determined:

$$C_{ef.k} = \left(\frac{\sum_i SP_{i,n+1} w_i}{\sum_i SP_{i,k} w_i} - 1\right),$$

where $k = 1, 2, \ldots, n$.

The economic effect caused by differences in the technical level of products is manifested in the change in the price of analogs: the price increases if the coefficient of technical improvement meets the condition:

$$C_{ef.k} > 0 \text{ for } k = 1, 2, \ldots, n.$$

5. Intellectual capital forms the number of $k_{\min}$, analog having the lowest price of $P_{k\min}$ in the market.

Then, the additional revenue from the sale of the product on the market will be

$$\boldsymbol{AR = N \times P_{kmin} \times \left(\frac{\sum_i SP_{i,n+1} w_i}{\sum_i SP_{i,kmin} w_i} - 1\right)},$$

where N is physical volume of production.

6. Additional net profit obtained by improving the technological level of products is equal to

$$ANP = N \times P_{k\min} \times \left(\frac{\sum_i SP_{i,n+1} w_i}{\sum_i SP_{i,kmin} w_i} - 1\right) \times r \times (1 - f),$$

where r is the profitability ratio, taken, for example, for high-tech industries is equal to 0.2; f is tax rate.

7. In the distribution of additional profit between the licensee and the licensor, the ratio of the cost of exploration and technological development to the total cost, including additional development of technologies and production organization, is used (1:4:16) [15], which is determined by the licensee's cost factor (k_{cl}).

If the licensee conducts both exploratory research and technology development, then

$$k_{cl} = \frac{1+4}{1+4+16} = \frac{5}{21} \approx 0{,}24.$$

When conducting only exploratory research,

$$k_{cl} = \frac{1}{21} \approx 0{,}05.$$

In the development of technology $k_{cl} = \frac{4}{21} \approx 0{,}20$.

8. The additional net profit from the implementation of the m-th sample of equipment, improved on the basis of the IP at the time tm, is described by the equation

$$ANP\ (t_m) = N_{m.\max} \times P_{k.\min} \times \left(\frac{\sum_i SP_{i.n+1} w_i}{\sum_i SP_{i.kmin} w_i} - 1\right) \times r_m \times (1 - f_m) \times k_{clm} \times \left(1 - \frac{4}{T_m^2}\left(t_m - \frac{T_m}{2}\right)^2\right),$$

where $N_{m.\max}$ is maximum physical volume of production sold on the market; $P_{k.\min}$ is the price of the m-th sample of technical intellectual capital and the market;

T_m is total life time of production of a particular kind;

$$t_m = t_{m0} + \Delta t_m,$$

where t_{m0} is the beginning of the commercialization of the m-th sample of technical intellectual capital in relation to the time of market entry pioneer model of the intellectual capital; Δt_m is the time period of commercialization of the original technology, measured in years.

9. Commercialization of products is carried out within Δt_m years, therefore, taking into account the discount determined by the formula

$$E_m = R_{infm} + R_{risk.m} + R_{inf.m} \times R_{risk.m},$$

where $R_{inf.m}$ is the inflation index at the time of commercialization of the m-th sample; $R_{risk.m}$ is the risk index from the implementation of the m-th sample of the improved technical and intellectual capital, the ratio to calculate the additional net profit from the implementation of the m-th sample of the technical and intellectual capital received by the licensee (researcher, developed by the intellectual capital of the technology and a specialist in engineering and production organization), depending on the coefficient of k_{clm}, is described by the

$$ANP\ (F_m) = \sum_{n}^{F_m} N_{m.\max} \times P._{k.\min} \times \left(\frac{\sum_i SP_{i,n+1} w_i}{\sum_i SP_{i,kmin} w_i} - 1\right) \times \times$$

$$r_m \times (1 - f_m) \times k_{clm} \times \left(1 - \frac{4}{T_m^2}\left(t_m - \frac{T_m}{2}\right)^2\right) \times (1 + E_m)^{-n},$$

where $t_m = t_{m0} + \Delta t_m$; $F_m = T_m - t_{0m}$.

