ISSN 2713-3192
DOI 10.15622/ia.2022.5.21
http://ia.spcras.ru

ТОМ 21 № 5

ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

INFORMATICS AND AUTOMATION

СПИИРАН
WWW.SPIRAS.NW.RU
СПб ФИЦ РАН

Санкт-Петербург
2022

INFORMATICS AND AUTOMATION

Volume 21 № 5, 2022

Scientific and educational journal primarily specialized in computer science, automation, robotics, applied mathematics, interdisciplinary research
Founded in 2002

Founder and Publisher
St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS)

Editorial office address
SPC RAS, 14-th line V.O., 39 litera A , St. Petersburg, Russia, 199178
e-mail: ia@spcras.ru, web: http://ia.spcras.ru
The journal is indexed in Scopus
The journal is published under the scientific-methodological supervision of Department for Nanotechnologies and Information Technologies of the Russian Academy of Sciences

ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

Том 21 № 5, 2022

Научный, научно-образовательный журнал с базовой специализацией
в области информатики, автоматизации, робототехники, прикладной математики
и междисциплинарных исследований.
Журнал основан в 2002 году

Учредитель и издатель

Федеральное государственное бюджетное учреждение науки
«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук»
(СПб ФИЦ РАН)

Адрес редакции

199178, г. Санкт-Петербург, 14-я линия В.О., д. 39, литер А
e-mail: ia@spcras.ru, сайт: http://ia.spcras.ru

Журнал индексируется в международной базе данных Scopus

Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий,
в которых должны быть опубликованы основные научные результаты диссертации
на соискание ученой степени доктора и кандидата наук»

Журнал выпускается при научно-методическом руководстве Отделения нанотехнологий
и информационных технологий Российской академии наук

CONTENTS

СОДЕРЖАНИЕ

DOI 10.15622/ia.21.5.1

L. UTKIN , A. KONSTANTINOV

RANDOM SURVIVAL FORESTS INCORPORATED BY THE NADARAYA-WATSON REGRESSION

Utkin L., Konstantinov A. **Random Survival Forests Incorporated by the Nadaraya-Watson Regression.**

Abstract. An attention-based random survival forest (Att-RSF) is presented in the paper. The first main idea behind this model is to adapt the Nadaraya-Watson kernel regression to the random survival forest so that the regression weights or kernels cannbsp;be regarded as trainable attention weights under important condition that predictions of the random survival forest are represented in the form of functions, for example, the survival function and the cumulative hazard function. Each trainable weight assigned to a tree and a training or testing example is defined by two factors: by the ability of corresponding tree to predict and by the peculiarity of an example which falls into a leaf of the tree. The second main idea behind Att-RSF is to apply the Huber's contamination model to represent the attention weights as the linear function of the trainable attention parameters. The Harrell's C-index (concordance index) measuring the prediction quality of the random survival forest is used to form the loss function for training the attention weights. The C-index jointly with the contamination model lead to the standard quadratic optimization problem for computing the weights, which has many simple algorithms for its solution. Numerical experiments with real datasets containing survival data illustrate Att-RSF.

Keywords: machine learning, random survival forest, survival analysis, Harrell's C-index, cumulative hazard function, attention mechanism, Huber's contamination model.

1. Introduction. Survival analysis can be regarded as an important and fundamental tool for modelling applications using time-to-event data [1]. In various spheres of life, including, medicine, reliability, safety, finance and economics, we encounter time-to-event data. Therefore, many machine learning models have been proposed to deal with time-to-event data and to solve the corresponding problems in the framework of survival analysis [2–5].

Three types of survival models can be considered [4]. Models of the first type, called parametric models, assume that a probability distribution of time to event is known, but its parameters are unknown and should be estimated. Models of the second type, called semi-parametric models, do not assume any probability time-to-event distribution, but assume that there is some known functional dependence between covariates and the model outcomes. The well-known semi-parametric model is the Cox proportional hazards model [6] which can be regarded as a regression model. Models of the third type, called non-parametric, do not use any information about a probability time-to-event distribution as well as a relationship between covariates and the model outcomes. The well-known non-parametric survival model is the Kaplan-Meier model [4]. An important peculiarity of many survival models is that

their outcomes are functions, for example, survival functions, hazard functions, cumulative hazard functions, but not point-valued data.

Following the Cox model, many of its modifications overcoming some disadvantages of the Cox model have been developed, for example, models based on the Lasso method [7], models generalizing the Cox model by using neural networks [3], the support vector machine [8], survival trees [9], random survival forests (RSFs) [10] as an extension of the original random forest (RF) [11]. Due to the small number of tuning parameters, due to the ability to deal with both low and high-dimensional data, due to adaptability to data, RSFs became a popular tool for survival analysis of time-to-event data in many applications. RSFs have demonstrated their efficiency in solving many real problems [12–18].

One of the ways to improve RSF is to replace the standard averaging with the weighted sum of the tree survival functions. Following this idea, the corresponding weighted RSF was proposed in [19]. According to the weighted RSF, every tree is assigned by a weight which is computed by solving an optimization problem maximizing the concordance error rate called C-index [20]. The main disadvantage of the weighted RSF is that it uses weights which do not depend on each example and are defined only by the corresponding survival tree. This fact reduces the RSF accuracy. In order to overcome this difficulty, we propose the attention-based RSF (Att-RSF). The idea behind Att-RSF is to adapt the Nadaraya-Watson regression to the original RSF. In other words, every survival tree jointly with an example, which falls into the tree, is considered as a term in the Nadaraya-Watson regression with a weight which is trained through its trainable parameters. The idea to assign weights to trees in the RF in accordance with the tree importance and with the example importance is not new, and it was proposed in [21] where attention weights are trained by solving the quadratic optimization problem. It turns out that this idea to consider the RF as the Nadaraya-Watson regression can be extended to RSF taking into account the RSF peculiarities which differ RSF from the RF. In particular, we propose to optimize the model parameters in accordance with the C-index as a measure of the RSF accuracy instead of the simple difference between predicted values and true labels used in the RF. This leads to a quite different optimization problem.

Similar to the attention-based RF [21], the Huber's ϵ-contamination model [22] is introduced to define the trainable parameters of the attention weights such that these trainable parameters of weights are optimally selected from an arbitrary adversary distribution. The ϵ-contamination model allows us to introduce weights being a linear function of the C-index.

Our contributions can be summarized as follows:

1. A new attention-based RSF model is proposed. According to the model, the trainable attention mechanism is incorporated into RSF to improve the accuracy of obtained predictions.

2. The proposed attention-based RSF can be regarded as an adaptation of the Nadaraya-Watson kernel regression to RSF. Moreover, we extend the Nadaraya-Watson kernel regression to predictions in the form of functions, for example, the cumulative hazard function.

3. Numerical experiments with real datasets are provided to justify Att-RSF, and to compare it with original RSFs [10] and the weighted RSF proposed in [19].

The paper is organized as follows. Related work devoted to machine learning models in survival analysis, the attention mechanism and the weighted RFs can be found in Section 2. Section 3 provides basic definitions of survival analysis, RSFs and the Nadaraya-Watson regression jointly with the attention mechanism. The main ideas of the Att-RSF and algorithms for training optimal attention parameters are considered in Section 4. Numerical experiments with well-known public real data illustrating the proposed Att-RSF model and comparing it with the available survival machine learning models are given in Section 5. Concluding remarks are provided in Section 6.

2. Related work. Machine learning models in survival analysis. Many survival machine learning models dealing with time-to-event data have been developed and investigated to predict survival time or other survival measures. A comprehensive review of the recent survival machine learning models is presented by [4]. The most popular survival model is the semi-parametric Cox proportional hazards model [6] which establishes a linear relationship between the covariates and the distribution of survival times. Tibshirani [7] presented a modification based on the Lasso method. Similar Lasso modifications, for example, the adaptive Lasso, were also proposed by several authors [23, 24] The linear relationship can be viewed in some applications as a disadvantage which can be resolved by relaxing the linear relationship assumption and extending the Cox model to more complex models [2, 25, 26]. At the present time, survival models can be regarded as extensions of many well-known machine learning models, for example, the Lasso models [23], SVM [8], decision trees [9], neural networks [2, 26, 27], etc.

We pay attention to random survival forests (RSFs) which can be regarded as one of the most powerful and efficient tools for survival analysis especially when the training data are tabular. Various implementations and modifications of RSFs were considered and studied in [14, 15, 17, 19, 28, 29]. To improve the available RSF models, it is proposed to incorporate the attention

mechanism with trainable parameters into RSFs, which allows us to take into account the importance of trees in RSF as well as the importance of every training or testing example.

Attention mechanism. Many attention-based models have been developed to improve the performance of classification and regression algorithms. Detailed and comprehensive surveys of attention models can be found in [30–35]. It is important to point out that attention models are mainly applied to the natural language processing, including text classification, translation, etc., to the computer vision area, including image-based analysis, visual question answering, etc. However, time-to-event data in many applications have a tabular form. An attempt to incorporate the attention mechanism into the RF was made in [21]. Following this work, we try to incorporate the attention mechanism into RSF by using peculiarities of survival models which include: predictions in the form of functions of time, the model accuracy measure in the form of the C-index, and censored data. The proposed attention mechanism can be also regarded as an extension of the weighted RFs.

Weighted RFs. Various models and methods have been developed to implement the weighted RFs. They can be divided into two groups. The first group consists of models which are based on assigning weights to decision trees in accordance with some criteria to improve the classification and regression models [36–38]. This group contains models using weights of classes to take into account imbalanced datasets [39]. However, the assigned weights in the aforementioned works are not trainable parameters. Therefore, models [40, 41] from the second group use trainable weights of trees such that the weights are trained by solving optimization problems in accordance with a certain loss function for the whole RF. The model of the weight assigning in [21] differs from the above models because weights are assigned depending on trees and each example.

Our aim is to incorporate the attention weights with trainable parameters into RSF and to propose simple algorithms for training the parameters.

3. Preliminaries.

3.1. Survival analysis. The i-th patient in survival analysis is represented by a triplet $(\mathbf{x}_i, \delta_i, T_i)$, where $\mathbf{x}_i \in \mathbb{R}^m$ is the feature vector characterizing the patient; T_i is the time to an event of interest; δ_i is the indicator of event observation, in particular, $\delta_i = 1$ if the corresponding event is observed (an uncensored observation), $\delta_i = 0$ if the event is not observed and the corresponding time to event is greater than T_i (a censored observation). Survival analysis aims to estimate time T to the event for a new patient having feature vector $\mathbf{x}$ on the basis of a training set D consisting of n triplets $(\mathbf{x}_i, \delta_i, T_i)$,

$i = 1, ..., n$. We will use the term "patient" to represent arbitrary subjects or objects.

It is important to point out that only a part of the patients will experience the event of interest during the course of the experiment. Other patients will not experience the event of interest after the expiration of the study. Therefore, observation is said to be censored in survival analysis when information on time to the corresponding event of interest is not available, i.e., we have some information about the patient survival time, but we do not know the survival time exactly. One should distinguish between right censoring, left censoring and interval censoring. Right censoring occurs when the study of a patient ends before the event has occurred. Left censoring is when the event of interest has already occurred before studying. This is a very rare case. Interval censoring is a combination of left and right censoring. It should be noted that right censoring is the most common type of censoring, therefore, we consider only this type. Parameter δ_i indicates whether the i-th event is censored or not.

Important concepts in survival analysis are the survival function (SF) and the cumulative hazard function (CHF). The SF $S(t|\mathbf{x})$ is a function of time t defined as the probability of surviving up to time t, i.e.: $S(t|\mathbf{x}) = \Pr\{T > t|\mathbf{x}\}$. The CHF $H(t|\mathbf{x})$ is also a function of time defined through the SF as follows:

$$H(t|\mathbf{x}) = -\ln S(t|\mathbf{x}). \tag{1}$$

Many survival machine learning models have been developed in the last decades. In order to compare the models, special measures are used differently from the standard accuracy measures accepted in machine learning classification and regression models. The most popular measure in survival analysis is Harrell's C-index (concordance index) [20]. It estimates the probability that, in a randomly selected pair of patients, the patient that fails first had the worst predicted outcome. In fact, this is the probability that the event times of a pair of patients are correctly ranked. C-index does not depend on choosing a fixed time for evaluation of the model and takes into account censoring of patients [42].

Let us consider the training set D consisting of n triplets $(\mathbf{x}_i, \delta_i, T_i)$. We consider possible or admissible pairs $\{(\mathbf{x}_i, \delta_i, T_i), (\mathbf{x}_j, \delta_j, T_j)\}$ for $i \leqslant j$. Then the C-index is calculated as the ratio of the number of pairs correctly ordered by the model to the total number of admissible pairs. A pair is not admissible if the events are both right-censored or if the earliest time in the pair is censored. If the C-index is equal to 1, then the corresponding survival model is supposed to be perfect. If the C-index is 0.5, then the model is not better than random guessing.

Let $t_1, ..., t_N$ denote predefined time points of the corresponding N distinct event times. If the output of a survival model is the predicted SF $S(t)$, then the C-index is formally calculated as [4]:

$$C = \frac{1}{M} \sum_{i:\delta_i=1} \sum_{j:t_i<t_j} \mathbf{1}\left[S(t_i|\mathbf{x}_i) - S(t_j|\mathbf{x}_j) > 0\right]. \quad (2)$$

Here M is the number of all comparable or admissible pairs; $\mathbf{1}[\cdot]$ is the indicator function taking value 1 if its argument is true, and 0 if the argument is false.

3.2. Random survival forests. In spite of the efficiency of deep neural networks, RSFs can be regarded as one of the best models for survival analysis due to their properties especially when the tabular training data are used. Therefore, we modify RSFs to improve their prediction capacity.

A general algorithm for constructing RSFs can be represented as follows [43]:

1. Q subsets of training data are selected to build Q trees in RSF. Each subset excludes on average 37% of the data, is called out-of-bag data (OOB data).

2. Each survival tree is built on the corresponding subset. At each node of the tree, $\sqrt{m}$ candidate variables are randomly selected. The node is split using the candidate variable that maximizes the survival difference between daughter nodes.

3. Each tree is built to full size under the constraint that a terminal node should have no less than $d > 0$ unique events. Here d is a tuning parameter which is chosen to get the best results.

4. CHFs or SFs are calculated for each tree. The ensemble CHF or the ensemble SF are obtained by averaging CHFs or SFs of trees.

5. Using out-of-bag data, prediction errors for the ensemble CHF or the ensemble SF are calculated.

The accuracy of RSF predictions is defined by a splitting rule. A good split maximizes survival difference across the two sets of data [43]. There are several splitting rules used in RSF [4, 43]. We do not consider them because the proposed approach does not depend on a splitting rule.

Before computing the ensemble CHF or the ensemble SF having CHF or SFs of trees, we consider how to compute the CHF for the k-th terminal node of a tree. Let $\{t_{j,k}\}$ be a set of $N(k)$ distinct event times in terminal node k of the q-th tree such that $t_{1,k} < t_{2,k} < ... < t_{N(k),k}$ and $Z_{j,k}$ and $Y_{j,k}$ equal to the number of events and patients at risk at time $t_{j,k}$. The CHF for node k is

defined by using the Nelson–Aalen estimator as follows:

$$H_k(t) = \sum_{t_{j,k} \leqslant t} Z_{j,k}/Y_{j,k}. \tag{3}$$

If the i-th patient with features $\mathbf{x}_i$ falls into node k, then one can say that $H(t|\mathbf{x}_i) = H_k(t)$. The ensemble CHF for the i-th patient is obtained by averaging CHFs of all Q trees, i.e.,

$$H_f(t|\mathbf{x}_i) = \frac{1}{Q}\sum_{q=1}^{Q} H_q(t|\mathbf{x}_i). \tag{4}$$

The SF can be obtained from $H_q(t|\mathbf{x}_i)$ as follows:

$$S_q(t|\mathbf{x}_i) = \exp\left(-H_q(t|\mathbf{x}_i)\right). \tag{5}$$

Ishwaran et al. [43] proposed another ensemble estimate using OOB data. Suppose that tree q is built on a set of OBB examples with indices from set O_q. The OOB prediction for each training example $\mathbf{x}_i$ uses only the trees that did not have $\mathbf{x}_i$ in their bootstrap sample. If to denote the indicator function as $\mathbf{1}(i \in O_q)$, then the OOB ensemble CHF for the i-th training example is estimated as:

$$H_f(t|\mathbf{x}_i) = \frac{\sum_{q=1}^{Q} \mathbf{1}(i \in O_q) \cdot H_q(t|\mathbf{x}_i)}{\sum_{q=1}^{Q} \mathbf{1}(i \in O_q)}. \tag{6}$$

3.3. Attention mechanism and the Nadaraya-Watson regression. The idea of the attention mechanism can clearly be explained by using the Nadaraya-Watson kernel regression model [44, 45]. If there is a training set $\{(\mathbf{x}_1, y_1), ..., (\mathbf{x}_n, y_n)\}$ consisting of n examples, where $\mathbf{x}_i \in \mathbb{R}^m$ is a feature vector and $y_i \in \mathbb{R}$ is the corresponding label, then the regression output prediction z, associated with a new input feature vector $\mathbf{x}$, can be estimated as the weighted average in the form of the Nadaraya-Watson kernel regression model [44, 45]:

$$z = \sum_{i=1}^{n} \alpha(\mathbf{x}, \mathbf{x}_i) y_i. \tag{7}$$

Here $\alpha(\mathbf{x}, \mathbf{x}_i)$ is the attention weight which measures how the feature vector $\mathbf{x}$ is far from the feature vector $\mathbf{x}_i$ from the training set. The closer $\mathbf{x}$ to $\mathbf{x}_i$, the greater the corresponding weight $\alpha(\mathbf{x}, \mathbf{x}_i)$. Generally, arbitrary distance

functions satisfying the above condition can be regarded as the attention weights. One of the sets of the functions is the kernel set because a kernel K can be regarded as a scoring function estimating how vector $\mathbf{x}_i$ is close to vector $\mathbf{x}$. Hence, the attention weights can be represented as:

$$\alpha(\mathbf{x}, \mathbf{x}_i) = \frac{K(\mathbf{x}, \mathbf{x}_i)}{\sum_{j=1}^{n} K(\mathbf{x}, \mathbf{x}_j)}. \tag{8}$$

In terms of the attention mechanism [46], vector $\mathbf{x}$, vectors $\mathbf{x}_i$ and labels y_i are called *query*, *keys* and *values*, respectively. Weights $\alpha(\mathbf{x}, \mathbf{x}_i)$ can be extended by incorporating trainable parameters. For example, if we take the Gaussian kernel with a trainable vector of parameters $\mathbf{w} = (w_1, ..., w_n)$, then the attention weight can be represented as:

$$\alpha(\mathbf{x}, \mathbf{x}_i, \mathbf{w}) =$$
$$= \text{softmax}\left(-\left\|\mathbf{x} - \mathbf{x}_i\right\|^2 |\mathbf{w}\right) = \frac{\exp\left(w_i \|\mathbf{x}-\mathbf{x}_i\|^2\right)}{\sum_{j=1}^{n} \exp\left(w_i \|\mathbf{x}-\mathbf{x}_i\|^2\right)}. \tag{9}$$

There exist several definitions of attention weights and the corresponding attention mechanisms, for example, the additive attention [46], multiplicative or dot-product attention [47,48]. We use a new attention mechanism which is based on the weighted RSFs training and the Huber's ϵ-contamination model.

4. Attention-based RSF.

4.1. Queries, keys and values in RSFs. The main idea behind Att-RSF is to adapt the Nadaraya-Watson regression to the original RSF. It can be done if we consider a prediction of each tree as a *value* in the terminology of the attention mechanism, define the parametric attention weight for each tree in a specific way, and find a simple way to compute the trainable parameters of the attention weight in accordance with some objective function which is responsible for the survival model accuracy.

First, predictions of trees as values in the Nadaraya-Watson regression are SFs $S_q(t|\mathbf{x}_i)$ or CHFs $H_q(t|\mathbf{x}_i)$, $q = 1, ..., Q$. Parameters of the attention weights are proposed to define through the Huber's ϵ-contamination model. The objective function depending on the trainable parameters of the Huber's ϵ-contamination model is proposed to define by using an approximation of the RSF C-index which is maximized to get the optimal attention to trainable parameters. Moreover, the approximation of the C-index is carried out in a way which leads to the standard quadratic optimization problem.

Denote a set of leaf nodes belonging to the k-th tree as $\mathcal{Q}^{(k)} = \{q_1^{(k)}, ..., q_{s_k}^{(k)}\}$, where $q_i^{(k)}$ is the i-th leaf in the k-th tree, $k = 1, ..., Q$; s_k is the number of leaves in the k-th tree. Suppose that an example $\mathbf{x}$ falls into the i-th leaf, i.e., into leaf $q_i^{(k)}$. Let us also introduce the mean vector $\mathbf{A}_k(\mathbf{x})$ as the mean of training example vectors, which fall into the i-th leaf of the k-th tree, i.e., there holds:

$$\mathbf{A}_k(\mathbf{x}) = \frac{1}{\#\mathcal{J}_i^{(k)}} \sum_{j \in \mathcal{J}_i^{(k)}} \mathbf{x}_j, \tag{10}$$

where $\mathcal{J}_i^{(k)}$ is the index set of examples which fall into leaf $q_i^{(k)}$, and there holds $\mathcal{J}_i^{(k)} \cap \mathcal{J}_l^{(k)} = \varnothing$ for arbitrary two leaves with indices i and l in the k-th tree such that $i \neq l$; $\#\mathcal{J}_i^{(k)}$ is the number of elements in $\mathcal{J}_i^{(k)}$.