10. To determine the cost of creating a technology to improve the *m*-sample of technological product, we take into account the additional profit obtained through its commercialization in the ratio

$$E_m = ANP\ (F_m) \times (1 + r_m)^{-1},$$

where r_m is profitability of scientific and technical activities.

Thus, the licensor at the conclusion of the contract for the development of technology and/or exploratory research of engineering works and the organization of production should provide for the payment to the licensee of an amount equal to the *ANP* (F_m), in which the E_m is the cost of carrying out the relevant work.

Calculation of the value of intangible assets created based on intellectual property and aimed at creating a fundamentally new product or a fundamentally new technology

In the case of production of fundamentally new products or new technologies the value of IA is determined by the ratio

$$S_m = V_m \times N_E \times k_{clm},$$

where V_m is the expected revenue from the *m*-sample technology intellectual capital sold on the market, created on the basis of IP; N_E is the knowledge-intensive sector of the market, within which it is expected to sell fundamentally new products.

The value of the cost of work on the creation of technology in the *m*-sample at the profitability of r_m is determined by the formula

$$E_m = V_m \times N_E \times k_{clm} \times (1 + r)^{-1}.$$

The licensee at the time t_m from entering the market receives a profit, determined by the following ratio

$$Pr_l\ (t_m) = V_m \times r_m \times (1 - f_m) \times k_{clm} \times \left(1 - \frac{4}{T_m^2}\left(t_m - \frac{T_m}{2}\right)^2\right) \times (1 + E_m)^{-n}.$$

E_m defined in point 9 of the procedure of calculating the value of IA.

V. CONCLUSION

The paper considers various methods of economic evaluation of intellectual capital and intangible assets:

- methods of direct measurement of IC (DIC);
- market capitalization methods (MCM);
- methods of return on Assets (ROA);
- scoring methods (SC).

It is shown that traditional methods based on the principles of accounting cannot take into account many elements of intellectual capital and IA, in particular the scale and level of quality of interaction with consumers, staff qualifications, knowledge base, etc.

Analysis of methods of valuation of IA (intellectual value, organizational expenses goodwill) showed that the standard approaches such as cost income and comparative meet with many difficulties, because the IA is a non-standard object for evaluation. These methods do not take into account the price and technical characteristics of the products formed in the market. In this regard, a fundamentally new method of valuation of intangible assets created on the basis of the IP, which are fundamentally new to the market, or improving products on the market. The use of this method orients the manufacturer to organize the production of competitive products for the market.

The method of intellectual capital appraisal formulated in the result of innovative projects. It can be used to determine the price of the R&D contract. In this case, its main component is the expected result of the commercialization of the created intangible assets. Thus, the method of intellectual capital is a tool to stimulate the creation of the economically perspective R&D products.

In conclusion, it should be noted that the method of calculating the cost of IA, which is the result of innovative projects, presented in the article, can be used to determine the price of the contract for R&D. In this case, its main component is the expected result of the implementation on the market of created IA. Thus, the method is a tool to stimulate performers to create market-promising results of research and economically perspective R&D products.

VI. REFERENCES

[1] Эдвинссон, Л. Корпоративная долгота: Навигация в экономике, основанной на знания. М.: Инфра-М, 2005 [Edvinsson L. Corporate longitude. Navigation in the knowledge-based economy. Moscow: Infra-M, 2005].

[2] Sveiby K. E. The New Organizational Wealth: Managing and Measuring Knowledge Based Assets. San-Francisco: Berrett Koehler, 1997. URL: http://www.sveiby.com/articles/MeasureIntangibleAssets.html.

[3] Edvinsson L., Malone M. S. Intellectual Capital: Realizing your company's true value by finding its hidden brainpower. NY: Harper Business, 1997.

[4] Стюарт Т. Интеллектуальный капитал. Новый источник богатства организаций // Новая постиндустриальная волна на Западе. Антология. М.: Академия, 1999. С. 372–400. [Stuart T. Intellectual Capital. A new source of wealth of organizations // The new wave in the post-industrial West. Anthology. Moscow: Academia, 1999. P. 372–400].