It should be noted that a single example can fall only into one leaf from $\mathcal{Q}^{(k)}$. Therefore, there is no need to use the index of the leaf in the notation for $\mathbf{A}_k(\mathbf{x})$. Mean values $\mathbf{A}_k(\mathbf{x})$ play the role of *keys* in the terminology of the attention mechanism. Indeed, every leaf node localizes a set of examples from the training set, which are close to each other. Since the tree prediction is the average of SFs or CHFs associated with examples $\mathbf{x}_j$, $j \in \mathcal{J}_i^{(k)}$, from the local set, then it makes sense to average the corresponding feature vectors. In fact, $\mathbf{A}_k(\mathbf{x})$ can be regarded as a prototype of examples localized by the leaf node. This implies that the proposed Att-RSF deals with local subsets of examples as *keys* and *values*, but not with separate examples. It is important to point out that the attention-based model can be detailed to deal with every example. However, the number of trainable parameters rapidly increases in this case and may lead to overfitting and worse results.

We denote the CHF and the SF of example $\mathbf{x}$, which falls into leaf $q_i^{(k)}$, as $H_k(t|\mathbf{x})$ and $S_k(t|\mathbf{x})$, respectively. If an example $\mathbf{x}$ (training or testing) falls into leaf $q_i^{(k)} \in \mathcal{Q}^{(k)}$ of the k-th tree, then distance $d(\mathbf{x}, \mathbf{A}_k(\mathbf{x}))$ shows how far the feature vector $\mathbf{x}$ is from the mean feature vector of all examples which fall into leaf $q_i^{(k)}$. We use the L_2-norm for the distance definition, i.e., $d(\mathbf{x}, \mathbf{A}_k(\mathbf{x})) = \|\mathbf{x} - \mathbf{A}_k(\mathbf{x})\|^2$. Note that each tree has only a single leaf which an example falls into.

The final RSF prediction $H(t|\mathbf{x})$ for a testing example $\mathbf{x}$ is defined as:

$$H(t|\mathbf{x}) = \frac{1}{Q} \sum_{k=1}^{Q} H_k(t|\mathbf{x}). \tag{11}$$

Let us return to the definition of the Nadaraya-Watson regression model and rewrite it in terms of RSF as follows: $H_k(t|\mathbf{x})$,

$$H(t|\mathbf{x}) = \sum_{k=1}^{Q} \alpha\left(\mathbf{x}, \mathbf{A}_k(\mathbf{x}), \mathbf{w}\right) \cdot H_k(t|\mathbf{x}). \tag{12}$$

Here $\alpha\left(\mathbf{x}, \mathbf{A}_k(\mathbf{x}), \mathbf{w}\right)$ is the attention weight which does not depend on time t and conforms with the relevance of "mean example" $\mathbf{A}_k(\mathbf{x})$ to vector $\mathbf{x}$ and satisfies condition:

$$\sum_{k=1}^{Q} \alpha\left(\mathbf{x}, \mathbf{A}_k(\mathbf{x}), \mathbf{w}\right) = 1, \tag{13}$$

$\mathbf{w}$ is a vector of the trainable attention parameters which will be defined below in accordance with the model modification.

The same can be written for SFs as:

$$S(t|\mathbf{x}) = \sum_{k=1}^{Q} \alpha\left(\mathbf{x}, \mathbf{A}_k(\mathbf{x}), \mathbf{w}\right) \cdot S_k(t|\mathbf{x}). \tag{14}$$

In terms of the attention mechanism, $H_k(t|\mathbf{x})$, $k = 1, ..., Q$, are *values*, $\mathbf{A}_k(\mathbf{x})$, $k = 1, ..., Q$, are *keys*, and $\mathbf{x}$ is the *query*. A scheme of the introduced terms is depicted in Figure 1. It illustrates a survival tree with the leaf $q_i^{(k)}$ where the vector $\mathbf{x}$ falls into.

4.2. Attention weights and the ϵ-contamination model. The next question is how to define the attention weights $\alpha\left(\mathbf{x}, \mathbf{A}_k(\mathbf{x}), \mathbf{w}\right)$ depending on the trainable parameters $\mathbf{w}$ to compute the parameters. Incorporating weights into the softmax function as it is shown in (9) leads to a computationally hard optimization problem. Moreover, it will be seen below that the optimization problem for computing the attention weights is constrained, and it is difficult to solve it by using the gradient-based algorithms.

Taking into account the above, we use a simple representation of attention weights proposed in [21], which leads to the linear or quadratic optimization problem whose solution is the optimal vector $\mathbf{w}$ of trainable parameters. The representation is based on applying the Huber's ϵ-contamination model [22] which is represented as $F = (1 - \epsilon) \cdot P + \epsilon \cdot R$.

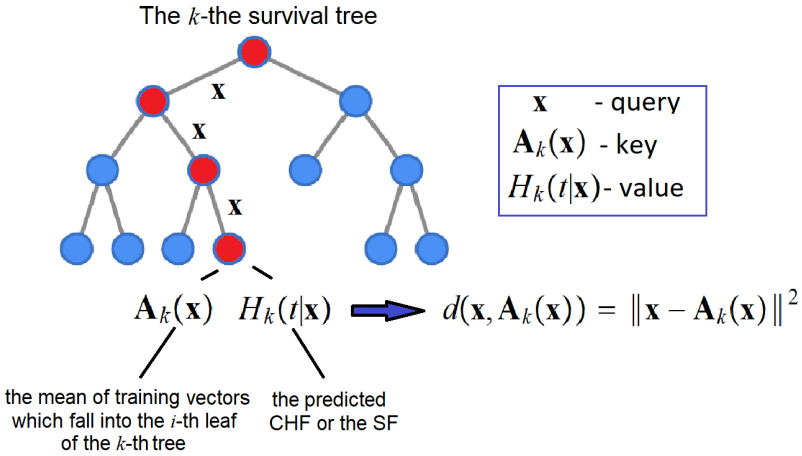


Fig. 1. A scheme of the introduced terms related to a survival tree and the attention mechanism, which include the query, keys and values

Here $P = (p_1, ..., p_Q)$ is a discrete probability distribution contaminated by another probability distribution denoted $R = (r_1, ..., r_Q)$, i.e., $p_1 + ... + p_Q = 1$ and $r_1 + ... + r_Q = 1$; the contamination parameter $\epsilon \in [0, 1]$ controls the degree of the contamination. It follows from the definition of the contamination model that R is a point in the unit simplex denoted as $U(1, Q)$ and having dimensionality Q. Hence, the subset of points F produced by the ϵ-contamination model is a subset of the unit simplex such that its center is the distribution P, its size is defined by hyperparameter ϵ. In particular, if $\epsilon = 1$, then the subset of points F coincides with the unit simplex, and if $\epsilon = 0$, then the subset of points F is reduced to point P.

If we assume that p_k is a result of the softmax operation, i.e., $p_k = \text{softmax}\left(-\|\mathbf{x} - \mathbf{A}_k(\mathbf{x})\|^2\right)$, and the probability r_k is nothing else but the trainable parameter w_k, i.e., $r_k = w_k$ for all $k = 1, ..., Q$, then the attention weight can be regarded as a result of contamination of the softmax operation, i.e., the attention weights $\alpha(\mathbf{x}, \mathbf{A}_k(\mathbf{x}), \mathbf{w})$ can be represented as follows:

$$\alpha(\mathbf{x}, \mathbf{A}_k(\mathbf{x}), \mathbf{w}) = \\ = (1 - \epsilon) \cdot \text{softmax}\left(-\|\mathbf{x} - \mathbf{A}_k(\mathbf{x})\|^2\right) + \epsilon \cdot w_k, \; k = 1, ..., Q. \quad (15)$$

It can be seen from the above that the softmax function depends only on the distance between $\mathbf{x}$ and $\mathbf{A}_k(\mathbf{x})$ and does not depend on trainable parameters. This implies that it can be regarded as a constant for every $\mathbf{x}$ which falls into the k-th tree. Moreover, the attention weight is linearly depends on trainable parameters $\mathbf{w} =(w_1, ..., w_Q)$. The contamination parameter ϵ can be regarded as a tuning parameter and its optimal value can be selected by using the standard validation procedures. All vectors $\mathbf{w}$ satisfy condition $\mathbf{w} \cdot \mathbf{1}^{\mathrm{T}} = 1$, where $\mathbf{1}$ is the unit vector, and they form the unit simplex $U(1, Q)$. It is important to note that condition (13) is satisfied when we use the ϵ-contamination model because, according to this model, F is a probability distribution.

The property of the attention weights that the softmax operation does not depend on the trainable parameters is very important because it allows us to avoid complex computations for optimizing these parameters. These approaches are united by one idea of the linear approximation of softmax operation [34, 49, 50]. In contrast to the approximation approaches, we propose a quite different model where the softmax operation does not have trainable parameters, and the attention weights inherently depend on these parameters.

4.3. Optimization problem for computing trainable parameters. The next question is how to train the attention parameters $\mathbf{w}$ in order to compute the attention weights. In order to answer this question, we return to the C-index defined in (2) as an important measure for evaluation of the model accuracy and for comparison of different survival models. If to assume that the predicted SF of RSF depends on the trainable attention parameters $\mathbf{w}$, then the C-index should be expressed through these parameters. Then it can be maximized with respect to $\mathbf{w}$. This implies that our first aim is to write C-index as a function of $\mathbf{w}$. Let us rewrite (2) taking into account that the SF of the whole RSF is determined by the attention weights $\alpha\left(\mathbf{x}_i, \mathbf{A}_k(\mathbf{x}_i), \mathbf{w}\right)$ through the Nadaraya-Watson regression (see (14)):

$$C(\mathbf{w}) = \frac{1}{M} \sum_{i:\delta_i=1} \sum_{j:t_i<t_j} \mathbf{1}\left[S(t_i|\mathbf{x}_i, \alpha(\mathbf{x}_i)) - S(t_j|\mathbf{x}_j, \alpha(\mathbf{x}_j)) > 0\right]. \quad (16)$$

Here $\alpha(\mathbf{x}_i)$ is the short notation of the vector of the attention weights $\alpha\left(\mathbf{x}_i, \mathbf{A}_k(\mathbf{x}_i), \mathbf{w}\right)$, $k = 1, ..., Q$; $S(t_i|\mathbf{x}_i, \alpha(\mathbf{x}_i))$ is the ensemble predicted SF depending on the vector $\alpha(\mathbf{x}_i)$ of the attention weights of trees. The C-index depends on $\mathbf{w}$ through the attention weights. We use the short notation $C(\mathbf{w})$ in order to avoid the long expression for C as a function of the SFs and the attention weights.

The survival attention-based model learning means to compute optimal values of the trainable parameters $\mathbf{w}$ of the attention, which maximize the C-index $C(\mathbf{w})$ over values of non-negative weights w_q, $q = 1, ..., Q$, under constraint $\mathbf{w} \cdot \mathbf{1}^{\mathrm{T}} = 1$. In sum, we can write the following optimization problem:

$$\mathbf{w}_{opt} = \max_{\mathbf{w}} C(\mathbf{w}), \tag{17}$$

subject to $\mathbf{w} \cdot \mathbf{1}^{\mathrm{T}} = 1$ or $\mathbf{w} \in U(1, Q)$.

Figure 2 shows a scheme of the training process for computing the attention weights assigned to every tree in RSF. After training the original RSF, vectors $\mathbf{A}_k(\mathbf{x}_s)$ and functions $H_k(t|\mathbf{x}_s)$ are taken for all $k = 1, ..., Q$ and $s = 1, ..., n$. Pairs $(\mathbf{A}_k(\mathbf{x}_s), H_k(t|\mathbf{x}_s))$ allow us to write the optimization problem for maximizing the C-index $C(\mathbf{w})$ as a function of the trainable weights $\mathbf{w}$. Having optimal trainable parameters $\mathbf{w}_{opt}$, we can compute the attention weights $\alpha(\mathbf{x}, \mathbf{A}_k(\mathbf{x}), \mathbf{w})$ and then to find $H(t|\mathbf{x})$ by using (12).

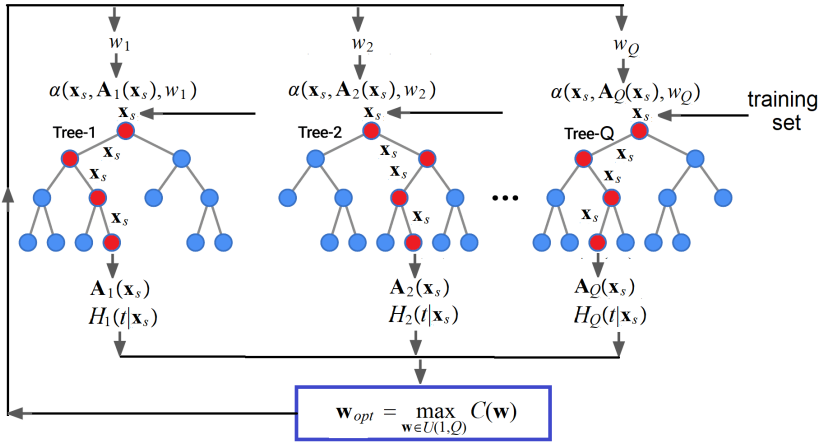


Fig. 2. A scheme of training the attention weights assigned to every tree in the RSF taking an example $\mathbf{x}_s$ from the training set under the condition that the output of the k-th tree is the pair $(\mathbf{A}_k(\mathbf{x}_s), H_k(t|\mathbf{x}_s))$

It should be noted that it is difficult to solve the optimization problem (17) with the indicator functions in the objective function because it is a hard combinatorial problem. Moreover, the ensemble predictive measure is the CHF because it is the weighted sum of the tree CHFs in contrast to the SF which cannot be linearly expressed through the tree weights. However, it has been

shown in [19] that a similar optimization problem can be solved by replacing SFs with CHFs and the indicator functions with hinge loss functions. Therefore, we first show that SFs can be replaced with the CHFs in the objective function. Indeed, it follows from (1) and from the monotonicity of SFs and CHFs that there holds:

$$\begin{aligned}&\mathbf{1}\left[S(t|\mathbf{x}_i) - S(t|\mathbf{x}_j) > 0\right] = \mathbf{1}\left[\ln S(t|\mathbf{x}_i) - \ln S(t|\mathbf{x}_j) > 0\right]\\&= \mathbf{1}\left[H(t|\mathbf{x}_j) - H(t|\mathbf{x}_i) > 0\right].\end{aligned} \tag{18}$$

Hence, the objective function (17) can be rewritten as follows:

$$C(\mathbf{w}) = \frac{1}{M}\sum_{i:\delta_i=1}\sum_{j:t_i<t_j}\mathbf{1}\left[H(t_j|\mathbf{x}_j,\alpha(\mathbf{x}_j)) - H(t_i|\mathbf{x}_i,\alpha(\mathbf{x}_i)) > 0\right]. \tag{19}$$

Now we can use (12) to express $H(t_i|\mathbf{x}_i,\alpha(\mathbf{x}_i))$ in the objective function (19) through the CHFs $H_q(t|\mathbf{x})$, $q = 1, ..., Q$, obtained by every survival tree. Let us denote the set of all possible pairs (i, j) in (19), satisfying condition $\delta_i = 1$ for i and condition $t_i < t_j$ for j, as J. The objective function becomes:

$$\begin{aligned}&C(\mathbf{w}) = \tfrac{1}{M}\textstyle\sum_{(i,j)\in J}\\&1\left[\textstyle\sum_{q=1}^{Q}\left(\alpha_j^{(q)}(\mathbf{w})H_q(t|\mathbf{x}_j) - \alpha_i^{(q)}(\mathbf{w})H_q(t|\mathbf{x}_i)\right) > 0\right],\end{aligned} \tag{20}$$

where $\alpha_i^{(q)}(\mathbf{w}) = \alpha\left(\mathbf{x}_i, \mathbf{A}_q(\mathbf{x}_i), \mathbf{w}\right)$.

Let us return to the definition of $\alpha_i^{(q)}(\mathbf{w})$ by using the Huber's ϵ-contamination model as it is shown in (15). Then the inequality in (20) can be rewritten as:

$$\begin{aligned}&\left((1-\epsilon)\text{softmax}\left(-\left\|\mathbf{x}_j - \mathbf{A}_q(\mathbf{x}_j)\right\|^2\right) + \epsilon w_q\right)H_q(t|\mathbf{x}_j)\\&-\left((1-\epsilon)\text{softmax}\left(-\left\|\mathbf{x}_i - \mathbf{A}_q(\mathbf{x}_i)\right\|^2\right) + \epsilon w_q\right)H_q(t|\mathbf{x}_i)\\&> 0.\end{aligned}$$

Let us introduce the following notations:

$$\begin{aligned}D_q\left(t, \mathbf{x}_i, \mathbf{x}_j, \mathbf{w}\right) &= \alpha_j^{(q)}(\mathbf{w})\cdot H_q(t|\mathbf{x}_j) - \alpha_i^{(q)}(\mathbf{w})\cdot H_q(t|\mathbf{x}_i)\\&= (1-\epsilon)\cdot F_q(t, \mathbf{x}_i, \mathbf{x}_j) + \epsilon\cdot w_q\cdot G_q(t, \mathbf{x}_i, \mathbf{x}_j),\end{aligned} \tag{21}$$

where:

$$F_q(t, \mathbf{x}_i, \mathbf{x}_j) = \text{softmax}\left(-\|\mathbf{x}_j - \mathbf{A}_q(\mathbf{x}_j)\|^2\right) \cdot H_q(t|\mathbf{x}_j) - \text{softmax}\left(-\|\mathbf{x}_i - \mathbf{A}_q(\mathbf{x}_i)\|^2\right) \cdot H_q(t|\mathbf{x}_i), \quad (22)$$

$$G_q(t, \mathbf{x}_i, \mathbf{x}_j) = H_q(t|\mathbf{x}_j) - H_q(t|\mathbf{x}_i). \quad (23)$$

The above notations are introduced to simplify the complex expressions and to show how $D_q\left(t, \mathbf{x}_i, \mathbf{x}_j, \mathbf{w}\right)$ depends on the trainable parameters $\mathbf{w}$. Note that $F_q(\mathbf{x}_i, \mathbf{x}_j)$ and $G_q(\mathbf{x}_i, \mathbf{x}_j)$ do not depend on the trainable parameters $\mathbf{w}$ and are defined only by predictions of trees in the form of CHFs $H_q(t|\mathbf{x})$ and by examples which fall into the corresponding leaf nodes. We also do not include ϵ into $F_q(t, \mathbf{x}_i, \mathbf{x}_j)$ and $G_q(t, \mathbf{x}_i, \mathbf{x}_j)$ in order to highlight the hyperparameter in the optimization problem.

Hence, the following optimization problem can be written:

$$C(\mathbf{w}) = \max_{\mathbf{w} \in U(1,Q)} \frac{1}{M} \sum_{(i,j) \in J} \mathbf{1}\left[\sum_{q=1}^{Q} D_q\left(t, \mathbf{x}_i, \mathbf{x}_j, \mathbf{w}\right) > 0\right], \quad (24)$$

subject to $\mathbf{w} \cdot \mathbf{1}^{\mathrm{T}} = 1$ and: $w_q \geqslant 0$, $q = 1, ..., Q$.

Problem (24) is hard to be solved. Therefore, we propose to replace the indicator function with the hinge loss function $l(x) = \max(0, x)$ similarly to the replacement proposed by Van Belle et al. [51]. This replacement is also used in the support vector machine where the hinge loss function is regarded as a desirable approximation of the indicator function.

By adding the regularization term $R(\mathbf{w})$, the optimization problem can be written as:

$$\min_{\mathbf{w} \in U(1,Q)} \left\{ \sum_{(i,j) \in J} \max\left(0, \sum_{q=1}^{Q} D_q\left(\mathbf{x}_i, \mathbf{x}_j, \mathbf{w}\right)\right) + \lambda R(\mathbf{w}) \right\}. \quad (25)$$

Here λ is a hyper-parameter which controls the strength of the regularization. Let us introduce the variables:

$$\xi_{ij} = \max\left(0, \sum_{q=1}^{Q} D_q\left(\mathbf{x}_i, \mathbf{x}_j, \mathbf{w}\right)\right). \quad (26)$$

If we take the regularization term in the form $R(\mathbf{w}) = \|\mathbf{w}\|^2$, then the optimization problem can be written in the following form:

$$\min_{\mathbf{w}} \left\{ \sum_{(i,j)\in J} \xi_{ij} + \lambda \|\mathbf{w}\|^2 \right\}, \tag{27}$$

subject to $\mathbf{w} \in U(1, Q)$ and:

$$\xi_{ij} \geqslant \sum_{q=1}^{Q} D_q(\mathbf{x}_i, \mathbf{x}_j, \mathbf{w}), \;\; \xi_{ij} \geqslant 0, \;\; \{i, j\} \in J. \tag{28}$$

After substituting (21) into constraints (28), we get:

$$\xi_{ij} \geqslant \sum_{q=1}^{Q} \left((1-\epsilon) \cdot F_q(\mathbf{x}_i, \mathbf{x}_j) + \epsilon \cdot w_q \cdot G_q(\mathbf{x}_i, \mathbf{x}_j)\right),$$
$$\xi_{ij} \geqslant 0, \;\; \{i, j\} \in J. \tag{29}$$

We get the quadratic optimization problem with linear constraints (29) and $\mathbf{w} \in U(1, Q)$. The problem has $\#J + Q$ variables.