[5] Брукинг Э. Интеллектуальный капитал. СПб.: Питер, 2001. [Brooking E. Intellectual Capital. St Petersburg: Piter, 2001].

[6] Булыга Р.П. Методологические проблемы учета, анализа и аудита интеллектуального капитала. М.: Финансовая академия при Правительстве Российской Федерации, 2005. [Bulyga R.P. Methodological problems of accounting, analysis and audit of

intellectual capital. Moscow: Finance Academy under the Government of the Russian Federation, 2005].

[7] Абдикеев Н.М. Управление интеллектуальным капиталом организации // Инновационное развитие России: Проблемы и решения / Под ред. М.А. Эскиндарова, С.Н, Сильвестрова. М., 2014. С. 603–634. [Abdikeev N. M. (2014) Managing intellectual capital of the organization // Innovative development of Russia: problems and solutions / Ed. by M. A. Eskindarov, S. N. Silvestrov. Moscow, 2014. P. 603–634].

[8] Абдикеев Н.М. Киселев А.Д. Управление знаниями корпорации и реинжиниринг бизнеса. М.: Инфра-М, 2011. [Abdikeev N. M., Kiselev A. D. Knowledge management of corporation and reengineering of business. Moscow: Infra-M, 2011].

[9] Федотова М.А., Дресвянникова В.А., Лосева О.В. Интеллектуальный капитал организации: управление и оценка. М.: Финансовый университет при Правительстве Российской Федерации, 2014. [Fedotova M.A., Dresvyannikova V.A., Loseva O.V. et al. Intellectual Capital of the Organization: management and evaluation. Moscow: Financial University under the Government of the Russian Federation, 2014].

[10] Козырев А.Н., Макаров В.Л. Оценка стоимости нематериальных активов и интеллектуальной собственности. М.: Интерреклама, 2003. [Kozyrev A.N., Makarov V.L. Valuation of intangible assets and intellectual property. Moscow: Interreklama, 2003].

[11] Смирнов С.А. Оценка интеллектуальной собственности. М.: Финансы и статистика, 2002. [Smirnov S.A. Valuation of intellectual property. Moscow: Finance and statistics, 2002].

[12] Балакирева Н.М. Нематериальные активы: учет, аудит, анализ. М.: Эксмо, 2005. [Balakireva N. M. Intangible assets: accounting, auditing, analysis. Moscow: Eksmo, 2005].

[13] Белышева И.Г., Козлов Н.П. Нематериальные активы компании: классификация и учет // Акционерное общество: вопросы корпоративного управления. 2008. № 5. С. 34–37. [Belysheva I.G., Kozlov N.P. Intangible assets of company: cost evaluation // Joint-stock company: corporate society. 2008. N 5. P. 34–37].

[14] Леонов А.В., Смирнов С.С., Новanов Д.Г. Адаптивный подход к определению объемов ассигнований на развитие базовых и критических военных технологий // Вооружение и экономика. 2012. № 5(21).С. 47–59. [Leonov A.V. Smirnov S.S., Hovanov D.G. Adaptive approach to determining the volume of appropriations for the development of basic and critical military technologies // Weaponry and economics. 2012. N 5 (21). P. 47–59].

[15] Завлин П.Н., Васильев А.В. Оценка эффективности инноваций. СПб.: Бизнеспресса, 1998. [Zavlin P.N., Vasiliev A.V. Assessment of efficiency of innovations. St Petersburg: Business press, 1998].

Passenger Traffic Forecasting for New Urban Railways in Moscow

Dmitry Namiot
Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University; Russian University of Transport
Moscow, Russia
dnamiot@gmail.com

Oleg Pokusaev[1], Alexander Chekmarev[2]
Center for High-Speed Digital Transport
Russian University of Transport
Moscow, Russia
o.pokusaev@rut.digital, a.chekmarev@rut.digital

Abstract. **This article is devoted to the description of the forecasting model of passenger traffic of new urban rail lines. It is a question of new through railway routes, which will be organized in Moscow. They will pass through the city, connecting together the stations located in the suburbs. The schedule of such trains, comfort, and pricing policy will undoubtedly make them a popular transportation tool that will compete with the Moscow metro. This popularity will result in an increase in passenger traffic on existing fragments of these railway lines, the assessment of which is the task considered in this article. The article looks at the existing methods of measuring passenger traffic, and suggests a model for assessing its possible change.**

Key words: **time series; traffic forecast; digital urbanism.**

I. INTRODUCTION

In this paper, we discuss the problems of forecasting of passenger traffic for new lines of urban railways in Moscow. In 2017, Moscow announced plans [1] to create links between the existing radial rail routes within the city, it is so-called Moscow railways diameters (Fig. 1).