In spite of the superficial simplicity of the problem (27) and (29), it has a huge amount of constraints. Therefore, to simplify it, we propose its relaxation in the following way. K constraints are randomly selected from all constraints and are used in the optimization problem. Repeating random selections several times and solving the obtained optimization problems, the obtained trainable parameters $\mathbf{w}$ are averaged and the results are used to compute the attention weights.

Let us consider two important special cases when $\epsilon = 0$ and $\epsilon = 1$. In the first case ($\epsilon = 0$), the subset of training parameters is reduced to the point $F_q(t, \mathbf{x}_i, \mathbf{x}_j)$. This implies that the optimization problem should not be solved. The attention weights are not trainable and have the simple form:

$$\alpha(\mathbf{x}, \mathbf{A}_k(\mathbf{x})) = \text{softmax}\left(-\|\mathbf{x} - \mathbf{A}_k(\mathbf{x})\|^2\right), \; k = 1, ..., Q. \tag{30}$$

This implies that case $\epsilon = 0$ can be regarded as a special case of non-parametric attention mechanism. In the second case ($\epsilon = 1$), the subset of training parameters coincides with the unit simplex $U(1, Q)$ such that the

constraints are reduced to:

$$\xi_{ij} \geqslant \sum_{q=1}^{Q} w_q \left(H_q(t|\mathbf{x}_j) - H_q(t|\mathbf{x}_i) \right), \; \xi_{ij} \geqslant 0. \tag{31}$$

The above coincides with the weighted RSF proposed in [19]. This case does not take into account the distance $\|\mathbf{x} - \mathbf{A}_k(\mathbf{x})\|^2$, i.e., the attention weight does not depend on the feature vector $\mathbf{x}$ and is defined only by some ability of every tree averaged over the training set.

The application of the Huber's ϵ-contamination model significantly simplifies the training and testing phases of Att-RSF because we solve the standard quadratic optimization problem with the convex objective function and linear constraints instead of the complex gradient-based optimization algorithms. At the same time, the complexity of the training phase for computing the optimal trainable parameters of the attention is defined by the complexity of solving the quadratic optimization problem. It depends on the number of selected linear constraints K and on the number of repetitions of the optimization problem solving with restricted numbers of constraints. At the same time, the testing phase is very simple, and it is defined by computing the attention weights $\alpha\left(\mathbf{x}, \mathbf{A}_k(\mathbf{x}), \mathbf{w}\right)$ in (15) under the condition that the attention parameters $\mathbf{w}$ are known. The simplicity of the testing phase allows us to get predictions in case of intensive online data traffic whereas the same cannot be realized in the training phase.

5. Numerical experiments. In order to study how Att-RSF outperforms the original RSF [10] and the weighted RSF [19], we compare Att-RSF with these models. The proposed Att-RSF as well as the original RSF and the weighted RSF are tested on the following real benchmark datasets.

The **Primary Biliary Cirrhosis (PBC) Dataset** consists of information about 418 patients with primary biliary cirrhosis of the liver from the Mayo Clinic trial [52], 257 of whom have censored data. Each patient is described by 17 features such as age, sex, ascites, hepatom, spiders, edema, bili and chol, etc. The dataset can be downloaded via the "randomForestSRC" R package.

The **German Breast Cancer Study Group 2 (GBSG2) Dataset** contains observations of 686 women [53]. Each woman is described by 10 features: age of the patients in years, menopausal status, tumor size, tumor grade, number of positive nodes, hormonal therapy, progesterone receptor, estrogen receptor, recurrence-free survival time, censoring indicator (0 - censored, 1 - event). The dataset can be obtained via the "TH.data" R package.

The **Chronic Myelogenous Leukemia Survival (CML) Dataset** is simulated according to the structure of the data by the German CML Study

Group used in [54]. The dataset consists of 507 observations with 7 features: a factor with 54 levels indicating the study center; a factor with levels trt1, trt2, trt3 indicating the treatment group; sex (0 = female, 1 = male); age in years; risk group (0 = low, 1 = medium, 2 = high); censoring status (FALSE = censored, TRUE = dead); time survival or censoring time in days. The dataset can be obtained via the "multcomp" R package (cml).

The **Bladder Cancer Dataset (BLCD)** [55] (Chapter 21) consists of observations of 86 patients after surgery assigned to placebo or chemotherapy (thiopeta). The endpoint is time to recurrence in months. Data on the number of tumors removed at surgery was also collected. The dataset is available at http://www.stat.rice.edu/~sneeley/STAT553/Datasets/survivaldata.txt.

The **Lupus Nephritis Dataset (LND)** [56] consists of observations of 87 persons with lupus nephritis. followed for 15+ years after an initial renal biopsy (the starting point of follow-up). This data set only contains time to death/censoring, indicator, duration and log(1+duration), where duration is the duration of untreated disease prior to biopsy. The dataset is available at http://www.stat.rice.edu/~sneeley/STAT553/Datasets/survivaldata.txt.

The **Heart Transplant Dataset (HTD)** consists of observations of 69 patients receiving heart transplants [57]. This dataset is available at http://lib.stat.cmu.edu/datasets/stanford.

The **Veterans' Administration Lung Cancer Study (Veteran) Dataset** [57] consists of observations of 137 males with advanced inoperable lung cancer. The patients were randomly assigned to either a standard chemotherapy treatment or a test chemotherapy treatment. Several additional variables were also measured on the patients. The dataset can be obtained via the "survival" R package.

The **Wisconsin Prognostic Breast Cancer (WPBC) dataset** [58] contains records representing follow-up data for one breast cancer case observed by Dr. Wolberg in 1984. WPBC consists of 198 instances having 34 features. The dataset can be obtained via the "TH.data" R package.

The **Gastric Cancer Dataset (GCD)** [59] contains data on the survival of 90 patients (4 features) with locally advanced, non-resectable gastric carcinoma. The dataset can be obtained via the "coxphw" R package.

The **Microarray Breast Cancer Gene expression profiling dataset (MBC)** [60] is for predicting the clinical outcome of breast cancer. It contains 4707 expression values on 78 patients with survival information. The dataset can be obtained via the "randomForestSRC" R package.

Att-RSF is implemented by means of a software in Python. The software implementing the weighted RSF is available at https://github.com/andruekonst/weighted-random-survival-forest.

In order to evaluate the C-index for each dataset, we perform a cross-validation with 100 repetitions by taking 75% of examples from each dataset for training and 25% of examples for testing. Examples for training and testing are randomly selected in each run. Different values for hyperparameters λ and ϵ have been tested, choosing those leading to the best results. The number of survival trees in RSF is 200. The number K of constraints randomly selected from all constraints in the optimization problem for computing optimal trainable parameters $\mathbf{w}$ is 3000, and the number of solutions to the optimization problem with different subsets of constraints is 10.

Table 1 illustrates the C-indices of the original RSF, the weighted RSF (WRSF), and the Att-RSF model obtained for the above datasets under the condition the depth of survival trees in the models is equal to 2. It can be seen from Table 1 that Att-RSF outperforms RSF as well as WRSF for all considered datasets.

Table 1. Comparison of the C-index obtained for RSF, WRSF and Att-RSF by using different datasets

Dataset	RSF	WRSF	Att-RSF
PBC	0.878	0.931	0.943
GBSG2	0.879	0.928	0.959
BLCD	0.876	0.926	0.987
CML	0.869	0.923	0.979
LND	0.875	0.924	0.940
HTD	0.859	0.931	0.931
Veteran	0.870	0.929	0.970
WPBC	0.914	0.943	0.969
GCD	0.518	0.524	0.692
MBC	0.802	0.868	0.914

To formally show the outperformance of the proposed Att-RSF model, we apply the t-test which has been proposed and described by Demsar [61] for testing whether the average difference in the performance of two models, Att-RSF and WRSF, is significantly different from zero. Since we use differences between accuracy measures of Att-RSF and WRSF, then they are compared with 0. The t-statistics is distributed in accordance with the Student distribution with $10-1$ degrees of freedom. The obtained p-value and the 95% confidence interval for the mean 0.046 are $p = 0.0135$ and $[0.012, 0.079]$, respectively. One can see from the t-test that Att-RSF clearly outperforms WRSF due to condition $p < 0.05$. Better results can be carried out for models Att-RSF and RSF. We get the p-value and the 95% confidence interval for the mean 0.094,

which are $p = 1.4 \cdot 10^{-5}$ and $[0.069, 0.120]$, respectively. The second test also demonstrates the outperformance of Att-RSF in comparison with RSF.

The next interesting question is how the C-index depends on the contamination hyperparameter ϵ for different datasets. The C-index as a function of the contamination hyperparameter for the CML dataset is depicted in Figure 3 by the solid line with the circle markers. For comparison purposes, lines with the triangle and square markers correspond to RSF and WRSF, respectively. It can be seen from Figure 3 that Att-RSF provides the best results (the largest C-index) when $\epsilon = 0.75$. However, the optimal choice of the contamination hyperparameter depends on a considered dataset. For example, Figure 4 illustrates the case when the optimal contamination hyperparameter is equal to 1 for the HTD dataset. This implies that WRSF outperforms Att-RSF for all ϵ and has the same C-index for $\epsilon = 1$.

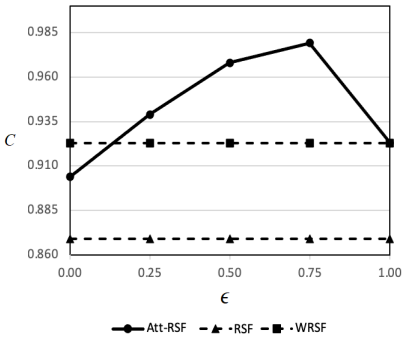


Fig. 3. The dependence of the C-index on the contamination hyperparameter for the CML dataset

6. Conclusion. A new RSF model based on using the attention mechanism has been presented in the paper. The first main idea behind this model is to adapt the Nadaraya-Watson kernel regression to RSF. The second main idea is to apply the Huber's ϵ-contamination model in order to represent the attention weights as the linear function of the trainable attention parameters. Att-RSF has demonstrated outperforming results in comparison with RSF and WRSF for the most considered datasets. This implies that Att-RSF can be an accurate tool for survival analysis especially when tabular data are used for training.

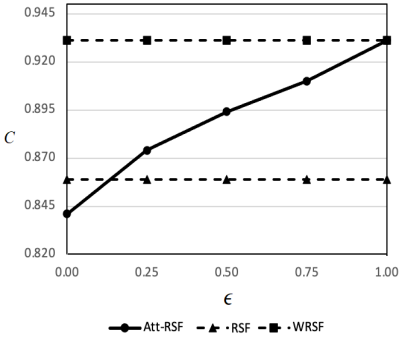


Fig. 4. The dependence of the C-index on the contamination hyperparameter for the HTD dataset

We have considered only one type of the trainable parameters, namely, parameters produced by the Huber's ϵ-contamination model. However, many other types of parameters can be proposed, for example, parameters of the softmax operation. It should be noted that additional parameters lead to more complex optimization problems for computing the attention weights which cannot be solved in a simple way. However, they can be solved by using the gradient-based algorithms and can provide more accurate predictions. The study of other attention mechanisms is an interesting direction for further research.

Another interesting direction for further research is to consider measures of the model accuracy different from the used C-index, for example, the Brier score. The use of other measures may lead to better results.

It should be pointed out that the softmax operation in the attention mechanism defined by the Gaussian kernel in the Nadaraya-Watson kernel regression can be also replaced with other operations if considering different kernels. This consideration and the analysis of the kernel types can be also regarded as an interesting direction for further research.

References

1. Hosmer D., Lemeshow S., May S. Applied Survival Analysis: Regression Modeling of Time to Event Data. — New Jersey : John Wiley & Sons, 2008.
2. DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network / Katzman J., Shaham U., Cloninger A., Bates J., Jiang T., and Kluger Y. // BMC medical research methodology. — 2018. — Vol. 18, no. 24. — P. 1–12.
3. A Deep Active Survival Analysis Approach for Precision Treatment Recommendations: Application of Prostate Cancer / Nezhad M., Sadati N., Yang K., and Zhu D. — 2018. —

Apr. — arXiv:1804.03280v1.
4. Wang P., Li Y., Reddy C. Machine Learning for Survival Analysis: A Survey // ACM Computing Surveys (CSUR). — 2019. — Vol. 51, no. 6. — P. 1–36.
5. Zhao L., Feng D. DNNSurv: Deep Neural Networks for Survival Analysis Using Pseudo Values. — 2020. — Mar. — arXiv:1908.02337v2.
6. Cox D. Regression models and life-tables // Journal of the Royal Statistical Society, Series B (Methodological). — 1972. — Vol. 34, no. 2. — P. 187–220.
7. Tibshirani R. The lasso method for variable selection in the Cox model // Statistics in medicine. — 1997. — Vol. 16, no. 4. — P. 385–395.
8. Survival SVM: a practical scalable algorithm. / Belle V. V., Pelckmans K., Suykens J., and Huffel S. V. // ESANN. — 2008. — P. 89–94.
9. Bou-Hamad I., Larocque D., Ben-Ameur H. A review of survival trees // Statistics Surveys. — 2011. — Vol. 5. — P. 44–71.
10. Ishwaran H., Kogalur U. Random Survival Forests for R // R News. — 2007. — Vol. 7, no. 2. — P. 25–31.
11. Breiman L. Random forests // Machine learning. — 2001. — Vol. 45, no. 1. — P. 5–32.
12. Hu C., Steingrimsson J. Personalized Risk Prediction in Clinical Oncology Research: Applications and Practical Issues Using Survival Trees and Random Forests // Journal of Biopharmaceutical Statistics. — 2018. — Vol. 28, no. 2. — P. 333–349.
13. Relative Risk Forests for Exercise Heart Rate Recovery as a Predictor of Mortality / Ishwaran H., Blackstone E., Pothier C., and Lauer M. // Journal of the American Statistical Association. — 2004. — Vol. 99. — P. 591–600.
14. Mogensen U., Ishwaran H., Gerds T. Evaluating Random Forests for Survival Analysis using Prediction Error Curves // Journal of Statistical Software. — 2012. — Vol. 50, no. 11. — P. 1–23.
15. Random survival forests for dynamic predictions of a time-to-event outcome using a longitudinal biomarker / Pickett K., Suresh K., Campbell K., Davis S., and Juarez-Colunga E. // BMC Medical Research Methodology. — 2021. — Vol. 21, no. 1. — P. 1–14.
16. Schmid M., Wright M., Ziegler A. On the use of Harrell's C for clinical risk prediction via random survival forests // Expert Systems with Applications. — 2016. — Vol. 63. — P. 450–459.
17. Wright M., Dankowski T., Ziegler A. Unbiased split variable selection for random survival forests using maximally selected rank statistics // Statistics in Medicine. — 2017. — Vol. 36, no. 8. — P. 1272–1284.
18. Zhou L., Wang H., Xu Q. Survival forest with partial least squares for high dimensional censored data // Chemometrics and Intelligent Laboratory Systems. — 2018. — Vol. 179. — P. 12–21.
19. A weighted random survival forest / Utkin L., Konstantinov A., Chukanov V., Kots M., Ryabinin M., and Meldo A. // Knowledge-Based Systems. — 2019. — Vol. 177. — P. 136–144.
20. Evaluating the yield of medical tests / Harrell F., Califf R., Pryor D., Lee K., and Rosati R. // Journal of the American Medical Association. — 1982. — Vol. 247. — P. 2543–2546.
21. Utkin L., Konstantinov A. Attention-based Random Forest and Contamination Model // Neural Networks. — 2022. — Vol. 154. — P. 346–359.
22. Huber P. Robust Statistics. — New York : Wiley, 1981.
23. Witten D., Tibshirani R. Survival analysis with high-dimensional covariates // Statistical Methods in Medical Research. — 2010. — Vol. 19, no. 1. — P. 29–51.
24. Zhang H., Lu W. Adaptive Lasso for Cox's proportional hazards model // Biometrika. — 2007. — Vol. 94, no. 3. — P. 691–703.

25. Support vector methods for survival analysis: a comparison between ranking and regression approaches / Belle V. V., Pelckmans K., Huffel S. V., and Suykens J. // Artificial intelligence in medicine. — 2011. — Vol. 53, no. 2. — P. 107–118.
26. Zhu X., Yao J., Huang J. Deep convolutional neural network for survival analysis with pathological images // 2016 IEEE International Conference on Bioinformatics and Biomedicine. — IEEE. — 2016. — P. 544–547.
27. Image-based Survival Analysis for Lung Cancer Patients using CNNs / Haarburger C., Weitz P., Rippel O., and Merhof D. — 2018. — Aug. — arXiv:1808.09679v1.
28. Decision tree for competing risks survival probability in breast cancer study / Ibrahim N., Kudus A., Daud I., and Bakar M. A. // International Journal Of Biological and Medical Research. — 2008. — Vol. 3, no. 1. — P. 25–29.
29. Wang H., Zhou L. Random survival forest with space extensions for censored data // Artificial intelligence in medicine. — 2017. — Vol. 79. — P. 52–61.
30. An attentive survey of attention models / Chaudhari S., Mithal V., Polatkan G., and Ramanath R. — 2019. — Apr. — arXiv:1904.02874.
31. Correia A., Colombini E. Attention, please! A survey of neural attention models in deep learning. — 2021. — Mar. — arXiv:2103.16775.
32. Correia A., Colombini E. Neural Attention Models in Deep Learning: Survey and Taxonomy. — 2021. — Dec. — arXiv:2112.05909.
33. A Survey of Transformers / Lin T., Wang Y., Liu X., and Qiu X. — 2021. — Jul. — arXiv:2106.04554.
34. Random Features for Kernel Approximation: A Survey on Algorithms, Theory, and Beyond / Liu F., Huang X., Chen Y., and Suykens J. — 2021. — Jul. — arXiv:2004.11154v5.
35. Niu Z., Zhong G., Yu H. A review on the attention mechanism of deep learning // Neurocomputing. — 2021. — Vol. 452. — P. 48–62.
36. Ronao C., Cho S.-B. Random Forests with Weighted Voting for Anomalous Query Access Detection in Relational Databases // Artificial Intelligence and Soft Computing. ICAISC 2015. — Cham : Springer. — 2015. — Vol. 9120 of Lecture Notes in Computer Science. — P. 36–48.
37. Xuan S., Liu G., Li Z. Refined Weighted Random Forest and Its Application to Credit Card Fraud Detection // Computational Data and Social Networks. — Cham : Springer International Publishing. — 2018. — P. 343–355.
38. Zhang X., Wang M. Weighted Random Forest Algorithm Based on Bayesian Algorithm // Journal of Physics: Conference Series. — IOP Publishing. — 2021. — Vol. 1924. — P. 1–6.
39. Weighted vote for trees aggregation in Random Forest / Daho M., Settouti N., Lazouni M., and Chikh M. // 2014 International Conference on Multimedia Computing and Systems (ICMCS). — IEEE. — 2014. — April. — P. 438–443.
40. Utkin L., Kovalev M., Meldo A. A deep forest classifier with weights of class probability distribution subsets // Knowledge-Based Systems. — 2019. — Vol. 173. — P. 15–27.
41. Utkin L., Kovalev M., Coolen F. Imprecise weighted extensions of random forests for classification and regression // Applied Soft Computing. — 2020. — Vol. 92, no. Article 106324. — P. 1–14.
42. Development and validation of a prognostic model for survival time data: application to prognosis of HIV positive patients treated with antiretroviral therapy / May M., Royston P., Egger M., Justice A., and Sterne J. // Statistics in Medicine. — 2004. — Vol. 23. — P. 2375–2398.
43. Random Survival Forests / Ishwaran H., Kogalur U., Blackstone E., and Lauer M. // Annals of Applied Statistics. — 2008. — Vol. 2. — P. 841–860.