Fig. 1. The planned radial rail routes in Moscow

New trains are launched on existing lines of the railway and only the connectivity of the system is changed. There is a through motion, which was absent earlier. This through movement connects two branches of the railway that existed before. Before the introduction of such a single scheme of traffic, various branches of the railway were connected by means of public transport already in Moscow. In addition to increasing the connectivity of the system, the new lines will increase the frequency of movement. Also, high-quality trains will be used here, to which a convenient form of payment will be added (a combined smart transport card with city passenger transport). In fact, it looks like a new metro (the main type of public transport in Moscow). Naturally, one can expect the growth of passenger traffic for the railway lines, which formed the basis for new through-routes. The evaluation (estimation) of this increase is the task considered in this paper. The need for such an assessment is caused both by economic considerations (it is necessary to understand the economics of the new project) and technical considerations; it is necessary to answer the question whether the existing infrastructure (stations, transitions for urban public transport systems, etc.) will be sufficient to serve a new flow of passengers.

The rest of the work is structured as follows. Section 2 examines the available means of measuring passenger traffic. In Section 3, we discuss our model, which is put in the assessment of the increase in passenger traffic after the launch of new radial lines.

II.ON TRAFFIC MEASUREMENTS

At present, for the railways in Moscow, the existing passenger traffic can be measured quite accurately. Under the existing scheme, at most stations passengers must present (validate) travel documents at the entrance and at the exit. In terms of social networks, this corresponds (is analogous) to check-in marks (at the input) and check-out (on exit) [2]. This increases the value of information on the use of railway stations comparing with other transport data. From the validation information, we can immediately restore the passenger's route. Traditionally, most travel documents in Moscow (in the Moscow region) are validated only at the entrance (check-in). Accordingly, to restore the actual route,

we must use some heuristic algorithms. A review of such algorithms is, for example, in our paper [3]. The idea is that the breaks (gaps) in the use of travel documents are just the delimiters of trips. For example, if we noted the use of a travel card (in Moscow, for example, it is a Troyka card) at point A, and then after a relatively long break at point B, then we can assume that the passenger traveled from A to B, then, for example, was at work and after its completion makes a new trip. The process of route restoring based on the data of telecommunication operators also relies on heuristics [4]. For example, the place where calls are made in the morning and in the late evening is considered "home", in the daytime is considered "work" and so on. Of course, for such classification, we can impose additional conditions too. For example, we can set a minimum total time of stay for a month to recognize the place as a working place. And in the case of information on the validation for railway stations, the route is known [5].

Technically, data on validations (checking/inspection of travel documents) are presented as separate text files, each of which describes passes for a particular station in one month. One entry (a line in the file) corresponds to one pass (to the entry or exit). The data is completely anonymous and does not contain any personal information. There may be a type of travel document used (for example, it is a preferential tariff or not, it is a one-time or reused travel document, etc.), but there is basically no identification of documents at all. The size of each such log depends, of course, on the use of a particular station in a particular month and varies between 20 and 70 Mb.

Fields that are contained in the records:

- date and time;
- price characteristic (full or reduced ticket);
- type of benefits (federal, local, etc.);
- ticket type (one-way ticket, one-way ticket, one-way ticket, subscription, etc.);
- a starting station;
- an end station.

This kind of information allows to analyze summary data on the stations (total entries/exits) and, if necessary, analyze the distribution of passengers on the entries/exits too. For example, the following pictures show hourly distributions of entries and exits at a suburban station near Moscow. This is a typical picture with working traffic, where we have a peak at the entrances in the morning and a peak at the exits in the evening (in the morning people go to Moscow, in the evening they return back from work) (fig. 2, 3). Note that this picture is not necessarily for all stations.