44. Nadaraya E. On estimating regression // Theory of Probability & Its Applications. — 1964. — Vol. 9, no. 1. — P. 141–142.
45. Watson G. Smooth regression analysis // Sankhya: The Indian Journal of Statistics, Series A. — 1964. — P. 359–372.
46. Bahdanau D., Cho K., Bengio Y. Neural machine translation by jointly learning to align and translate. — 2014. — Sep. — arXiv:1409.0473.
47. Luong T., Pham H., Manning C. Effective approaches to attention-based neural machine translation // Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. — The Association for Computational Linguistics. — 2015. — P. 1412–1421.
48. Attention is all you need / Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser L., and Polosukhin I. // Advances in Neural Information Processing Systems. — 2017. — P. 5998–6008.
49. Rethinking Attention with Performers / Choromanski K., Likhosherstov V., Dohan D., Song X., Gane A., Sarlos T., Hawkins P., Davis J., Mohiuddin A., Kaiser L., Belanger D., Colwell L., and Weller A. // 2021 International Conference on Learning Representations. — 2021.
50. Schlag I., Irie K., Schmidhuber J. Linear transformers are secretly fast weight programmers // International Conference on Machine Learning 2021. — PMLR. — 2021. — P. 9355–9366.
51. Support vector machines for survival analysis / Belle V. V., Pelckmans K., Suykens J., and Huffel S. V. // Proceedings of the Third International Conference on Computational Intelligence in Medicine and Healthcare (CIMED2007). — 2007. — P. 1–8.
52. Fleming T., Harrington D. Counting processes and survival aalysis. — Hoboken, NJ, USA : John Wiley & Sons, 1991.
53. Sauerbrei W., Royston P. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials // Journal of the Royal Statistics Society Series A. — 1999. — Vol. 162, no. 1. — P. 71–94.
54. Randomized comparison of interferon-alpha with busulfan and hydroxyurea in chronic myelogenous leukemia. The German CML study group / Hehlmann R., Heimpel H., Hasford J., Kolb H., Pralle H., Hossfeld D., Queisser W., Loeffler H., Hochhaus A., and Heinze B. // Blood. — 1994. — Vol. 84, no. 12. — P. 4064–4077.
55. Pagano M., Gauvreau K. Principles of biostatistics. — Pacific Grove, CA : Duxbury, 2000.
56. Abrahamowicz M., MacKenzie T., Esdaile J. Time-dependent hazard ratio: modelling and hypothesis testing with application in lupus nephritis // JASA. — 1996. — Vol. 91. — P. 1432–1439.
57. Kalbfleisch J., Prentice R. The Statistical Analysis of Failure Time Data. — New York : John Wiley and Sons, 1980.
58. Street W., Mangasarian O., Wolberg W. An inductive learning approach to prognostic prediction // Proceedings of the Twelfth International Conference on Machine Learning. — San Francisco : Morgan Kaufmann. — 1995. — P. 522–530.
59. Stablein D., Carter J., Novak J. Analysis of Survival Data with Nonproportional Hazard Functions // Controlled Clinical Trials. — 1981. — Vol. 2. — P. 149–159.
60. Gene expression profiling predicts clinical outcome of breast cancer / Veer L. V., Dai H., Vijver M. V. D., He Y., Hart A., Mao M., Peterse H., Kooy K. V. D., Marton M., Witteveen A., and Schreiber G. // Nature. — 2002. — Vol. 12. — P. 530–536.
61. Demsar J. Statistical comparisons of classifiers over multiple data sets // Journal of Machine Learning Research. — 2006. — Vol. 7. — P. 1–30.

Utkin Lev — Ph.D., Dr.Sci., Professor, Head of the institute, Institute of computer science and technology, Peter the Great St.Petersburg Polytechnic University. Research interests: machine learning, imprecise probability theory, decision making. The number of publications — 300. lev.utkin@gmail.com; 29, Politekhnicheskaya St., 195251, St. Petersburg, Russia; office phone: +7(812)775-0530.

Konstantinov Andrei — Ph.D., Graduate student, Peter the Great St. Petersburg Polytechnic University. Research interests: machine learning, computer vision and image processing. The number of publications — 15. andrue.konst@gmail.com; 29, Politekhnicheskaya St., 195251, St. Petersburg, Russia; office phone: +7(911)954-5565.

Acknowledgements. This work is supported by the Russian Science Foundation under grant 21-11-00116.

УДК 004.8 **DOI** 10.15622/ia.21.5.1

Л.В. Уткин, А.В. Константинов,

СЛУЧАЙНЫЙ ЛЕС ВЫЖИВАЕМОСТИ И РЕГРЕССИЯ НАДАРАЯ-УОТСОНА

Уткин Л.В., Константинов А.В. **Случайный лес выживаемости и регрессия Надарая-Уотсона.**

Аннотация. В статье представлен случайный лес выживаемости на основе модели внимания (Att-RSF). Первая идея, лежащая в основе леса, состоит в том, чтобы адаптировать ядерную регрессию Надарая-Уотсона к случайному лесу выживаемости таким образом, чтобы веса регрессии или ядра можно было рассматривать как обучаемые веса внимания при важном условии, что предсказания случайного леса выживаемости представлены в виде функций времени, например, функции выживания или кумулятивной функции риска. Каждый обучаемый вес, присвоенный дереву и примеру из обучающей или тестовой выборки, определяется двумя факторами: способностью соответствующего дерева предсказывать и особенностью примера, попадающего в лист дерева. Вторая идея Att-RSF состоит в том, чтобы применить модель загрязнения Хьюбера для представления весов внимания как линейной функции обучаемых параметров внимания. С-индекс Харрелла (индекс конкордации) как показатель качества предсказания случайного леса выживаемости используется при формировании функции потерь для обучения весов внимания. Использование С-индекса вместе с моделью загрязнения приводит к стандартной задаче квадратичной оптимизации для вычисления весов, которая имеет целый ряд простых алгоритмов решения. Численные эксперименты с реальными наборами данных, содержащими данные о выживаемости, иллюстрируют предлагаемую модель Att-RSF.

Ключевые слова: машинное обучение, случайный лес выживаемости, функция выживаемости, С-индекс, кумулятивная функция риска, модель внимания, модель засорения Хьюбера.

Литература

1. Hosmer D., Lemeshow S., May S. Applied Survival Analysis: Regression Modeling of Time to Event Data. — New Jersey : John Wiley & Sons, 2008.
2. DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network / Katzman J., Shaham U., Cloninger A., Bates J., Jiang T., and Kluger Y. // BMC medical research methodology. — 2018. — Vol. 18, no. 24. — P. 1–12.
3. A Deep Active Survival Analysis Approach for Precision Treatment Recommendations: Application of Prostate Cancer / Nezhad M., Sadati N., Yang K., and Zhu D. — 2018. — Apr. — arXiv:1804.03280v1.
4. Wang P., Li Y., Reddy C. Machine Learning for Survival Analysis: A Survey // ACM Computing Surveys (CSUR). — 2019. — Vol. 51, no. 6. — P. 1–36.
5. Zhao L., Feng D. DNNSurv: Deep Neural Networks for Survival Analysis Using Pseudo Values. — 2020. — Mar. — arXiv:1908.02337v2.
6. Cox D. Regression models and life-tables // Journal of the Royal Statistical Society, Series B (Methodological). — 1972. — Vol. 34, no. 2. — P. 187–220.
7. Tibshirani R. The lasso method for variable selection in the Cox model // Statistics in medicine. — 1997. — Vol. 16, no. 4. — P. 385–395.

8. Survival SVM: a practical scalable algorithm. / Belle V. V., Pelckmans K., Suykens J., and Huffel S. V. // ESANN. — 2008. — P. 89–94.
9. Bou-Hamad I., Larocque D., Ben-Ameur H. A review of survival trees // Statistics Surveys. — 2011. — Vol. 5. — P. 44–71.
10. Ishwaran H., Kogalur U. Random Survival Forests for R // R News. — 2007. — Vol. 7, no. 2. — P. 25–31.
11. Breiman L. Random forests // Machine learning. — 2001. — Vol. 45, no. 1. — P. 5–32.
12. Hu C., Steingrimsson J. Personalized Risk Prediction in Clinical Oncology Research: Applications and Practical Issues Using Survival Trees and Random Forests // Journal of Biopharmaceutical Statistics. — 2018. — Vol. 28, no. 2. — P. 333–349.
13. Relative Risk Forests for Exercise Heart Rate Recovery as a Predictor of Mortality / Ishwaran H., Blackstone E., Pothier C., and Lauer M. // Journal of the American Statistical Association. — 2004. — Vol. 99. — P. 591–600.
14. Mogensen U., Ishwaran H., Gerds T. Evaluating Random Forests for Survival Analysis using Prediction Error Curves // Journal of Statistical Software. — 2012. — Vol. 50, no. 11. — P. 1–23.
15. Random survival forests for dynamic predictions of a time-to-event outcome using a longitudinal biomarker / Pickett K., Suresh K., Campbell K., Davis S., and Juarez-Colunga E. // BMC Medical Research Methodology. — 2021. — Vol. 21, no. 1. — P. 1–14.
16. Schmid M., Wright M., Ziegler A. On the use of Harrell's C for clinical risk prediction via random survival forests // Expert Systems with Applications. — 2016. — Vol. 63. — P. 450–459.
17. Wright M., Dankowski T., Ziegler A. Unbiased split variable selection for random survival forests using maximally selected rank statistics // Statistics in Medicine. — 2017. — Vol. 36, no. 8. — P. 1272–1284.
18. Zhou L., Wang H., Xu Q. Survival forest with partial least squares for high dimensional censored data // Chemometrics and Intelligent Laboratory Systems. — 2018. — Vol. 179. — P. 12–21.
19. A weighted random survival forest / Utkin L., Konstantinov A., Chukanov V., Kots M., Ryabinin M., and Meldo A. // Knowledge-Based Systems. — 2019. — Vol. 177. — P. 136–144.
20. Evaluating the yield of medical tests / Harrell F., Califf R., Pryor D., Lee K., and Rosati R. // Journal of the American Medical Association. — 1982. — Vol. 247. — P. 2543–2546.
21. Utkin L., Konstantinov A. Attention-based Random Forest and Contamination Model // Neural Networks. — 2022. — Vol. 154. — P. 346–359.
22. Huber P. Robust Statistics. — New York : Wiley, 1981.
23. Witten D., Tibshirani R. Survival analysis with high-dimensional covariates // Statistical Methods in Medical Research. — 2010. — Vol. 19, no. 1. — P. 29–51.
24. Zhang H., Lu W. Adaptive Lasso for Cox's proportional hazards model // Biometrika. — 2007. — Vol. 94, no. 3. — P. 691–703.
25. Support vector methods for survival analysis: a comparison between ranking and regression approaches / Belle V. V., Pelckmans K., Huffel S. V., and Suykens J. // Artificial intelligence in medicine. — 2011. — Vol. 53, no. 2. — P. 107–118.
26. Zhu X., Yao J., Huang J. Deep convolutional neural network for survival analysis with pathological images // 2016 IEEE International Conference on Bioinformatics and Biomedicine. — IEEE. — 2016. — P. 544–547.

27. Image-based Survival Analysis for Lung Cancer Patients using CNNs / Haarburger C., Weitz P., Rippel O., and Merhof D. — 2018. — Aug. — arXiv:1808.09679v1.
28. Decision tree for competing risks survival probability in breast cancer study / Ibrahim N., Kudus A., Daud I., and Bakar M. A. // International Journal Of Biological and Medical Research. — 2008. — Vol. 3, no. 1. — P. 25–29.
29. Wang H., Zhou L. Random survival forest with space extensions for censored data // Artificial intelligence in medicine. — 2017. — Vol. 79. — P. 52–61.
30. An attentive survey of attention models / Chaudhari S., Mithal V., Polatkan G., and Ramanath R. — 2019. — Apr. — arXiv:1904.02874.
31. Correia A., Colombini E. Attention, please! A survey of neural attention models in deep learning. — 2021. — Mar. — arXiv:2103.16775.
32. Correia A., Colombini E. Neural Attention Models in Deep Learning: Survey and Taxonomy. — 2021. — Dec. — arXiv:2112.05909.
33. A Survey of Transformers / Lin T., Wang Y., Liu X., and Qiu X. — 2021. — Jul. — arXiv:2106.04554.
34. Random Features for Kernel Approximation: A Survey on Algorithms, Theory, and Beyond / Liu F., Huang X., Chen Y., and Suykens J. — 2021. — Jul. — arXiv:2004.11154v5.
35. Niu Z., Zhong G., Yu H. A review on the attention mechanism of deep learning // Neurocomputing. — 2021. — Vol. 452. — P. 48–62.
36. Ronao C., Cho S.-B. Random Forests with Weighted Voting for Anomalous Query Access Detection in Relational Databases // Artificial Intelligence and Soft Computing. ICAISC 2015. — Cham : Springer. — 2015. — Vol. 9120 of Lecture Notes in Computer Science. — P. 36–48.
37. Xuan S., Liu G., Li Z. Refined Weighted Random Forest and Its Application to Credit Card Fraud Detection // Computational Data and Social Networks. — Cham : Springer International Publishing. — 2018. — P. 343–355.
38. Zhang X., Wang M. Weighted Random Forest Algorithm Based on Bayesian Algorithm // Journal of Physics: Conference Series. — IOP Publishing. — 2021. — Vol. 1924. — P. 1–6.
39. Weighted vote for trees aggregation in Random Forest / Daho M., Settouti N., Lazouni M., and Chikh M. // 2014 International Conference on Multimedia Computing and Systems (ICMCS). — IEEE. — 2014. — April. — P. 438–443.
40. Utkin L., Kovalev M., Meldo A. A deep forest classifier with weights of class probability distribution subsets // Knowledge-Based Systems. — 2019. — Vol. 173. — P. 15–27.
41. Utkin L., Kovalev M., Coolen F. Imprecise weighted extensions of random forests for classification and regression // Applied Soft Computing. — 2020. — Vol. 92, no. Article 106324. — P. 1–14.
42. Development and validation of a prognostic model for survival time data: application to prognosis of HIV positive patients treated with antiretroviral therapy / May M., Royston P., Egger M., Justice A., and Sterne J. // Statistics in Medicine. — 2004. — Vol. 23. — P. 2375–2398.
43. Random Survival Forests / Ishwaran H., Kogalur U., Blackstone E., and Lauer M. // Annals of Applied Statistics. — 2008. — Vol. 2. — P. 841–860.
44. Nadaraya E. On estimating regression // Theory of Probability & Its Applications. — 1964. — Vol. 9, no. 1. — P. 141–142.
45. Watson G. Smooth regression analysis // Sankhya: The Indian Journal of Statistics, Series A. — 1964. — P. 359–372.

46. Bahdanau D., Cho K., Bengio Y. Neural machine translation by jointly learning to align and translate. — 2014. — Sep. — arXiv:1409.0473.
47. Luong T., Pham H., Manning C. Effective approaches to attention-based neural machine translation // Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. — The Association for Computational Linguistics. — 2015. — P. 1412–1421.
48. Attention is all you need / Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser L., and Polosukhin I. // Advances in Neural Information Processing Systems. — 2017. — P. 5998–6008.
49. Rethinking Attention with Performers / Choromanski K., Likhosherstov V., Dohan D., Song X., Gane A., Sarlos T., Hawkins P., Davis J., Mohiuddin A., Kaiser L., Belanger D., Colwell L., and Weller A. // 2021 International Conference on Learning Representations. — 2021.
50. Schlag I., Irie K., Schmidhuber J. Linear transformers are secretly fast weight programmers // International Conference on Machine Learning 2021. — PMLR. — 2021. — P. 9355–9366.
51. Support vector machines for survival analysis / Belle V. V., Pelckmans K., Suykens J., and Huffel S. V. // Proceedings of the Third International Conference on Computational Intelligence in Medicine and Healthcare (CIMED2007). — 2007. — P. 1–8.
52. Fleming T., Harrington D. Counting processes and survival aalysis. — Hoboken, NJ, USA : John Wiley & Sons, 1991.
53. Sauerbrei W., Royston P. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials // Journal of the Royal Statistics Society Series A. — 1999. — Vol. 162, no. 1. — P. 71–94.
54. Randomized comparison of interferon-alpha with busulfan and hydroxyurea in chronic myelogenous leukemia. The German CML study group / Hehlmann R., Heimpel H., Hasford J., Kolb H., Pralle H., Hossfeld D., Queisser W., Loeffler H., Hochhaus A., and Heinze B. // Blood. — 1994. — Vol. 84, no. 12. — P. 4064–4077.
55. Pagano M., Gauvreau K. Principles of biostatistics. — Pacific Grove, CA : Duxbury, 2000.
56. Abrahamowicz M., MacKenzie T., Esdaile J. Time-dependent hazard ratio: modelling and hypothesis testing with application in lupus nephritis // JASA. — 1996. — Vol. 91. — P. 1432–1439.
57. Kalbfleisch J., Prentice R. The Statistical Analysis of Failure Time Data. — New York : John Wiley and Sons, 1980.
58. Street W., Mangasarian O., Wolberg W. An inductive learning approach to prognostic prediction // Proceedings of the Twelfth International Conference on Machine Learning. — San Francisco : Morgan Kaufmann. — 1995. — P. 522–530.
59. Stablein D., Carter J., Novak J. Analysis of Survival Data with Nonproportional Hazard Functions // Controlled Clinical Trials. — 1981. — Vol. 2. — P. 149–159.
60. Gene expression profiling predicts clinical outcome of breast cancer / Veer L. V., Dai H., Vijver M. V. D., He Y., Hart A., Mao M., Peterse H., Kooy K. V. D., Marton M., Witteveen A., and Schreiber G. // Nature. — 2002. — Vol. 12. — P. 530–536.
61. Demsar J. Statistical comparisons of classifiers over multiple data sets // Journal of Machine Learning Research. — 2006. — Vol. 7. — P. 1–30.

Уткин Лев Владимирович — д-р техн. наук, профессор, директор, институт компьютерных наук и технологий, Санкт-Петербургский политехнический университет Петра Великого. Область научных интересов: машинное обучение, неточная теория вероятностей, принятие решений. Число научных публикаций — 300. lev.utkin@gmail.com; улица Политехническая, 29, 195251, Санкт-Петербург, Россия; р.т.: +7(812)775-0530.

Константинов Андрей Владимирович — канд. техн. наук, аспирант, Санкт-Петербургский политехнический университет Петра Великого. Область научных интересов: машинное обучение, компьютерное зрение и обработка изображений. Число научных публикаций — 15. andrue.konst@gmail.com; Политехническая улица, 29, 195251, Санкт-Петербург, Россия; р.т.: +7(911)954-5565.

Поддержка исследований. Работа выполнена при поддержке Российского научного фонда в рамках гранта 21-11-00116.

УДК 004.91 **DOI** 10.15622/ia.21.5.2

Н.И. Юсупова, Г.Р. Воробьева, Р.Х. Зулкарнеев

ПОДХОД К ИНТЕГРАЦИИ РАЗНОРОДНЫХ ИСТОЧНИКОВ МЕДИЦИНСКИХ ДАННЫХ НА ОСНОВЕ МИКРОСЕРВИСНОЙ АРХИТЕКТУРЫ

Юсупова Н.И., Воробьева Г.Р., Зулкарнеев Р.Х. **Подход к интеграции разнородных источников медицинских данных на основе микросервисной архитектуры.**

Аннотация. Задача обработки медицинской информации в настоящее время в нашей стране и за рубежом решается посредством разнородных медицинских информационных систем, преимущественно локального и регионального уровней. Постоянно возрастающий объем и сложность накапливаемой информации наряду с необходимостью обеспечения прозрачности и преемственности обработки медицинских данных (в частности, к примеру, по бронхолегочным заболеваниям) в различных организациях требует разработки нового подхода к интеграции их разнородных источников. При этом важным требованием к решению поставленной задачи является возможность веб-ориентированной реализации, что позволит сделать соответствующие приложения доступными широкому кругу пользователей без высоких требований к их аппаратно-программным возможностям. В работе рассматривается подход к интеграции разнородных источников медицинской информации, который основан на принципах построения микросервисных веб-архитектур. Каждый модуль обработки данных может быть использован независимо от других программных модулей, предоставляя универсальную точку входа и результирующий набор данных в соответствии с принятой схемой данных. Последовательное выполнение этапов обработки предполагает передачу управления соответствующим программным модулям в фоновом режиме по принципу Cron. В схеме декларируется два вида схем данных – локальная (от медицинских информационных систем) и глобальная (для единой системы хранения), между которыми предусмотрены соответствующие параметры отображения по принципу построения XSLT-таблиц. Важной отличительной особенностью предлагаемого подхода представляется модернизация системы хранения медицинской информации, заключающейся в создании зеркальных копий основного сервера с периодической репликацией соответствующей информации. При этом взаимодействие между клиентами и серверами хранилищ данных осуществляется по типу систем доставки контента с созданием сеанса соединения между конечными точками по принципу ближайшего расстояния между ними, рассчитанного по формуле гаверсинусов. Проведенные вычислительные эксперименты над тестовыми данными по бронхолегочным заболеваниям показали эффективность предложенного подхода как для загрузки данных, так и для их получения отдельными пользователями и программными системами. В целом показатель реактивности соответствующим веб-ориентированных приложений был улучшен на 40% при стабильном соединении.

Ключевые слова: медицинские данные, хранилища данных, интеграция данных, веб-приложения, система доставки контента, микросервисная архитектура.

1. Введение. В настоящее время биомедицина и здравоохранение все чаще классифицируются как области интенсивного использования данных (так называемые «data-intensive field»), где специалисты постоянно сталкиваются с необходимостью обработки и анализа большого объема разнородной информации. При

этом разнородная медицинская информация представляет собой данные различного формата и назначения, получаемые из различных источников по различным протоколам. Для конкретизации решаемой задачи здесь и далее разнородные медицинские данные представляют собой медико-клиническую информацию, а также административно-социальные данные задействованных в соответствующих медицинских процессах объектов, субъектов и ресурсов. Медико-экономическая информация в данной работе не рассматривается и является предметом дальнейших исследований авторов.