To analyze the data, a cloud tool from Google Collaboratory [6] was used. Technically, this is a cloud implementation of Jupyter notebook. It does not require any software installation, everything is accessible through the browser. The data files were stored on Google Drive.

Thus, the processing of validator data allows receiving a correspondence matrix: for each pair of stations, specify the number of passengers traveling between the two stations. And this kind of matrix can be built in any time slice: for an hour, a day, a month, etc.

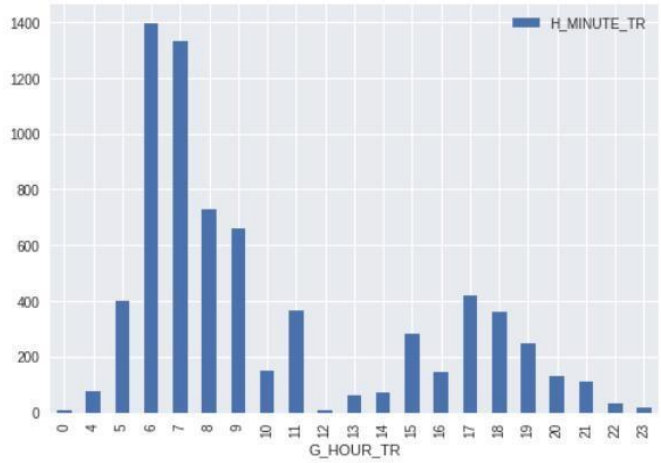


Fig. 2. Entries in suburb

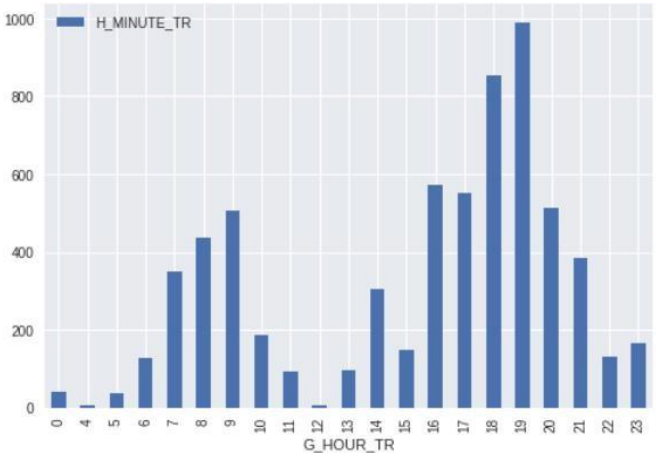


Fig. 3. Exits in suburb

This set is an objective data on the existing passengers of the railway. What can serve as a basis for assessing the potential demand for such a service?

In our case, we used data collected by mobile operators. Telecommunication operators in order to bill users for the services they provide constantly collect information about the activity of users (mobile devices). This is not related to any tracking of subscribers, this is purely an economic event (they need to know the duration of the calls made in order to calculate their cost, the numbers for which calls were made, since this can also affect their cost, etc.). In the same log, system software puts information about mobile tower which is currently serving a specific mobile device. Accordingly, the fact of changing this information is reflected too. A cellular tower (base station) has a very specific location (specific geographical coordinates). Accordingly, the location of the mobile device (change of location) can be associated with geographic coordinates. In fact, it looks more

complicated than described here (it is necessary to simulate radio signal propagation, know the location of other base stations, antenna patterns, etc.), but the result will be exactly this; we can estimate the geographic position of the mobile device.

Naturally, this estimate is made with some accuracy. This accuracy, obviously, is related to the number of base stations that participate in the calculations. Therefore, we can say that in urban conditions (high density of base stations) this accuracy will be higher than, for example, in a province where fewer people live and fewer base stations deployed. In general, under current conditions, we can expect the accuracy of several hundred meters (for example, a square of 500×500 meters). Note that locating with some kind of rigidity corresponds also needed to observe the interests of privacy of mobile subscribers. If movements are issued for all mobile subscribers from a certain geographical square, then it is not possible to track the movements of any individual subscriber.

As a result, the data of mobile operators make it possible to build a correspondence matrix (origin-destination matrix) for geographical squares. For each pair of squares (say, 500×500 meters), you can get the number of people moving from one square to another. Again, it could be done in any time frame: hour, day, month, etc. Of course, the total size of such a geographic grid can vary and the total size of the data depends on the grid. In our case, it covered a sufficient area in the suburbs of Moscow. Sufficiency, in this case, it assumes either the direct availability of a railway running through Moscow or the possibility of getting by transport to the station of this road at a reasonable time.

Information collected on geographical squares (in any time frame) can be aggregated, so that we can obtain information on the movements for administrative areas. Usually, for administrative regions, we have the population statistics. It means that we can use this statistics for verification of collected data.

The result is an objective picture of the movement of mobile devices (subscribers) from areas adjacent to railway stations. From consideration, we can immediately exclude moving to close distances (we assume that these are pedestrians). The remaining movements will be transport movements. Strictly speaking, transport modeling starts here. We suppose that we are considering a specific area in a suburb of Moscow. All these moving mobile subscribers must be somehow sorted by mode of transport. It is known exactly how many they used the railway. If there is information on the use of buses in the suburbs (for example, validation of transport cards), the difference between the total number of subscribers and the number of passengers of the railway and buses will give the number of those who will get to Moscow by cars. This is the current picture. After opening the urban railway, it will change. Flows should be redistributed between modes of transport (in fact, the main idea is that some of the motorists will transfer to rail transport), there may be new passengers who did not travel before because of the fact that getting to the city was uncomfortable.

III.On traffic models

Firstly, from the analysis of operator data on movements, for example, for a month, you can determine the nature of displacements by the type of "home-work." This is a standard approach, which was used in the earliest works on the analysis of such data [7]. The idea is to relate the time of day and user activity. If the mobile subscriber is constantly active (calls, writes SMS, etc.) from a certain region in the evening, then this region can be recognized as his "home". Similarly, the constant activity from a particular region during the daytime leads to the classification of this region as a "work". The term "permanently" defines here the time required for the classification. For example, it could be 7 days during a month. This value is given empirically. In paper [8], the algorithm assumes the location from which a user departs in the morning and to which he or she returns at night is home. It also infers that the location of the longest recurring stays during weekday daytime hours is the user's workplace.

After this classification, two things become available. Firstly, we can describe the displacement in terms of home-work. This can be a more stable pattern of displacement. Secondly, it becomes possible to estimate the actual population by geographical objects (administrative units, if aggregate information from squares).

Estimating the number, in this case, is the first point in the assessment of passenger traffic. This is obviously the upper estimate of possible railroad traffic. This is the theoretically possible number of passengers. There are simply no more of them.

At the same time, we assume that the distribution of "new" passengers by stations will be exactly the same as the current distribution of railway passengers by stations. In other words, each station will receive its proportional increase in passengers.

From a practical point of view, in most areas (stations), the figures obtained exceeded the physical capacity (capacity) of the respective stations.

There is one important observation about the classification of movements by home-work pattern. It is not necessary that both sites for a particular subscriber will be classified. For example, subscribers with traveling related work. They do not stay long in one place to classify their activities there. But at the same time, they create a load on the transport system. On the other hand, there may be subscribers who do not really move anywhere. In this case, the home area for them will be determined; there will be really registered telephone activities at the place of residence.

The analysis of passenger activity by stations shows a very high level of constancy. If, for example, we build a one hour aggregated graph for the entrances to a typical Moscow suburban station (see Fig. 2), then such time series for different days of the week will be very similar. The degree of similarity of this kind of time series, measured by different metrics, is quite large. For example, in our work, we have successfully used a shape-based similarity measure, it is Angular Metric for Shape Similarity (AMSS) [9].

This approach treats a time series as a vector sequence and focus on the shape of the data and compares data shapes by employing a variant of cosine similarity. It is illustrated in Fig. 4.

The cosine similarity metrics minimize the influence of outliers in similarity computation, where outliers are defined as much bigger or smaller data points than their immediate neighbors [9].