В качестве источников медицинских данных выступают клинические данные для поддержки принятия решений различной специализации в виде стандартизированных данных из электронных историй болезни, данные с датчиков мониторинга и различных записывающих лабораторных устройств, данные неотложной помощи, лекарственных препаратах и пр. Здесь же фигурируют и административно-социальные данные пациентов, врачей, медицинских учреждений. При этом одной из основных задач развития технологий обработки таких данных является соблюдение преемственности в лечении пациентов, обеспечение прозрачности клинико-диагностических процессов, повышение эффективности лечения за счет оперативного доступа ко всей необходимой информации.

На сегодняшний день достигнут значительный прогресс в области обработки и хранения медицинской информации [1]. Однако интеграция и управление постоянно растущими объемами разнородных медицинских данных часто сопряжены с рядом проблем, связанных прежде всего с трансформацией непосредственно самой отрасли, обусловленной накоплением массивов цифровых данных, их аналитической обработкой и применением полученных результатов в процессе принятия соответствующих решений. При этом акцент смещается на преобразование крупномасштабных разнородных медицинских данных в информационные продукты и стратегии, что является, в частности, одной из ключевых правительственных инициатив в области персонализованной медицины. Так, с 2011 по 2018 год реализован первый проект по созданию единой государственной информационной системы в сфере здравоохранения, с 2018 года разрабатывается второй проект по созданию единого цифрового контура.

Другой областью повышенного интереса к интеграции и хранению данных в медицине является разработка стратегий для эффективного анализа постоянно растущего массива электронных медицинских карт (ЭМК). При этом одна из основных проблем при

извлечении знаний из электронных медицинских карт заключается в том, что электронные медицинские записи представляют собой крайне разнородные источники данных со сложным массивом количественных, качественных и транзакционных данных. Разрозненные типы данных включают коды МКБ (используемые в основном для ценообразования и оплаты больничных процедур), биохимические и лабораторные тесты, клинические (текстовые) заметки, исторические архивы медицинских вмешательств, методов лечения и даже доставки фармацевтических препаратов. Эти источники данных часто собираются десятками людей (иногда с субъективными критериями) для каждого случая. Следовательно, данные ЭМК довольно сложно анализировать, особенно если рассматривать институциональные и даже многоцентровые уровни с большим количеством пациентов.

Клинические и биомедицинские данные бывают самых разных форматов; они зачастую сложны, неоднородны, слабо аннотированы и, как правило, не структурированы, а объем данных постоянно растет. Каждая из этих проблем: размер, разнообразие, форматирование, сложность, неоднородность, плохие аннотации и отсутствие структуры создают проблемы для эффективного применения в процессе принятия решений.

Проблема усугубляется отсутствием единого подхода к организации медицинских данных, несогласованностью и разнородностью медицинских информационных систем на уровне как отдельных медицинских учреждений, так и территориальных областей. Здесь же важно отметить проблему низкой сопряженности сведений о различных случаях лечения одного и того же человека в конкретной медицинской организации (в частности, при лечении бронхолегочных заболеваний). В подавляющем большинстве случаев в ЭМК дискретно фиксируется только один случай лечения пациента (от поступления до выписки), однако непосредственно интеграция данных предполагает создание единой системы, обеспечивающий доступ ко всем передвижениям пациента в рамках как отдельной организации, так и страны в целом.

Для решения обозначенных проблем требуется подход к интеграции данных, что, безусловно, сопряжено с решением вопросов взаимодействия медицинских информационных систем между собой. Очевидным решением здесь является использование единого информационного хранилища медицинских данных (в частности, по бронхолегочным заболеваниям), обеспечивающего возможность

оперативного импортирования и экспортирования разнородных данных в / из любой информационной системы.

При этом необходимо обеспечить интеграцию источников данных, которые, в свою очередь, существенно различаются по типу (видео, звук, текст, графика, компьютерная анимация и т.д.), назначению, способу записи и возможности удаленной обработки. Также важно учитывать, что зачастую при хранении результатов клинических обследований на различном медицинском оборудовании используются разнообразные форматы данных, нередко специфичные для определенного типа оборудования и не доступные для обработки общедоступными средствами.

2. Состояние вопроса. В настоящее время известен целый ряд российских и зарубежных научно-исследовательских работ, посвященных обозначенному вопросу. Однако наибольшее внимание при этом уделяется созданию такого формата представления медицинских данных, который позволял бы независимо от используемых аппаратно-программных платформ импортировать и экспортировать информацию, а также осуществлять ее обработку, анализ и графическую интерпретацию [2, 3]. Так, к примеру, известны результаты в области разработки единого формата представления медицинской информации, основанного на онтологическом подходе к представлению объектов и субъектов медицинского обслуживания [3-5]. На сегодняшний день формат доступен не только на российском, но и на международном уровне. Однако в настоящий момент вопрос его внедрения в существующие медицинские информационные системы остается открытым.

Другое направление исследований связано с разработкой непосредственно подхода к интеграции без привязки к формату представления данных. Так, согласно некоторым научным школам в данной области, интеграция должна быть проведена поэтапно, начиная от уровня отдельно взятого пациента и данных медицинских приборов и аппаратов до представления сводных данных в единую государственную информационную систему в сфере здравоохранения [6]. В большинстве исследований по интеграции в основу положен онтологический подход, а данные рекомендуется представлять в общепринятом формате HL7 FHIR [7], а также с применением технологий облачных сервисов для реализации хранения больших объемов медицинских данных из гетерогенных источников.

Для обобщения мирового опыта в области интеграции разнородных источников медицинских данных были проанализированы результаты известных научных исследований по

этому вопросу. Результаты сравнительного анализа нескольких распространенных подходов к интеграции медицинской информации приведены в таблице 1.

Обобщая представленные в таблице 1 данные, можно резюмировать следующее. Большинство известных подходов являются узкоспециализированными и представляют собой решение задач обработки и анализа данных по одному или нескольким заболеваниям. При этом реализующая их программная архитектура, как правило, слабо масштабируется – введение новых функций и ресурсов требует существенных трудовых и вычислительных затрат. В большинстве случаев предлагаемые решения локальны и ориентированы на реализацию в пределах одного медицинского учреждения, что сопряжено с соответствующей физической организацией систем долговременного хранения медицинской информации. Немаловажным представляется тот факт, что практически каждое известное решение предполагает применение собственного формата медицинских данных, полученных в результате интеграции гетерогенных источников. При этом ни один из форматов не коррелирует в достаточной мере с известными стандартами представления медицинской информации (например, HL7 FHIR), что существенно затрудняет интеграцию соответствующих решений со сторонними информационными системами. Кроме того, в подавляющем большинстве случаев имеют место десктопные приложения, что, в свою очередь, сопряжено с высокими требованиями к вычислительным ресурсам конечного пользователя. Выполнение запросов к большим объемам медицинской информации так же не оптимизировано, что приводит к низкой реактивности соответствующих приложений.

В этой связи возникает актуальная задача разработки подхода к интеграции данных из разнородных медицинских информационных систем, обеспечивающего возможность веб-ориентированного оперативного обращения с различных интерфейсов взаимодействия для последующей обработки полученных результирующих данных с выгрузкой в установленные структуры, к примеру, в виде электронной медицинской карты пациента.

Таблица 1. Сравнительный анализ подходов к интеграции разнородной медицинской информации

Название, организация, назначение	Преимущества	Недостатки
Масштабируемые информационные системы [8] *Совместные исследования университетов США (Albert Einstein College of Medicine, Beth Israel Deaconess Medical Center, Tufts University School of Medicine и др.)* Цель: создание инфраструктуры оцифрованных медицинских данных по всему миру	– открытый код – веб-интерфейс для доступа к данным – положительные отзывы пользователей с точки зрения эргономики приложения	– в настоящее время система предназначена для медицинской информации только по интенсивной терапии – используется собственный формат данных (нет поддержки стандарта FHIL) – низкая реактивность приложения
Цифровая платформа для анализа и визуализации эпидемиологических данных [9] *ICMR-National Institute of Malaria Research, Индия* Цель: разработка универсального интерфейса, который позволит проводить быстрый и интерактивный анализ эпидемиологических данных по малярии	– простой и понятный интерфейс – единый доступ к большим эпидемиологическим данным, накопленным в рассматриваемом регионе	– только настольные приложения – высокие требования к вычислительным мощностям пользователей – система предназначена для обработки региональных данных по малярии – на данный момент нет возможности масштабирования

Продолжение Таблицы 1

Глобальный хаб для медицинских данных [10] *Clalit Research Institute, Израиль* Цель: разработка единого информационного пространства медицинских данных для реализации программ персонализированной медицины	– наличие единого информационного пространства в рамках региона – использование облачных технологий для хранения данных – веб-ориентированный интерфейс, доступный широкому кругу пользователей без высоких требований к вычислительным ресурсам	– только региональный масштаб – низкая реактивность – сложность масштабирования
Интеграция данных из разнородных источников [11] *HealthCore Inc, США* Цель: описание и оценка процесса интеграции данных из нескольких дополнительных источников для проведения исследований результатов лечения пациентов с раком легкого	– наличие единого информационного пространства в рамках региона	– рассматриваются только 4 источника данных, масштабирование не предусмотрено на настоящий момент – данные представлены в собственном, отличном от стандартного, формате (негативно сказывается на интероперабельности)
Интеграция систем клинических и лабораторных исследований [12] *Институт медицинской информатики, Германия* Цель: создание единой онтологии для	– в основе лежит онтологический подход, позволяющий выявить связи (в том числе неявные) между ресурсами, объектами и субъектами	– в настоящее время существует только десктопная версия, сопряженная с высокими требованиями к вычислительным ресурсам пользователя – не предусмотрено масштабирование системы

Продолжение Таблицы 1

электронных медицинских карт университетской больницы Эрлангена		– требуется стандартизации целевой онтологии и внутренней модели данных, а также интеграции дополнительных отображений в стандартизированные терминологии
Технологии интеграции и визуализации данных (DIVT) [13] *Institute of Biomaterials and Biomedical Engineering, Канада* Цель: анализ эффективности принятия решений врачами на основе интегрированных наборов медицинских данных	– проводится подробный анализ взаимосвязи доступа пользователей с эффективностью использования их времени, точностью принятия решений и когнитивной нагрузкой	– в качестве информационной базы рассматриваются не данные электронных медицинских карт, а рекомендации для врачей интенсивной терапии – не рассматриваются непосредственно подходы к интеграции данных, фокус на оценке эффективности применения единого информационного пространства
Интеграция источников структурированной и неструктурированной медицинской информации [14] *University of Antwerpen, Бельгия* Цель: сравнительная оценка эффективности медицинских прогнозов на основании изолированных и интегрированных источников данных	– рассматриваются возможности применения ранней (одна схема для всех источников) и поздней интеграции (отдельная схема для каждого источника) данных – предложена схема ансамблирования методов ранней и поздней интеграции данных для повышения эффективности медицинского кодирования на их основе	– на текущий момент результаты не масшабируемы и ориентированы на локальное применение в рамках одного учреждения – данные представлены в собственном, отличном от стандартного, формате (негативно сказывается на интероперабельности) – в настоящее время существует только

Продолжение Таблицы 1

		десктопная версия, сопряженная с высокими требованиями к вычислительным ресурсам пользователя – не предусмотрено масштабирование системы
Интеграция данных для прецизионной персонализированной медицины [15] *National Institute of Cardiology 'Ignacio Chávez', Мексика* Цель: развитие персонализированной медицины с использованием методов искусственного интеллекта для интегрированных медицинских данных	– предложен подход к созданию информационной медицинской экосистемы как центра интеграции данных из распределенных источников	– не рассматриваются гетерогенные источники данных – формат данных отличен от стандартизованного или рекомендованного соответствующим организациями
Health Mining [16] *University of Catania, Италия* Цель: создание системы интеграции источников медицинских данных для анализа и поддержки принятия решений на этой основе	– подход ориентирован на интеграцию трех типов данных: клинические данные, социальные данные и данные IoT	– используется только централизованное хранение – приложение предназначено только для локального использования – расширение решения на большее число данных и организаций сопряжено с колоссальными вычислительными затратами

Продолжение Таблицы 1

Облачное хранилище открытых биомедицинских данных из разнородных источников [17] *Stanford University, США* Цель: разработка масштабируемых интеллектуальных инфраструктур биомедицинских данных для повышения качества биомедицинских исследований.	– используются облачные технологии, что при единой точке доступа создает широкие возможности для централизованного хранения больших данных, – основой является онтологический подход, позволяющий выявить связи между задействованными в процессе ресурсами.	– низкая реактивность веб-приложений из-за ограниченных выделенных ресурсов общедоступных конечных точек SPARQL, – сложность сопровождения проекта из-за семантической неоднородности источников данных.

3. Формализация задачи интеграции медицинских данных. В общем виде задача интеграции разнородных источников медицинских данных может быть описана следующим образом. Пусть представлены *N* источников медицинской информации:

$$A = [A_1, \dots, A_N]. \quad (1)$$

При этом множество кортежей данных из источника A_i можно обозначить как:

$$D^{A_i} = \{d_1^{A_i}, \dots, d_m^{A_i}\}. \quad (2)$$

Тогда информационный поток, поступающий от атомарного источника, в момент времени *t* описывается (в контексте обработки медицинских данных) линейной моделью вида:

$$y(t) = y(t_0) + v(t - t_0), \quad (3)$$

где *t* – стартовое время отсчета, $y(t_0)$ – количество сообщений за время *t*, *v* – средняя скорость увеличения информационного потока.

Следовательно, флуктуация каждого информационного потока может быть определена как:

$$\sigma(t_n) = \sqrt{\frac{1}{n}\sum_{i=0}^{n}[y(t_i) - (y(t_0) + v(t_i - t_0))]^2}. \quad (4)$$

Каждый источник характеризуется собственной схемой организации данных, предполагающей описание элементов данных посредством одной из известных моделей (реляционной, сетевой, иерархической, объектной и пр.). Здесь и далее схемы, используемые отдельными источниками данных, предлагается обозначать как «локальные». Пусть в контексте локальной схемы данные, поступающие от одного удаленного источника, представляют собой упорядоченный набор значений, где каждый кортеж может быть описан отношением вида:

$$X = [x_1, x_2 \dots, x_N], \quad (5)$$

где x_1,...,x_n – значения исходного кортежа для n параметров.

При этом предполагается, что данные из разнородных источников аккумулируются централизованно в хранилище, где, в свою очередь, используется глобальная схема данных, к которой должны быть трансформированы все локальные схемы:

$$Y = [y_1, y_2 \dots, y_m], \quad (6)$$

где y_1,...,y_m – значения кортежа в глобальном хранилище для m параметров.

Соответственно, для включения источника в единую систему данных необходим дополнительный компонент трансформации данных, обеспечивающий автоматическое отображение локальной схемы в глобальную с применением соответствующей схемы отображения (по аналогии с известной моделью XML – XSL(T)). Указанный компонент можно представить в виде функции отображения следующим образом:

$$f_p: \{a_1, a_2, \dots, a_n\} \xrightarrow{g} \{b_1, b_2, \dots, b_m\}; g = \{g_1, g, \dots, g_m\}, \quad (7)$$

где f_p – функция преобразования кортежа данных, задающая правила формирования нового кортежа из элементов заданного кортежа; $\{a_1, a_2, \dots, a_n\}$ – исходный кортеж данных; $\{b_1,b_2,\dots,b_m\}$ – формируемый кортеж данных; g=$\{g_1, g_2,\dots,g_m\}$ – эталонный набор параметров, который необходимо реализовать в формируемом кортеже данных.

Одним из значимых аспектов предлагаемого подхода является формирование результирующего набора данных как совокупности непосредственно содержательных данных, а также метаданных, которые характеризуют обстоятельства их получения (к примеру, информация о медицинском учреждении, используемом лабораторном оборудовании и пр.). Если рассматривать информационный поток от удаленного источника в качестве совокупности отдельных (необязательно взаимосвязанных) информационных блоков, то сказанное можно представить следующим образом:

$$D = \{X, M\}, \tag{8}$$

где X – массив кортежей содержательных данных, M – массив кортежей с метаданными для одного источника.

Для описания метаданных предлагается рассматривать источники данных в качестве группы ресурсов в соответствии с принятыми стандартами и спецификациями представления медицинской информации (в данном случае имеется в HL7 FHIR [5]). В общем виде каждый k-й ресурс R представлен совокупностью своих p характеристик – параметров соответствующего провайдера:

$$R^k = \{r_1: \{e_1, f_1\}; \dots; r_p: \{e_p, f_p\}; type\}, \tag{9}$$

где каждая характеристика r_i характеризуется парой ключ e_i и значение f_i; type – флаг типа ресурса.

При этом для повышения информативности ресурсы предлагается разделить на событийные и участники. К первым относятся атрибуты, характеризующие процесс получения данных (параметры оборудования, к примеру), а ко вторым – атрибуты, характеризующие непосредственно участников процесса формирования набора данных (пациенты, медицинский персонал и пр.):

$$\begin{gathered} R = Rs \cup Rp: \forall k\ R^k\mathrm{s} = \{r1: \{e1, f1\}; \dots; rp: \{ep, fp\};' \mathrm{s}'\}; \\ R^k\mathrm{p} = \{r1: \{e1, f1\}; \dots; rp: \{ep, fp\};' \mathrm{p}'\}. \end{gathered} \tag{10}$$

Последовательная обработка кортежей данных из поступающего от провайдера информационного потока (получение данных, их нормализация и приведение к стандартному виду типа HL7) в терминах кортежей данных может быть описано следующим образом. Пусть каждый k-й этап обработки данных приводит к

формированию нового кортежа данных. Этот факт может быть описан таким образом, что зафиксированное значение всех элементов множества X относительно одного значения параметра k представляет вектор состояния:

$$C_i = \langle X_1, k_1; \dots; X_n, k_n \rangle. \quad (11)$$

Тогда множество всех возможных векторов состояний $C = \{C_i, i = 1, \dots, |C|\}$ составляет полное множество состояний получаемых от источника данных, которое можно представить в виде $C = \prod_i k_i$ (в данном случае соотношение приведено для одного кортежа информационного потока, получаемого от отдельно рассматриваемого источника данных).

По аналогии формирование единого информационного пространства можно описать последовательной сменой состояний каждого его кортежа. При этом представляется целесообразным введение так называемого начального этапа $k = 0$, при котором элементы каждого кортежа единого информационного пространства представляют собой неопределенные (NULL) значения. На последнем k_n этапе формируется окончательный набор кортежей в соответствии с описанными выше функциями отображения и трансформации.

4. Подход к территориальному распределению хранилищ данных. Предлагаемый подход к интеграции данных из разнородных источников основан на их последовательном преобразовании с учетом соответствующих метаданных. Ожидаемым результатом данного процесса является формирование единого централизованного хранилища медицинской информации, построенного по принципу CDN (Content Delivery Network [19]).

Выбор CDN в качестве базового для решения выявленной проблемы обусловлен следующим. На протяжении нескольких лет авторами проводились исследования, связанные с повышением реактивности веб-приложений, ориентированных на данные (рассматривались различные прикладные пространственные данные, в частности, геофизические [21, 23]). Как и ожидалось, было установлено, что ключевым фактором, определяющим характеристики реактивности приложений, является объем и сложность запрашиваемых клиентской стороной размещенных на сервере данных. Однако вместе с тем проведенные исследования показали зависимость скорости отклика от расстояния между источником запроса (клиентом) и размещением запрашиваемых данных. В целом, это известная в веб-разработке проблема, заключающаяся в том, что

время ожидания ответа от сервера при направлении к нему запроса со стороны клиента тем больше, чем территориально дальше находится источник запроса от его цели. Известное решение – технология CDN – положена в основу предложенного подхода к модернизации физической организации системы хранения информации.

Известно, что в целом при использовании CDN-подхода распределение соединения «клиент – сервер» осуществляется по-другому. Запрос клиента автоматически переадресуется к географически ближайшему кэширующему серверу в составе CDN, что позволяет доставлять контент существенно быстрее по сравнению с аналогичным запросом, направленным к основному серверу. При этом существенного снижается пропускная нагрузка на основной сервер за счет задействования дополнительных серверов для кэширования данных [20].

Предполагается, что физически единое информационное хранилище должно быть разделено на несколько территориальных кластеров, объединяемых через внешние интерфейсы в общую систему данных. Каждый территориальный кластер включает в себя несколько источников данных, в качестве которых выступают медицинские учреждения и организации, размещенные в пределах некоторого пространственного кластера с заданной геопространственной привязкой. С заданной периодичностью необработанные данные от соответствующих источников поступают в территориально ближайшее хранилище данных выделенного пространственного кластера (рисунок 1).

В результате единое информационное хранилище медицинских данных должно представлять собой географически распределенную сетевую инфраструктуру из множества источников данных по схеме типа «звезда». В центре указанной структуры размещается хранилище данных, для всех остальных связанных с ним хранилищ осуществляется систематическая репликация всех данных, что позволяет хранить копию имеющихся данных непосредственно в каждом территориальном кластере. Поскольку предполагается работа непосредственно с данными, то представляется целесообразной именно процедура репликации, а не кэшировании, как это предусмотрено известным подходом для сети доставки контента.

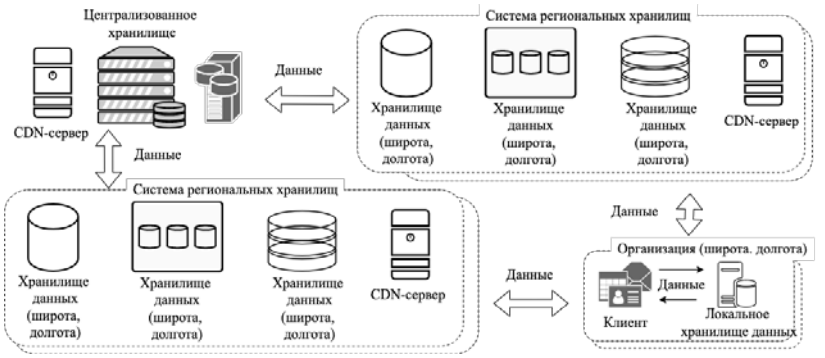


Рис. 1. Фрагмент схемы территориального CDN-распределения источников данных

Для загрузки и получения данных (так же в соответствии с принципами CDN) возможно применение технологии GeoDNS. Данная технология позволяет привязывать к одному доменному имени одновременно несколько IP-адресов. При анализе пришедшего от клиента запроса по соответствующему IP-адресу определяется географическое положение пользователя. В зависимости от полученного значения пользователь для получения / загрузки данных отправляется в территориально ближайшее хранилище данных. При этом расчет расстояния между соответствующими точками, соединенными дугой (расстояние по дуге большого круга), осуществляется по известной формуле гаверсинусов, которая представляет собой расстояние от центральной точки дуги, измеряемой удвоенным данным углом, до центральной точки хорды, стягивающей дугу:

$$d = 2r \arcsin\left(\sin^2\left(\frac{\varphi_2-\varphi_1}{2}\right) + \cos(\varphi_1)\cos(\varphi_2)\sin^2\left(\frac{\lambda_2-\lambda_1}{2}\right)\right), \quad (12)$$

где d – это центральный угол между двумя точками, лежащими на большой дуге; r – радиус сферы; φ_1 и φ_2 – широта первой и второй точек в радианах; λ_1 и λ_2 – долгота первой и второй точек в радианах.

Известно, что в CDN используется принцип кэширования по первому обращению, что означает максимальное время ожидания для пользователя, обратившегося к основному серверу первым. При этом все последующие пользователи будут получать данные, кэшированные на ближайшей к ним точке присутствия. Для преодоления ограничений, накладываемых этой схемой, используются технологии

регионального извлечения: соседние серверы, входящие в состав CDN, забирают контент друг у друга, а не обращаются к оригинальному серверу.

По аналогии с обозначенной схемой представляется целесообразным применить такой же подход к территориальному распределению и репликации данных в рамках предлагаемого единого хранилища медицинских данных. Пользователь, отправивший запрос на получение или загрузку данных, переадресуется к ближайшей точке присутствия и получает доступ к соответствующей копии основного хранилища медицинских данных. При этом если ближайшая точка присутствия не может найти данные (например, между сеансами репликации данных с основным сервером), то начинается поиск по соседним точкам присутствия, откуда и будет направлен ответ пользователю. Иными словами, имеет место подход, также известный как «tiered distribution» (или «многоуровневая раздача»).

Таким образом, за счет сокращения физического расстояния между потребителями и источниками данных можно уменьшить время отклика соответствующего контента. Поскольку одним из основных технических требований является возможность применения данных в веб-ориентированной среде, то время отклика от сервера при получении запроса от клиента на работу с информацией является критичным с точки зрения эргономики веб-приложений. Поэтому сокращение времени отклика за счет использования CDN-подобного подхода к организации хранения данных позволит при этом и повысить реактивность использующих их приложений.

Здесь представляется целесообразным отметить, что в качестве потребителя данных может выступать не только конечный пользователь, но и сторонние программные системы и приложения. При этом способ обращения к данным не зависит от типа обращающегося к ним клиента, поскольку IP-адрес источника запроса может быть выделен из заголовка любого пришедшего на сторону сервера запроса (предполагается, что взаимодействие между клиентом и сервером при этом осуществляется по стандартному интернет-протоколу HTTP(s)).

Еще одним положительным моментом от использования CDN-подхода для работы с медицинскими данными является снижение риска потери доступа к данным из-за потери основного сервера. В случае, если основной сервер или ближайшее для обращающегося к данным клиента хранилище по техническим причинам недоступны, запрос за счет принципа многоуровневой раздачи будет успешно выполнен. Иными словами, данные доступны все время, пока

восстанавливается работоспособность или доступность основного сервера или связанного с ним в сеть хранилища данных.

Уязвимость предлагаемого подхода сопряжена с привязкой к группе IP-адресов для управления распределением клиентских запросов между составляющими единого хранилища медицинских данных. В таком случае возможны блокировки, обусловленные блокировками сервисов, которые являются близкими по группе IP соответствующего CDN-провайдера. В результате таких блокировок возможна и блокировка территориальных компонент всего единого хранилища медицинских данных. Проблема решаема путем запроса изменения IP-адреса у CDN-провайдера.

Предложенная схема может быть (с некоторыми модификациями) адаптирована под различные геополитические регионы. Схема типа «звезда» по хранению данных может быть преобразована в схему типа «снежинка». В этом случае к выделяется основной сервер на федеральном уровне и по одному региональному серверу на уровне каждой области или региона страны. Данные от поставщиков поступают в реплицированное ближайшее хранилище, откуда передаются в региональное хранилище и далее – на основной сервер.

В этом случае процедура репликации должны быть асинхронно рассредоточена между различными уровнями поставщиков данных: от источников к региональным хранилищам, от региональных хранилищ – к основному серверу, и обратно. Техническая сложность реализации такого подхода, а также сопряженные с этим финансовые затраты представляются целесообразными в контексте существенного повышения реактивности использующих данные приложений.

5. Научная новизна предложенных решений. Отличительной особенностью предложенного подхода к интеграции данных является дополнение метаданных геопространственной меткой, характеризующей территориальное размещение источника данных на основании соответствующих параметров геолокации. Если представить сказанное в терминах HL7 FHIR [5] (как говорилось выше), то каждый k-й ресурс-источник данных R с p характеристиками – параметрами, необходимо дополнить следующим образом:

$$R^k = \{r1: \{e_1, f_1\}; \dots; r_p: \{e_p, f_p\}; \text{type}, \text{loc}: \{\text{lat}, \text{lng}\}\}, \quad (13)$$

где каждая характеристика *loc* характеризуется парой атрибутов lat и lng для представления географических широты и долготы источника-ресурса соответственно.

Глобальное информационное хранилище представляет собой совокупность кортежей, прошедших процедуру трансформации в соответствии с установленной схемой данных и дополненных соответствующей ссылкой на метаданные источника данных. Последние, в свою очередь, и содержат необходимую для позиционирования объекта геопространственную информацию. Так, множество кортежей данных из источника A_i можно обозначить как:

$$D^{A_i} = \left\{d_1^{A_i}, \dots, d_m^{A_i}, R^{A_i}\right\}, \tag{14}$$

где R^{A_i}– ссылка на метаданные источника A_i.

В соответствии с принципами CDN предлагается создание комплекса связанных хранилищ данных, репликация которых с централизованным хранилищем позволяет хранить данные в актуальном состоянии, а геопространственная метка обеспечивает возможность обработки запросов с учетом геопространственной локации их источников. Соответствующая информация представлены в метаданных хранилища медицинской информации:

$$\text{loc: \{lat: } \{x, x\_\text{val}\}, \text{lng: } \{y, y\_\text{val}\}\}, \tag{15}$$

где параметр loc метаданных произвольного источника данных представлен двумя парами параметров: соответственно ключ *x* и значение *x_val* для географической широты lat; ключ *y* и значение *y_val* для географической долготы lng местоположения источника данных. При этом при необходимости пространственная характеристика может быть задана в отличной от геодезической системе координат.

При получении данных от удаленного источника выполняется обработка соответствующего запроса, что предполагает непосредственно извлечение и анализ его геопространственной метки. Поскольку взаимодействие между клиентом и сервером предполагает применение протокола HTTP(HTTPS), соответствующая информация может быть получена посредством извлечения IP-адреса клиента непосредственно из заголовка запроса (Request, RemoteAddr):

$$\begin{gathered}\text{Request} = \{\text{Header}, \text{Body}\};\ \text{Header} = \{\text{General}, \text{Additional}\};\\ \text{Additional} = \{\text{Vary}, \text{Accept} - \text{Ranges}, \text{RemoteAddr}\},\end{gathered} \tag{16}$$

где Request – пакет запроса от клиента, Header – параметры заголовка запроса, Body – параметры тела запроса, General и Additional – основная и дополнительная части заголовка запроса соответственно, Vary – параметр, характеризующий возможность использования кэшированных ответов, Accept-Ranges – маркер составного запроса, RemoteAddr – IP-адрес хоста, с которого поступил запрос.

Параметр RemoteAddr трансформируется в пару значений типа «широта – долгота». Полученная метка сравнивается последовательно со всеми n зарегистрированными хранилищами данных R на предмет выявления кратчайшего расстояния между ними (по большой дуге):

$$\begin{gathered}\text{RemoteAddr} = \{\text{lat: } \{\text{x}, \text{x\_val}\}, \text{lng: } \{\text{y}, \text{y\_val}\}\}; \\ \min(\ \ |\text{RemoteAddr}, \text{Ri.loc}|, \text{i} = 1, \dots, \text{n}).\end{gathered} \tag{17}$$

На основании полученного результата определяется то хранилище, которое физически ближе остальных находится к источнику запроса. Данное хранилище в текущем сеансе запроса помечается как целевое:

$$\begin{gathered}\text{Target} = R^k : \forall l \neq k \ , \\ |\text{RemoteAddr}, R^l.loc| > \ |\text{RemoteAddr}, R^k.loc|.\end{gathered} \tag{18}$$

На последующем шаге формируется совокупность кортежей данных, передаваемая непосредственно в целевое хранилище Target. При этом выделенная из запроса по IP-адресу геопространственная метка также дополняет каждый из сформированных кортежей.

В результате формируется массив геопространственных данных, к которым, в свою очередь, применимы не только статистические модели и методы, но также и методы геостатистики, позволяющие в том числе оценить особенности соответствующего пространственного распределения соответствующих объектов, субъектов и процессов, а также оценить специфику их пространственной анизотропии при работе непосредственно с глобальным хранилищем.

Помимо геопространственной метки кортежи данных сопровождаются и временной меткой, что позволяет провести их обработку методами анализа пространственно-временных данных и оценить характер их анизотропии не только в пространстве, но и во времени.

В результате соответствующий кортеж, поступивший в запросе от источника данных A_i, при направлении в целевое хранилище будет иметь следующий вид:

$$D^{A_i} = \left\{d_1^{A_i}, \dots, d_m^{A_i}, R^{A_i}, \text{loc}, \text{timetag}\right\}, \quad (19)$$

где R^{A_i}– ссылка на метаданные источника A_i, loc – пространственная характеристика источника данных, заданная по ключу парой координат типа «широта – долгота», timetag – временная метка получения данных в формате UTC.

Для синхронизации с главным информационным хранилищем выполняется периодическая процедура репликации данных. При этом метки источников данных в соответствующих кортежах дополняются меткой соответствующего им целевого хранилища, принявшего запрос на размещение данных:

$$D^{A_i} = \left\{d_1^{A_i}, \dots, d_m^{A_i}, R^{A_i}, \text{loc}, \text{timetag}, \text{target}\right\}, \quad (20)$$

где target – ссылка на метаданные целевого хранилище данных.

Помимо непосредственно медицинских данных геопространственная метка должна сопровождать и запрос на их извлечение, поступающий от потребителя данных и характеризующий его физическое местоположение по результатам процедуры геолокации. Соответствующая информация по аналогии с вышеизложенным форматом запроса содержится в его заголовке (Request, RemoteAddr) и представляет собой IP-адрес источника запроса. На основании IP-адреса выделяется непосредственно геопространственная метка в виде пары геопространственных координат (географические широта и долгота).

Далее (аналогично обработке запроса на загрузку данных) выделенные геопространственные данные последовательно сопоставляются с метаданными распределенных хранилищ медицинских данных. Из них выявляется то хранилище, которое является ближайшим к исходной точке запроса с точки зрения соответствующего расстояния по большой дуге. Выделенное хранилище данных в рамках текущего сеанса взаимодействия с клиентом помечается как целевое.

В целевое хранилище направляется поступивший от клиента запрос на извлечение данных. При этом в результирующий набор кортежей, полученный при выполнении соответствующего запроса, добавляется пространственно-временная метка, характеризующая метаданные хранилища, из которого были извлечены данные, а также время формирования кортежа в формате UTC. Соответствующая информация помещается в заголовок отклика, формируемого на сервере.

Между основным и распределенным хранилищами выполняется процедура репликации данных. По умолчанию данная процедура должна выполняться с заданной периодичностью, например, один раз в сутки (каждые 24 часа). Соответственно обновление кортежей данных во всех хранилищах может быть представлено следующим образом:

$$D(t+\Delta t) \quad = \begin{Bmatrix} d_1\ (t+\Delta t), \dots, d_m(t+\Delta t), R \quad , \\ \text{loc}, (t++\Delta t)(UTC), \text{target} \end{Bmatrix}, \tag{21}$$

где t – временная характеристика запроса на загрузку данных, Δt – промежуток времени, по истечении которого данные должны быть обновлены и дополнены (при необходимости); R – ссылка на метаданные источника данных, loc – пространственная характеристика источника данных, заданная по ключу парой координат типа «широта – долгота», timetag – временная метка получения данных в формате UTC.

Вместе с тем во избежание трудоемких с точки зрения вычислительных затрат последовательных опросов источников «на местах» со стороны центрального хранилища данная процедура может быть дополнена следующим образом.

Представляется целесообразным введение дополнительного программного триггера, срабатывание которого привязано к изменению состава кортежей в соответствующем локальном хранилище информации. При выполнении заданного в программном сценарии условия на обновление данных выполняется процедуры их передачи (и обновления) сначала в центральное хранилище и далее в распределенные хранилища за исключением того, которое было непосредственно причиной срабатывания соответствующего триггера.

При этом во избежание возможных коллизий, связанных с началом одной операции обновления данных в центральном или распределенных хранилищах информации в архитектуру соответствующей информационной системы, целесообразно ввести программный диспетчер, в фоновом режиме отслеживающий входящий и исходящий информационные потоки в хранилищах.

6. Модульная схема интеграции данных. В работе [21] предложена и успешно апробирована схема создания единого информационного пространства геомагнитных данных на основе комбинирования принципов консолидации и федерализации информации. Представляется целесообразным несколько адаптировать предложенную схему под особенности создания и применения

медицинской информации с учетом описанной выше CDN-подобной организации физического хранения данных (рисунок 2).

Формирование централизованного хранилища медицинской информации осуществляется посредством фонового (по принципу Cron) последовательного выполнения связанных вычислительных процессов по получению, обработке и физическому размещению соответствующей информации, поступающей из территориально распределенных гетерогенных источников (в частности, разнородных медицинских информационных систем). При этом принцип Cron [22] предполагает составление графика рабочих процессов («workers»), которые запускаются автоматически (в фоновом режиме) и выполняются после настройки без непосредственного участия разработчика, сохраняя информацию о ходе своей работы в соответствующих программных логах, которые размещены на сервере в заданном каталоге (обычно корневом для проекта) и доступны для анализа.

Каждый этап обработки поступающих от гетерогенных источников медицинских данных представлен в виде одного или группы связанных программных модулей. При этом порядок выполнения вычислительных операций инкапсулируется таким образом, что каждый программный модуль может быть интерпретирован как сторонними, так и внешними потребителями как «черный ящик» с известными форматами представления входной и выходной информации.

Представляется целесообразным в контексте веб-ориентированной реализации рассматриваемого подхода применить элементы микросервисной архитектуры. В этом случае каждый отдельный функциональный модуль обработки данных (или группа связанных модулей) могут быть использованы независимо друг от друга, в том числе и сторонними программными системами и библиотеками. Для этого у каждого из представленных модулей-сервисов предполагается наличие собственного программного интерфейса (по принципу организации API – прикладных программных интерфейсов).

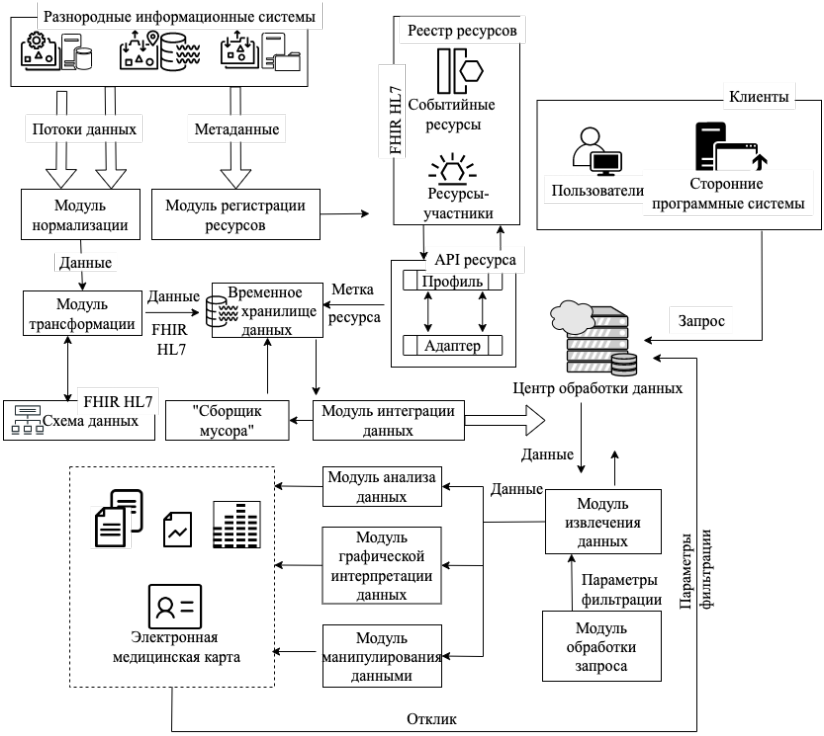


Рис. 2. Общая схема реализации предлагаемого подхода к интеграции данных

В результате обращение к каждому программному модулю осуществляется по его интерфейсу таким образом, что соответствующий вызов выполняется по названию (метке) интерфейса с передачей значений необходимых входных параметров и возвратом искомого результата по установленному протоколу. В качестве основного формата представления выходных данных используется JSON (JavaScript Object Notation – это формат, реализующий неструктурированное текстовое представление структурированных данных, основанное на принципе пар ключ-значение и упорядоченных списках [24]), который в настоящее время является фактически стандартом де-факто передачи данных в программных системах. В предлагаемой схеме выделены два типа данных – непосредственно содержательная информация (сами медицинские данные) и метаданные (информация об источниках данных). При этом содержательная информация, поступающая от внешних источников,

может быть представлена в структурированном, полуструктурированном и неструктурированном виде (в текстовом, CSV-подобных, мультимедиа, графических и иных форматах). В этой связи непосредственно данные непосредственно от своих провайдеров поступают в модуль нормализации, где проходят предварительную обработку для приведения к унифицированному виду. Так, к примеру, графические и мультимедиа данные преобразуются в BLOB-объекты [25], представляющие собой пригодные для обработки и анализа массивы двоичных данных, которые, как правило, применяются для хранения данных MIME-типа.

При этом для повышения эффективности дальнейшей обработки данные представляются в результирующем JSON-формате, что позволяет применять к ним библиотечные модели и методы трансформации. При этом сформированные BLOB-массивы также становятся частью выходных JSON-потоков в качестве секций формата CDATA (известного из спецификации W3C XML).

На последующем этапе данные формата JSON преобразуются в формат, специфичный для медицинских данных. В качестве такового выступает стандарт HL7, который в настоящее время применяется во многих медицинских учреждения в нашей стране и за рубежом для электронного обмена документами (особенно в тех медицинских организациях, где пациент получает непосредственно интенсивную помощь, в частности, в больницах). HL7 включает в себя концептуальные стандарты (HL7 RIM), стандарты приложений (HL7 CCOW), документальные стандарты (HL7 CDA), и стандарты обмена сообщениями (HL7 v2., v3.0 и HL7 FHIR). Согласно известным научным работам (например, [26]), самым актуальным и многообещающим из этих стандартов является HL7 FHIR (Health Level 7 - Fast Healthcare Interoperability Resources) [26].

Преобразование JSON-данных в указанный формат осуществляется на основании соответствующей схемы данных, используемой модулем трансформации. В схеме данных представлены основные параметры трансформации / отображения для последующего формирования результирующего набора данных.

Основой стандарта HL7 FHIR являются ресурсы (FHIR Resources). При этом каждый ресурс представляет собой отдельную независимую структурированную единицу информации, используемая при обмене медицинскими данными. Большинство ресурсов — это отображение реального мира в цифровых данных. Так, к примеру, на верхнем уровне абстракции предлагается выделять такие виды

ресурсов, как пациенты, обращения в медицинские организации и учреждения, а также результаты осмотров и исследований.

При дальнейшей детализации ресурсы декомпозируются на метаданные соответствующих провайдеров данных, которыми выступают конкретные медицинские информационные системы отдельных медицинских организаций и учреждений. Формирование ресурсов осуществляется вычислительным модулем регистрации ресурсов, на вход которого поступают метаданные источников медицинской информации, направляемой в рассматриваемое централизованное хранилище данных.