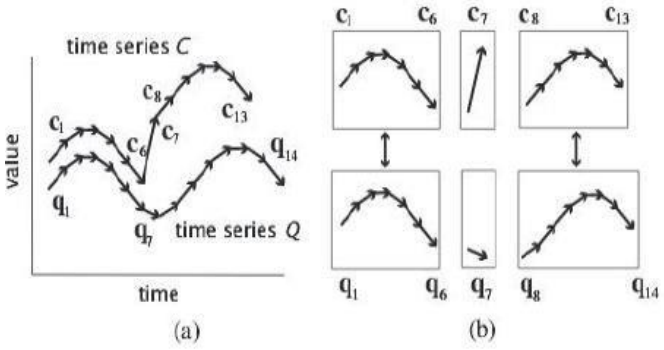


Fig. 4. Angular Metric for Shape Similarity [9]

This means that the railway traffic is fairly stable. The railway has its regular passengers who use it. And the fluctuations here (before the launch of new traffic conditions) are rather small.

The telecom data is here the only available information on the behavior of passengers. Historical data for this kind of projects are not available for obvious reasons. There are also no survey estimates. Accordingly, several heuristics were used to estimate the future passenger traffic, based on the characteristics of the current and the maximum possible passenger traffic [10]. Let's look at these heuristics.

A. We estimate the share of railway passengers for each of the stations of the new route. In other words: what percentage of all those moving from the area (according to the statistics of the mobile operator) is actually transported by the railway (according to the railway station validation data).

Determine the maximum percentage across all stations. And after that, we assume that such a percentage (coverage) will be at each station. So, we can recalculate traffic for all stations based on a new coverage.

Just for the explanation: 100% of the movements will not transfer to the railway. The maximum percentage of coverage by the railway is a real achievable share. Thus, we determine the new number of passengers (entrances) for the stations, and then these passengers are distributed to the other stations in the same proportion as it exists for the real data.

B. We can estimate the share of new passengers for railway stations according to seasonal trends. In summer, the number of residents who travel to the city from suburbs increases at the expense of those townspeople who live in dachas, but travel to the city.

Accordingly, we will get some increase in the number of people in the suburbs, which moves (according to the telecom operator) from the area to the city and back. Of these trips, some share will have to rail transport, which can be objectively determined according to the validators (according to the real data). The share of new "residents" who chose rail transport can be extrapolated to the entire population of the district and thus calculate the possible share of the railway in the transportation of residents in the area.

A possible explanation: the new "residents" do not have stable preferences in comparison with permanent residents, and their choice will reflect the objective correlation of transport shares.

C. Most of the inhabitants of the suburbs who travel at least once used the railroad. This assumption seems reasonable, since the railway, in any case, is the simplest way to travel to Moscow. So, the appearance of new passengers at the railway is possible only due to the fact that those who have traveled rarely will travel more often. Accordingly, in addition to counting the number of trips, it is necessary to calculate their frequency distribution. Technically, this is a grouping of data on the movement between areas (territories, cells) by the user ID. Data from operators contain some kind of subscriber identification (for example, IMEI or hash code, which replaces this identifier to preserve privacy). If we calculate the number of user movements per month from the area to the area, we can get useful frequency information too. E.g., it is some like this:

> *N* movements were made by *X* subscribers
> *N-1* displacements were made by *Y* subscribers *N-2* movements made by *Z* subscribers etc.

For reasons of privacy, we can also cluster the data by the number of displacements. For example, in steps of 5:

> from 5 to 10 movements made by *X* subscribers;
> from 10 to 15 movements made by *Y* subscribers;
> from 15 to 20 movements made by *Z* subscribers, etc.

This will give the distribution of residents (those who travel) on the frequency of travel. And this distribution shows the reserves of growth in travel. In this distribution, we will be interested in the low-frequency "tail". In fact, we are interested in aggregated data again (the number of trips) but only for a part of the inhabitants (for low-frequency travelers). We proposed to sum these low-frequency trips and then conduct something like a stress test. For the number of trips increases by 25% (50%, 100%) recalculate this percentage in thousands of passengers. Then we can share this amount of new passengers by the stations using the existing distribution of passengers (calculated by real validation data). And for each growth figure, it will be possible to assess whether there are enough trains (capacity of platforms, the throughput of turnstiles, etc.) and what will

be the increase in passengers at metro stations (when passengers change from railways to metro).