При этом предлагается выделение двух видов ресурсов. Первая группа обозначается как «событийные ресурсы» и включает в себя результаты всех действий, осуществляемых между пациентом и медицинским учреждением. Здесь могут быть представлены результаты осмотров, лабораторных исследований, программы лечения, обращения пациента в лечебное учреждение и пр.

Вторая группа ресурсов представляет собой характеристику участников взаимодействия в рассматриваемых процессах. Это непосредственно пациенты, медицинские учреждения, медицинский персонал, назначенные лечебные препараты и мероприятия, используемое лабораторное оборудования (фактически, здесь представлены как активные, так и пассивные участники медицинских процессов).

Все сформированные и зарегистрированные ресурсы размещаются в одноименном реестре, данные в котором представлены в соответствии со стандартом HL7 FHIR. При этом для унификации доступа к ресурсам каждому из них ставится в соответствие отдельный API, отличительной особенностью которого является самоидентифицируемость за счет представления метаданных в соответствующей структуре, именуемой профилем ресурса. Для программного доступа к ресурсу через его API используется его профиль, так называемая точка входа, принимающая входные данные и выдающая в результате уникальную метку соответствующего ресурса. Метка представляет собой свертку JSON-данных в структуру по спецификации JSON Web Token (JWT) [27, 28], что в итоге имеет вид объекта, который определен в открытом стандарте RFC 7519 [29].

Фоновое выполнение процесса трансформации содержательных данных предполагает наличие промежуточного (временного) хранилища, в котором (помимо хранения) выполняется дополнительная обработка данных, связанная с объединением

непосредственно данных и метки ресурса, характеризующего обстоятельства и источники их получения.

Данные из временного хранилища с заданной периодичностью поступают в модуль интеграции данных, в которой информационные потоки объединяются в единый массив и направляются в централизованное хранилище (центр обработки данных). Важно отметить, что в схему интеграции дополнительно включен так называемый «сборщик мусора», выполнение которого запускается завершением работы модуля интеграции данных [30, 31]. Указанный программный модуль отвечает за очищение временного хранилища данных после передачи обработанных данных в централизованное хранилище.

Непосредственно обработка пользовательских запросов (включая управление CDN-серверами) осуществляется посредством группы модулей, отвечающих за получение и декомпозицию поступивших с клиентской стороны запросов. Выделенные параметры запросы являются базой для выполнения вертикальной и горизонтальной фильтрации данных, размещенных в централизованном хранилище. Результаты извлечения данных поступают на вход одному или группе программных модулей, отвечающих за их анализ и интерпретацию конечному пользователю в виде наборов данных, графиков, таблиц, либо в формате электронной медицинской карты соответствующего пациента. Результаты интерпретации передаются на клиентскую сторону в качестве отклика и доступны как непосредственно для рендеринга, так и для применения сторонними средствами обработки данных.

7. Оценка эффективности предложенного подхода. Возможности предложенного подхода к интеграции данных из гетерогенных источников медицинской информации были проанализированы на примере тестовых данных, сформированных по аналогии с данными пациентов с бронхолегочными заболеваниями трех городских клинических больниц г. Уфа.

Данные представлены наборами значений параметров обследования 500 гипотетических пациентов и содержат информацию в текстовом и графическом форматах (составлены анонимизированные тестовые данные по каждому пациенту, характерные для программ лечения в рассмотренных медицинских учреждениях: электронные медицинские карты по каждому пациенту в формате XML, данные МКБ в формате CSV, биохимические и лабораторные тесты (в текстовом и графическом форматах), клинические заметки медицинского персонала (в текстовом формате), программы лечения

(в текстовом формате) и назначения фармацевтических препаратов (в текстовом формате)). Для повышения наглядности были развернуты тестовые сервера для хранения данных на виртуальном хостинге Beget со следующими характеристиками: веб-сервер с процессором 72 * Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz, Apache 2.4.5.1, сетевое журналируемое хранилище данных типа «ключ» – «значение» Redis (REmote DIctionary Server) [32].

Для имитации применения нескольких разнородных систем хранения созданы аналогичные копии сервера данных (в общей сложности были развернуты 4 копии), на которых в случайном порядке были размещены медицинские данные. Дополнительно также было выделено серверное хранилище с аналогичными техническими характеристиками для имитации централизованного хранилища (так называемого основного сервера), предназначенного для хранения результатов интеграции данных из соответствующих гетерогенных источников.

Вычислительный эксперимент проводился в несколько этапов для данных небольшого объема (порядка 150-200 Мб). На первом этапе осуществлялась загрузка данных с сервера, принятого за сервер отдельной организации. Далее информационный поток через обработку и унификацию данных перенаправлялся в хранилище следующего уровня (имитирующего региональный уровень). На завершающем этапе данные передаются на основной сервер.

Эксперименты были проведены по двум направлениям. С одной стороны, исследовались характеристики реактивности в случае применения стандартной монолитной веб-ориентированной архитектуры. С другой стороны, вычислительный эксперимент был проведен в рамках архитектуры по предложенному подходу к интеграции данных. Клиентская сторона при этом тестировалась со следующим характеристиками: CPU Intel Core i5 10300H ГГц, оперативная память 4 ГБ, скорость интернет-соединения ~52.4 Мбит/с. Результаты вычислительного эксперимента показали, что подход с монолитной архитектурой обеспечивает получение серверного отклика за 10-12 с при стабильном Интернет-соединении. Предложенный подход за счет применения CDN позволяет сократить значение указанного параметра до 6-7 с.

Вторая часть экспериментов с тестовыми данными была направлена на оценку изменения скорости получения серверного отклика при направлении простого одномерного запроса к центральному хранилищу. Подход, основанный на монолитной архитектуре и отсутствии территориального распределения системы

хранения данных, показал значение указанного параметра в 15-17 с, а предложенный подход с микросервисной архитектурой и CDN – 9-10 с. При этом были установлены пути повышения реактивности соответствующих веб-ориентированных приложений посредством применения технологии Node.js на сервере (предназначенной для работы с высоконагруженными приложениями) для определения ближайшей точки доступа с расчетом соответствующих расстояний по географическим координатам клиента [33].

8. Заключение. В настоящее время одним из магистральных направления развития информационных технологий является разработка систем цифрового здравоохранения, предназначенных в том числе для повышения эффективности процессов принятия решений на основе медицинских данных. Одним из важнейших факторов успешности реализации данного направления является повышение эффективности и прозрачности обработки соответствующих данных, а также реактивности соответствующих инструментально-программных средств. Значимым препятствием на пути решения поставленной задачи является постоянно возрастающие объем и сложность медицинских данных, а также разнообразие их источников. В конечном итоге требуется разработка подхода, который в результате своей программной реализации позволил бы решить поставленные задачи.

В работе предлагается подход к интеграции разнородных источников медицинской информации, отличительной особенностью которого является применение в архитектуре соответствующей программной системы принципов организации микросервисных приложений. Предложена многоуровневая организация системы сбора, обработки и хранения медицинской информации с дополнительным предоставлением метаданных, характеризующих источники информации и обстоятельства их получения (к примеру, используемое лабораторное оборудование и пр.).

Представляется целесообразным отметить еще одну отличительную особенность предложенного подхода, которая заключается в применении системы доставки контента (CDN) для организации системы хранения данных. Предложено создание копий основного сервера хранения данных с различной геопространственной привязкой, которая является определяющим фактором при направлении клиентского запроса на получение / загрузку соответствующей медицинской информации. Также

предусматривается (с заданной периодичностью) репликация данных между основным и зеркальными серверами-хранилищами данных.

Результаты проведенных вычислительных экспериментов (на тестовых данных, автоматически сформированных на основе реальных по бронхолегочным заболеваниям) в заданных технических условиях клиент-серверного веб-ориентированного взаимодействия показали, что применение предложенного подхода к решению перечисленных выше проблем позволит существенно сократить время получения отклика от серверной части (по сравнению с традиционным монолитным подходом) при загрузке данных в среднем на 39 % и при выполнении запросов на выборку – на 41%.

Литература

1. Snyder M., Zhou W. Big data and health // The Lancet. Digital Health. 2019. Vol. 1, iss. 6. P. E252-E-254
2. Комолов А.В. Обзор медицинских стандартов передачи электронной информации // Аллея науки. 2019. Т. 2. № 2(29). С. 909-913
3. Martínez-Costa C., Schulz S. HL7 FHIR: Ontological Reinterpretation of Medication Resources // Studies in Health Technology and Informatics. 2017. No. 235. P. 451–455. doi:10.3233/978-1-61499-753-5-451.
4. Mukhiya S., Rabbi F., Pun V. [et al.]. A GraphQL approach to Healthcare Information Exchange with HL7 FHIR // Procedia Computer Science. 2019. No. 160. P.338-345. doi:10.1016/j.procs.2019.11.082.
5. Hong N., Wang K., Wu S. [et al.] An Interactive Visualization Tool for HL7 FHIR Specification Browsing and Profiling // Journal of Healthcare Informatics Research. 2019. No. 3. doi:10.1007/s41666-018-0043-8.
6. Елоев М.С. Опыт внедрения медицинской информационной системы в многопрофильном амбулаторно-поликлиническом учреждении // Военно-медицинский журнал. 2014. Т. 335. № 9. С. 4-13
7. Alqudah A., Al-Emran M., Shaalan K. Medical data integration using HL7 standards for patient's early identification // PLOS ONE. 2021. No. 16. P. e0262067. doi:10.1371/journal.pone.0262067.
8. Brogan J., del Pilar M., López A. [et al.] Scalable data systems require creating a culture of continuous learning // EBioMedicine Home (Part of Lancet Discovery Science). 2021. Vol. 74, P. 103738, doi: https://doi.org/10.1016/j.ebiom.2021.103738
9. Prakash C., Amit Sh. National Institute of Malaria Research-Malaria Dashboard (NIMR-MDB): A digital platform for analysis and visualization of epidemiological data // The Lancet Regional Health. 2022. P. 100030.
10. Balicer R., Arnon A. Digital health nation: Israel's global big data innovation hub // The Lancet. 2017. Vol. 389, iss. 10088, p. 2451-2453. doi: https://doi.org/10.1016/S0140-6736(17)30876-0
11. Grabner M, Molife C, Wang L, Winfree K, Cui Z, Cuyun Carter G, Hess L. Data Integration to Improve Real-world Health Outcomes Research for Non–Small Cell Lung Cancer in the United States: Descriptive and Qualitative Exploration // JMIR Cancer 2021;7(2):e23161. DOI: 10.2196/23161
12. Mate S, Köpcke F, Toddenroth D, Martin M, Prokosch H-U, Bürkle T,et al. Ontology-Based Data Integration between Clinical and Research Systems // PLoS ONE. 2015. No. 10(1). P. e0116656. pmid:25588043.

13. Lin YL, Trbovich P, Kolodzey L, Nickel C, Guerguerian A. Association of Data Integration Technologies With Intensive Care Clinician Performance: A Systematic Review and Meta-analysis // JAMA Netw Open. 2019. No. 2(5). P. e194392. doi:10.1001/jamanetworkopen.2019.4392.
14. Scheurwegs E., Luyckx K. [et al.]. Data integration of structured and unstructured sources for assigning clinical codes to patient stays // Journal of the American Medical Informatics Association. 2016. Vol. 23, Iss. e1. P. e11–e19, https://doi.org/10.1093/jamia/ocv115
15. Martínez-García M., Hernández-Lemus E. Data Integration Challenges for Machine Learning in Precision Medicine // Front. Med. 2022. No. 8:784455. doi: 10.3389/fmed.2021.784455.
16. Di Stefano A., La Corte A., Scatá M. Health Mining: a new data fusion and integration paradigm // Proceedings of CIBB. 2014. Vol. 1. P. 98-107.
17. Kamdar M.R., Fernández J.D., Polleres A. [et al.] Enabling Web-scale data integration in biomedicine through Linked Open Data // Digit. Med. 2019. No. 2. P. 90. https://doi.org/10.1038/s41746-019-0162-5
18. Dhayne H., Haque R., Kilany R., Taher Y. In Search of Big Medical Data Integration Solutions. A Comprehensive Survey // IEEE Access. 2019. PP. 1-10. doi:10.1109/ACCESS.2019.2927491.
19. Kük E., Erel-Ozcevik M. Access protocol aware controller design for eMBB traffic in SD-CDN // Computer Networks. 2022. No. 205. P. 08686. doi:10.1016/j.comnet.2021.108686.
20. Zerwas J., Poese I., Schmid S., Blenk A. On the Benefits of Joint Optimization of Reconfigurable CDN-ISP Infrastructure // IEEE Transactions on Network and Service Management. 2021. PP. 105-112. Doi:10.1109/TNSM.2021.3119134.
21. Vorobev, A.; Soloviev, A.; Pilipenko, V.; Vorobeva, G.; Sakharov, Y. An Approach to Diagnostics of Geomagnetically Induced Currents Based on Ground Magnetometers Data // Appl. Sci. 2022, 12, 1522. https://doi.org/10.3390/app12031522
22. Choi, B. Python Network Automation Labs: cron and SNMPv3. In: Introduction to Python Network Automation. Apress, Berkeley, CA, 2021.doi:10.1007/978-1-4842-6806-3_15.
23. Vorobev, A.V., Pilipenko, V.A., Enikeev, T.A., Vorobeva, G.R. Geoinformation system for analyzing the dynamics of extreme geomagnetic disturbances from observations of ground stations // Computer Optics. 2020. No. 44(5). P. 782–790.
24. Barlas K., Stefaneas P. An Algebraic Specification / Schema for JSON // Journal of Engineering Research and Sciences. 2022. No. 1. doi:10.55708/js0105025.
25. Rajendran L., Veilumuthu R. An Efficient Distributed Model for XMLised Blob Data Generation // International Journal of Computer Applications. 2011. No. 22. doi:10.5120/2561-3519.
26. Yang Z., Jiang K., Lou M. [et al.] Defining health data elements under the HL7 development framework for metadata management // Journal of Biomedical Semantics. 2022. No. 13. doi:10.1186/s13326-022-00265-5.
27. Rahmatulloh A., Gunawan R., Nursuwars F. Performance comparison of signed algorithms on JSON Web Token // IOP Conference Series: Materials Science and Engineering. 2019. No. 550. P. 012023. doi:10.1088/1757-899X/550/1/012023.
28. Beltran V. Characterization of web single sign-on protocols // IEEE Communications Magazine. 2016. No. 54. P. 24-30. doi:10.1109/MCOM.2016.7514160.
29. Jones M., Bradley J., Sakimura N., JSON Web Token (JWT)., RFC 7519, doi:10.17487/RFC7519, May 2015, https://www.rfc-editor.org/info/rfc7519.

30. Cai Sh., Chen K., Liu M. [et al.] Garbage collection and data recovery for N2DB // Tsinghua Science and Technology. 2022. No. 27. P. 630-641. doi:10.26599/TST.2021.9010016.
31. Garcia A., May D., Nutting E. Integrated Hardware Garbage Collection // ACM Transactions on Embedded Computing Systems. 2021. No. 20. P. 1-25. doi:10.1145/3450147.
32. Zhang Q., Bernstein P., Berger D., Chandramouli B. Redy: remote dynamic memory cache // Proceedings of the VLDB Endowment. 2021. No. 15. P. 766-779. doi:10.14778/3503585.3503587.
33. Tserpes K., Pateraki M., Varlamis I. Strand: scalable trilateration with Node.js // Journal of Cloud Computing. 2019. No. 8. doi:10.1186/s13677-019-0142-y.

Юсупова Нафиса Исламовна — д-р техн. наук, профессор, кафедра вычислительной математики и кибернетики, ФГБОУ ВО Уфимский государственный авиационный технический университет. Область научных интересов: интеллектуальные методы обработки информации и управления с приложениями в социальных, экономических и технических системах. Число научных публикаций — 560. yussupova@ugatu.ac.ru; улица Карла Маркса, 12, 450008, Уфа, Россия; р.т.: +7(908)350-3285.

Воробьева Гульнара Равилевна — д-р техн. наук, профессор, доцент, кафедра вычислительной математики и кибернетики, ФГБОУ ВО Уфимский государственный авиационный технический университет. Область научных интересов: геоинформационные и веб-технологии, системы хранения и обработки информации. Число научных публикаций — 141. gulnara.vorobeva@gmail.com; улица Карла Маркса, 12, 450008, Уфа, Россия; р.т.: +7(908)350-3285.

Зулкарнеев Рустэм Халитович — д-р мед. наук, профессор, кафедра пропедевтики внутренних болезней, ФГБОУ ВО Башкирский государственный медицинский университет Минздрава России. Область научных интересов: пульмонология, медицинские информационные системы. Число научных публикаций — 30. zurustem@mail.ru; улица Карла Маркса, 9/1, 450076, Уфа, Россия; р.т.: +7(908)350-3285.

Поддержка исследований. Работа выполнена при финансовой поддержке РНФ (проект № 22-19-00471).

DOI 10.15622/ia.21.5.2

N. YUSUPOVA, G. VOROBEVA, R. ZULKARNEEV

APPROACH TO SOFTWARE INTEGRATION OF HETEROGENEOUS SOURCES OF MEDICAL DATA BASED ON MICROSERVICE ARCHITECTURE

Yusupova N., Vorobeva G., Zulkarneev R. **Approach to Software Integration of Heterogeneous Sources of Medical Data Based on Microservice Architecture.**

Abstract. The task of processing medical information is currently being solved in our country and abroad by means of heterogeneous medical information systems, mainly at the local and regional levels. The ever-increasing volume and complexity of the accumulated information, along with the need to ensure transparency and continuity in the processing of medical data (in particular, for bronchopulmonary diseases) in various organizations, requires the development of a new approach to integrating their heterogeneous sources. At the same time, an important requirement for solving the problem is the possibility of web-oriented implementation, which will make the corresponding applications available to a wide range of users without high requirements for their hardware and software capabilities. The paper considers an approach to the integration of heterogeneous sources of medical information, which is based on the principles of building microservice web architectures. Each data processing module can be used independently of other program modules, providing a universal entry point and the resulting data set in accordance with the accepted data schema. Sequential execution of processing steps implies the transfer of control to the corresponding program modules in the background according to the Cron principle. The schema declares two types of data schemas - local (from medical information systems) and global (for a single storage system), between which the corresponding display parameters are provided according to the principle of constructing XSLT tables. An important distinguishing feature of the proposed approach is the modernization of the medical information storage system, which consists in creating mirror copies of the main server with periodic replication of the relevant information. At the same time, the interaction between clients and data storage servers is carried out according to the type of content delivery systems with the creation of a connection session between end points based on the principle of the nearest distance between them, calculated using the haversine formula. The computational experiments carried out on test data on bronchopulmonary diseases showed the effectiveness of the proposed approach both for loading data and for obtaining them by individual users and software systems. Overall, the reactivity score of the corresponding web-based applications was improved by 40% on a stable connection.

Keywords: medical data, data warehouses, data integration, web applications, content delivery system, microservice architecture.

Yusupova Nafisa — Ph.D., Dr.Sci., Professor, Computational mathematics and cybernetics department, Ufa State Aviation Technical University. Research interests: intelligent methods of information processing and management with applications in social, economic and technical systems. The number of publications — 560. yussupova@ugatu.ac.ru; 12, Karl Marx St., 450008, Ufa, Russia; office phone: +7(908)350-3285.

Vorobeva Gulnara — Ph.D., Dr.Sci., Professor, Associate professor, Computational mathematics and cybernetics department, Ufa State Aviation Technical University. Research interests: geoinformation and web technologies, systems of information storing and processing.

The number of publications — 141. gulnara.vorobeva@gmail.com; 12, Karl Marx St., 450008, Ufa, Russia; office phone: +7(908)350-3285.

Zulkarneev Rustem — Ph.D., Dr.Sci., Professor, Department of propaedeutics of internal diseases, Bashkir State Medical University of the Ministry of Health of Russia. Research interests: pulmonology, medical information systems. The number of publications — 30. zurustem@mail.ru; 9/1, Karl Marx St., 450076, Ufa, Russia; office phone: +7(908)350-3285.

Acknowledgements. This research is supported by RSF (grant 22-19-00471).