D. For areas where it is possible to count motorists (as a difference between all movements and validations of travel documents on the railway and public transport), we can use historical data on the launch of urban railways abroad. The idea of replacing motorists on the railway, in fact, was always one of the main for urban railways. In the literature, it was noted that the railway to the cities is the only transport today that can move without traffic jams and, at the same time, provide a fairly high level of comfort. It is noted that the city railway is the only transport today, for which motorists are transplanted. Because, for example, using a bus often means meet the same traffic jams as on a personal car, but only with a lower level of comfort (as opposed to a personal car). But it is obvious that 100% of the replacement still does not happen. The figure that occurs in the literature is up to 20–25% of the users of vehicles that switched to the railway. This was determined, of course, after the projects were put into operation. That's exactly this figure - 20% of the number of motorists can be used to calculate the increase in the number of passengers in the railway.

IV.CONCLUSION

In this paper, the models of an estimation of passengers' traffic for new lines of a city railway projected in Moscow are considered. The initial data for building the forecast was information on the current load of the railway transport, data on the use of public transport (in some areas outside Moscow) and measurements based on the data of mobile operators and describing the movement of mobile subscribers in the Moscow region. The model offers a number of heuristic approaches to traffic estimation, which are combined with real data of existing rail traffic.

V. ACKNOWLEDGMENT

Some of the ideas presented here were previously condemned by us at data analysis conferences (AIST 2018, DAMDID 2018). We are grateful to the reviewers of these conferences. Their criticism and constructive comments allowed us to seriously improve our work.

VI. REFERENCES

[1] Shneps-Shneppe D. On Digital Signaling for Moscow City Railways // International Journal of Open Information Technologies. 2018. Vol. 6, N 6. P. 28–37.

[2] Namiot D. Sneps-Sneppe M. Customized check-in procedures // Smart Spaces and Next Generation Wired/Wireless Networking. Berlin; Heidelberg: Springer, 2011. P. 160–164.

[3] Namiot D., Sneps-Sneppe M. A Survey of Smart Cards Data Mining // Supplementary Proceedings of the Sixth International Conference on Analysis of Images, Social Networks and Texts (AIST 2017) Moscow, Russia, July 27-29, 2017. Moscow, 2017.

[4] Steenbruggen J., Borzacchiello M.T., Nijkamp P. et al. Mobile phone data from GSM networks for traffic parameter and urban spatial pattern assessment: a review of applications and opportunities // GeoJournal. 2013. Vol. 78, N 2. P. 223–243.

[5] Ratti C., Frenchmann D., Pulselli R.M. et al. Mobile landscapes: using location data from cell phones for urban analysis // Environment and Planning B: Planning and Design. 2006. Vol. 33, N 5. P. 727–748.

[6] Google Collaboratory https://research.google.com/colaboratory/unregistered.html.

[7] Wang J., Li J., He K. et al. Vulnerability analysis and passenger source prediction in urban rail transit networks // PloS one. 2013. Vol. 8. N 11. P. e80178.

[8] Mearian L. Ford, MIT use Bostonians' cellphone location data for traffic planning // Computerworld. URL: https://www.computerworld.com/article/3112845/cartech/ford-mit-use-bostonians-cellphone-location-data-for-city-trafficplanning.html.

[9] Nakamura T., Taki K., Nomiya H. et al. A shape-based similarity measure for time series data with ensemble learning // Pattern Analysis and Applications. 2013. Vol. 16, N 4. P. 535–548.

[10] Намиот Д.Е., Покусаев О.Н., Лазуткина В.С. О моделях пассажирского потока для городских железных дорог // International Journal of Open Information Technologies. 2018. Vol. 6, N 3. P. 9–14. [Namiot D., Pokusaev O., Lazutkina V. On passenger flow data models for urban railways // International Journal of Open Information Technologies. 2018. Vol. 6, N 3. P. 9–14.