References

1. Snyder M., Zhou W. Big data and health. The Lancet. Digital Health. 2019. Vol. 1, iss. 6. P. E252-E-254
2. Komolov A.V. [Review of medical standards for the transmission of electronic information]. Alleya nauki – Alley of Science. 2019. vol. 2. no. 2(29). pp. 909-913 (in Russ.)
3. Martínez-Costa C., Schulz S. HL7 FHIR: Ontological Reinterpretation of Medication Resources. Studies in Health Technology and Informatics. 2017. no. 235. pp. 451–455. doi:10.3233/978-1-61499-753-5-451.
4. Mukhiya S., Rabbi F., Pun V. [et al.]. A GraphQL approach to Healthcare Information Exchange with HL7 FHIR. Procedia Computer Science. 2019. no. 160. pp.338-345. doi:10.1016/j.procs.2019.11.082.
5. Hong N., Wang K., Wu S. [et al.] An Interactive Visualization Tool for HL7 FHIR Specification Browsing and Profiling. Journal of Healthcare Informatics Research. 2019. No. 3. doi:10.1007/s41666-018-0043-8.
6. Eloev M.S. [Experience in implementing a medical information system in a multidisciplinary outpatient clinic]. Voenno-medicinsky Journal – Military Medical Journal. 2014. vol. 335. no. 9. pp. 4-13
7. Alqudah A., Al-Emran M., Shaalan K. Medical data integration using HL7 standards for patient's early identification. PLOS ONE. 2021. no. 16. p. e0262067. doi:10.1371/journal.pone.0262067.
8. Brogan J., del Pilar M., López A. [et al.] Scalable data systems require creating a culture of continuous learning. EBioMedicine Home (Part of Lancet Discovery Science). 2021. Vol. 74, P. 103738, doi: https://doi.org/10.1016/j.ebiom.2021.103738
9. Prakash C., Amit Sh. National Institute of Malaria Research-Malaria Dashboard (NIMR-MDB): A digital platform for analysis and visualization of epidemiological data. The Lancet Regional Health. 2022. P. 100030.
10. Balicer R., Arnon A. Digital health nation: Israel's global big data innovation hub. The Lancet. 2017. Vol. 389, iss. 10088, p. 2451-2453. doi: https://doi.org/10.1016/S0140-6736(17)30876-0
11. Grabner M, Molife C, Wang L, Winfree K, Cui Z, Cuyun Carter G, Hess L. Data Integration to Improve Real-world Health Outcomes Research for Non–Small Cell Lung Cancer in the United States: Descriptive and Qualitative Exploration. JMIR Cancer 2021;7(2):e23161. DOI: 10.2196/23161
12. Mate S, Köpcke F, Toddenroth D, Martin M, Prokosch H-U, Bürkle T,et al. Ontology-Based Data Integration between Clinical and Research Systems. PLoS ONE. 2015. No. 10(1). P. e0116656. pmid:25588043.
13. Lin YL, Trbovich P, Kolodzey L, Nickel C, Guerguerian A. Association of Data Integration Technologies With Intensive Care Clinician Performance: A Systematic Review and Meta-analysis // JAMA Netw Open. 2019. No. 2(5). P. e194392. doi:10.1001/jamanetworkopen.2019.4392.

14. Scheurwegs E., Luyckx K. [et al.]. Data integration of structured and unstructured sources for assigning clinical codes to patient stays. Journal of the American Medical Informatics Association. 2016. Vol. 23, Iss. e1. P. e11–e19, https://doi.org/10.1093/jamia/ocv115
15. Martínez-García M., Hernández-Lemus E. Data Integration Challenges for Machine Learning in Precision Medicine. Front. Med. 2022. No. 8:784455. doi: 10.3389/fmed.2021.784455.
16. Di Stefano A., La Corte A., Scatá M. Health Mining: a new data fusion and integration paradigm. Proceedings of CIBB. 2014. Vol. 1. P. 98-107.
17. Kamdar M.R., Fernández J.D., Polleres A. [et al.] Enabling Web-scale data integration in biomedicine through Linked Open Data. Digit. Med. 2019. No. 2. P. 90. https://doi.org/10.1038/s41746-019-0162-5
18. Dhayne H., Haque R., Kilany R., Taher Y. In Search of Big Medical Data Integration Solutions. A Comprehensive Survey. IEEE Access. 2019. PP. 1-10. doi:10.1109/ACCESS.2019.2927491.
19. Kük E., Erel-Ozcevik M. Access protocol aware controller design for eMBB traffic in SD-CDN. Computer Networks. 2022. No. 205. P. 08686. doi:10.1016/j.comnet.2021.108686.
20. Zerwas J., Poese I., Schmid S., Blenk A. On the Benefits of Joint Optimization of Reconfigurable CDN-ISP Infrastructure. IEEE Transactions on Network and Service Management. 2021. PP. 105-112. Doi:10.1109/TNSM.2021.3119134.
21. Vorobev, A.; Soloviev, A.; Pilipenko, V.; Vorobeva, G.; Sakharov, Y. An Approach to Diagnostics of Geomagnetically Induced Currents Based on Ground Magnetometers Data. Appl. Sci. 2022, 12, 1522. https://doi.org/10.3390/app12031522.
22. Choi, B. Python Network Automation Labs: cron and SNMPv3. In: Introduction to Python Network Automation. Apress, Berkeley, CA, 2021.doi:10.1007/978-1-4842-6806-3_15.
23. Vorobev, A.V., Pilipenko, V.A., Enikeev, T.A., Vorobeva, G.R. Geoinformation system for analyzing the dynamics of extreme geomagnetic disturbances from observations of ground stations. Computer Optics. 2020. No. 44(5). P. 782–790.
24. Barlas K., Stefaneas P. An Algebraic Specification / Schema for JSON. Journal of Engineering Research and Sciences. 2022. No. 1. doi:10.55708/js0105025.
25. Rajendran L., Veilumuthu R. An Efficient Distributed Model for XMLised Blob Data Generation. International Journal of Computer Applications. 2011. No. 22. doi:10.5120/2561-3519.
26. Yang Z., Jiang K., Lou M. [et al.] Defining health data elements under the HL7 development framework for metadata management. Journal of Biomedical Semantics. 2022. No. 13. doi:10.1186/s13326-022-00265-5.
27. Rahmatulloh A., Gunawan R., Nursuwars F. Performance comparison of signed algorithms on JSON Web Token. IOP Conference Series: Materials Science and Engineering. 2019. No. 550. P. 012023. doi:10.1088/1757-899X/550/1/012023.
28. Beltran V. Characterization of web single sign-on protocols. IEEE Communications Magazine. 2016. No. 54. P. 24-30. doi:10.1109/MCOM.2016.7514160.
29. Jones M., Bradley J., Sakimura N., JSON Web Token (JWT)., RFC 7519, doi:10.17487/RFC7519, May 2015, https://www.rfc-editor.org/info/rfc7519.
30. Cai Sh., Chen K., Liu M. [et al.] Garbage collection and data recovery for N2DB. Tsinghua Science and Technology. 2022. No. 27. P. 630-641. doi:10.26599/TST.2021.9010016.
31. Garcia A., May D., Nutting E. Integrated Hardware Garbage Collection. ACM Transactions on Embedded Computing Systems. 2021. No. 20. P. 1-25. doi:10.1145/3450147.

32. Zhang Q., Bernstein P., Berger D., Chandramouli B. Redy: remote dynamic memory cache. Proceedings of the VLDB Endowment. 2021. No. 15. P. 766-779. doi:10.14778/3503585.3503587.
33. Tserpes K., Pateraki M., Varlamis I. Strand: scalable trilateration with Node.js. Journal of Cloud Computing. 2019. No. 8. doi:10.1186/s13677-019-0142-y.

DOI 10.15622/ia.21.5.3

I. SUROV

OPENING THE BLACK BOX: FINDING OSGOOD'S SEMANTIC FACTORS IN WORD2VEC SPACE

Surov I. **Opening the Black Box: Finding Osgood's Semantic Factors in Word2vec Space.**

Abstract. State-of-the-art models of artificial intelligence are developed in the black-box paradigm, in which meaningful information is limited to input-output interfaces, while internal representations are not interpretable. The resulting algorithms lack explainability and transparency, requested for responsible application. This paper addresses the problem by a method for finding Osgood's dimensions of affective meaning in multidimensional space of a pre-trained word2vec model of natural language. Three affective dimensions are found based on eight semantic prototypes, composed of individual words. Evaluation axis is found in 300-dimensional word2vec space as a difference between positive and negative prototypes. Potency and activity axes are defined from six process-semantic prototypes (perception, analysis, planning, action, progress, and evaluation), representing phases of a generalized circular process in that plane. All dimensions are found in simple analytical form, not requiring additional training. Dimensions are nearly orthogonal, as expected for independent semantic factors. Osgood's semantics of any word2vec object is then retrieved by a simple projection of the corresponding vector to the identified dimensions. The developed approach opens the possibility for interpreting the inside of black box-type algorithms in natural affective-semantic categories, and provides insights into foundational principles of distributive vector models of natural language. In the reverse direction, the established mapping opens machine-learning models as rich sources of data for cognitive-behavioral research and technology.

Keywords: semantics, dimension, Osgood, affective meaning, interpretation, word2vec, language, black box.

1. Introduction. Machine-learning approaches to natural language processing suffer from the interpretability problem. Popular models represent words by states of a hidden layer of a neural network, trained for various linguistic tasks on large corpora of texts [1–4]. Relevant regularities of natural language are then encoded in high-dimensional vectors, components of which say nothing to their developers and users. Although efficient for many tasks [5–8], this "black box" paradigm is problematic for applications of artificial intelligence (AI), requested to be transparent and explainable [9, 10].

Compared to the dimensionality of machine-learning (ML) representations, interpretable factors of human thought are much less in number. They are *evaluation* (pleasantness-unpleasantness), *activity* (active-passive, external-internal), and *potency* (strength, dominance, openness, freedom) [11, 12]. As established by Charles Osgood, these dimensions account for the majority of judgment variance for various objects in semantic-differential scales like good-bad, heavy-light, soft-hard, straight-curved, etc.

(ibid). Although widely known in cognitive science, these results are largely ignored in modern AI and ML research.

Previous attempts to establish the missing link achieved only partial success. *Evaluation*, *potency*, and *activity* are found to correlate with 208, 187, and 175 out of 300 raw word2vec dimensions, respectively [13]. Out of 280 principal components of the same model, these numbers are 38, 35, and 37 (ibid.), indicating that the achieved correspondence is far from useful. Nevertheless, word2vec and similar (distributional-vector) representations of language are known to encode specific types of semantic information that can be extracted by additional methods [14, 15].

An inspiring example is provided by self-organized semantic maps of natural languages [16, 17]. These maps, built from synonym-antonym relations by a specially designed algorithm (minimizing an energy-like cost function of word configuration), consistently identify Osgood's dimensions as the main principal components of English, German, Spanish, French, and Russian. Although expected to take a central place in biologically-inspired cognitive architectures [18], the present form of these maps is yet of demonstrative character. The combination of semantic interpretability with practicality of applied machine learning remains a challenge.

The present work addresses this problem for the word2vec model [1], baseline in ML approach to natural language processing. The paper describes a novel method to identify directions in the multidimensional word2vec space, responsible for Osgood's semantic factors. In contrast to the aforementioned attempts, the method does not rely on special variance properties of semantic dimensions, presupposed in principal component analysis. Instead, it finds unique dimensions based on a small sample of supervised data and a simple mathematical procedure. The theory and realization of the algorithm are described in Sections 2 and 3. Tests of stability and predictive potential of the developed method are reported in Section 4. Section 5 discusses the implications of the result.

2. Theory. The theory is developed in the following steps. After introducing the source data, Section 2.1 describes a method for finding the evaluation axis (Z) in the word2vec space. Section 2.2 then elaborates this logic to find the potency-activity (XY) plane. Finally, Section 2.3 shows how to use the obtained dimensions to extract Osgood's semantics of arbitrary word2vec representation.

Source data. This work uses a word2vec model of the English language, trained to predict skipped words based on their surroundings on Google News dataset of about 100 billion words [19]. The model encodes ~ 3 million English words w in 300-dimensional vectors $\vec{w}$. These vectors reflect regularities of

linguistic practice in algebraic relations of type:

$$\overrightarrow{Greece} - \overrightarrow{Athens} \approx \overrightarrow{Russia} - \overrightarrow{Moscow}, \tag{1a}$$

$$\overrightarrow{king} - \overrightarrow{man} \approx \overrightarrow{queen} - \overrightarrow{woman}, \tag{1b}$$

useful in natural language analysis [1, 20]. Neither individual dimensions of vectors (1), nor principal components of the model as a whole, however, have understandable meanings, leading to the aforementioned interpretability problem.

2.1. Evaluation Z axis. The finding of Osgood's evaluation axis in 300-dimensional word2vec space is suggested by the above identities. Both sides of (1a), for example, refer to the concept close to *country*, while sides of (1b) encode the notion of *mightiness*. In the same way, *pleasantness* can be defined as:

$$\begin{gathered}\overrightarrow{good} - \overrightarrow{bad}, \\ \overrightarrow{joy} - \overrightarrow{grief}, \\ \overrightarrow{well} - \overrightarrow{poor},\end{gathered} \tag{2}$$

and similar differences. To get a unitary definition, the sides of these pairs are combined in averaged vectors:

$$\vec{W}_{\mathrm{good}} = \frac{1}{N}\sum_{j=1}^{N} \vec{w}^{j}_{\mathrm{good}}, \qquad \vec{W}_{\mathrm{bad}} = \frac{1}{N}\sum_{j=1}^{N} \vec{w}^{j}_{\mathrm{bad}}, \tag{3}$$

where N in the number of individual words in "good" and "bad" *prototypes*. Difference between them then defines Osgood's *evaluation* axis:

$$\vec{Z} = \vec{W}_{\mathrm{good}} - \vec{W}_{\mathrm{bad}}, \tag{4}$$

averaging out semantic variations among lines of (2) as illustrated in Figure 1(a).

2.2. Potency – Activity XY plane. Osgood's *activity* and *potency* dimensions could be found by the same opponent-prototype method as used for the *evaluation* axis above. The XY plane then would be defined via "active-passive" and "strong-weak" prototype pairs analogous to (4). Four points, however, are excessive to define a plane; on the other hand, it is desirable to be able to use more prototypes to achieve the necessary precision of the resulting dimensions $\vec{X}$ and $\vec{Y}$. The rest of this section describes a method for this.

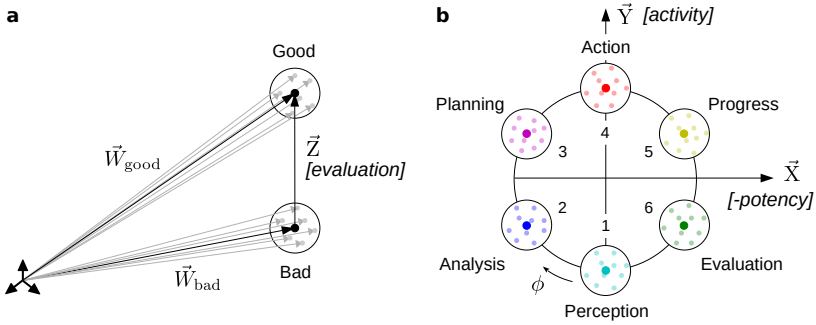


Fig. 1. Finding Osgood's semantic dimensions in the word2vec space: a) evaluation axis $\vec{Z}$ is defined as a difference between 300-dimensional "good" and "bad" prototypes (4); individual word vectors $\vec{w}^i$ and central prototypes (3) are shown in black and gray, respectively, b) finding the XY plane via six process-semantic prototypes, introduced in [21]

2.2.1. General idea. Suppose we have k word2vec prototypes of type (3):

$$\vec{W}_i = \frac{1}{N_i} \sum_{j=1}^{N} \vec{w}_i^j, \quad i = 1 \dots k, \tag{5}$$

that should have coordinates $(x_1, y_1) \dots (x_k, y_k)$ in the XY plane to be found. If these prototypes are indeed coplanar, their vectors $\vec{W}_i$ must be representable as linear superpositions of its basis vectors $\vec{X}$ and $\vec{Y}$. In the matrix form, illustrated in Figure 2, this is expressed by decomposition:

$$W = A * \Omega_{xy}, \tag{6}$$

where two rows of Ω_{xy} are the basis vectors $\vec{X}$ and $\vec{Y}$, k rows of W are word2vec prototypes $\vec{W}_i$, and k rows of A are the expected coordinates (x_i, y_i) of these prototypes in the XY plane.

If the coordinate matrix A is invertible, Ω_{xy} is obtained by multiplying the sides of (6) by A^{-1} from the left, so that:

$$\Omega_{xy} = A^{-1} * W. \tag{7}$$

Being rows of this matrix, the sought dimensions $\vec{X}$ and $\vec{Y}$ are then obtained as linear combinations of the prototype vectors $\vec{W}_i$ in 300-dimensional word2vec space.

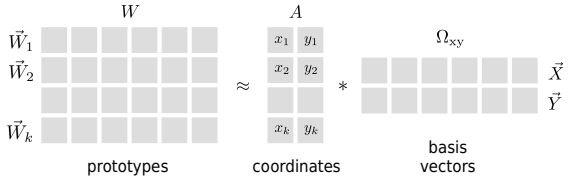


Fig. 2. Finding the potency-activity XY plane from several prototypes. The corresponding word2vec vectors $\vec{W}_i$ are represented by linear superposition of the basis vectors $\vec{X}$ and $\vec{Y}$ with coordinates $\{x_i, y_i\}$. Precise equality (6) is reached if the prototypes are coplanar. Basis vectors forming matrix Ω_{xy} are found by multiplying both sides by inverted coordinate matrix A^{-1} (7)

2.2.2. Specification of the prototypes. To find the XY plane, this study used the set of $k = 6$ prototypes close to that developed in [21]. These prototypes are *perception*, *analysis* (of novelty), *planning*, *action*, *progress*, and *evaluation* (of the result), forming a circular template for process representation in natural thinking (ibid)[1]:

1. **Perception** prototype describes observations, leading to the discovery of novelty. In the process of writing a paper this could include, for example, the accumulating new knowledge or demands for clarification of previous results. Realization of the fact that a new paper is needed turns the process to the next stage.

2. **Analysis** prototype accounts for the understanding of a newly identified factor. In the paper example, this is the thinking process revealing what exactly needs to be explained, and which problem will be resolved. The formalization of these requirements turns the process to the next stage.

3. **Planning** prototype describes a prospective vision of how the novelty will be handled. In the same example, this includes conceptualization of the paper and specification of its structure, approximate content of the sections, and selection of an appropriate journal for submission. The finalization of this project moves the process to the next stage.

4. **Action** prototype describes an implementation of the project. This is the process of making a paper with all accompanying activities: writing the text, producing the figures, selecting the references, and formatting the

[1]These prototypes are subsequent stages of a generalized cyclical process, most obviously derived from day-night and seasonal cycles. This progression is found in any process from taking a shower to running a space mission, with particular granularity chosen according to the desired precision. Listed prototypes are obtained from doubling the resolution of a basic three-stage sequence *analysis - action - evaluation*, identical to the baseline cybernetic *sense (evaluation) - think (analysis) - act (action)* control loop [22].

manuscript according to the journal's requirements. Submission of the paper finalizes this stage of the process.

5. **Progress** stage situates the obtained prototype in real circumstances based on the received feedback. The paper-making process includes answering to reviewer's questions, correcting the contents, and resubmission to another journal if necessary. The stage is finalized by the journal's decision.

6. **Evaluation** prototype accounts for subjective estimation of the result. In the case of a positive decision, it describes how well the paper achieved the initial goals, and estimates intended and unintended consequences. Alternatively, implications of a negative result, e.g. abandoning the project based on the received feedback, are summarized and recorded.

Individual words forming the prototypes were selected to represent their function in the abstract cyclical process. For example, various aspects and types of perception are accounted by observation, cognition, feedback, feeling, reflection, intuition, and the like; analysis, similarly, is strongly associated with novelty as its object, attention, thinking, reasoning, questioning, forming hypotheses and theories. As compared e.g. to *arm* or *table*, such concepts with definite process functions, stable across the majority of contexts, are small in number. Manually formed lists of such process-functional English terms for each prototype are given in Table 1.

2.2.3. Coordinate matrix. According to the cyclical topology of the process template, it can be visualized as a circular trajectory in the XY plane. Uniform discretization of this trajectory to six process stages is then expected to form a regular hexagon as shown in Figure 1(b). With the radius of the circle set to unity, corresponding coordinates of the prototype centers become:

$$\begin{aligned} x_i &= -\sin\Phi_i, \\ y_i &= -\cos\Phi_i, \qquad \Phi_i = 60^\circ * (i-1), \end{aligned} \tag{8}$$

where the process phase ϕ starts from the *perception* prototype with $\Phi_1 = 0^\circ$ and increases in steps of 60°. (Pseudo)inverse of the resulting coordinate matrix A (6), remarkably, is proportional to its transposition:

$$A^{-1} = \frac{1}{3}A^{\mathrm{T}} = \frac{1}{6}\begin{bmatrix} 0 & -\sqrt{3} & -\sqrt{3} & 0 & \sqrt{3} & \sqrt{3} \\ -2 & -1 & 1 & 2 & 1 & -1 \end{bmatrix}, \tag{9}$$

with coefficient $1/3$[2]. Substituting (9) in (7) produces the requested $\vec{X}$ and $\vec{Y}$ vectors in 300-dimensional word2vec space.

[2]Obtained as $2/k$. In this form, the analytical part of (9) holds for any number of prototypes, dividing the process circle into k equal sectors analogous to Figure 1(b).

Table 1. Individual terms, forming two Z-axis (3) and six XY-plane prototypes (5) shown in Figure 1

Prototype	**Individual terms**
Good	good light well fine yes
Bad	bad dark poor vice no
Perception	perception observation cognition feedback feeling reflection intuition insight introspection monitoring sensing data forecast prediction expectation contemplation anticipation
Analysis	analytics novelty hypothesis theory problem reason mystery question attention factor issue query puzzle challenge think
Planning	plan aim goal model concept intent purpose project principle plot motive strategy design map vision solve
Action	action act work duty develop implement manage deal compete cooperate execute accomplish produce construct engage fulfill
Progress	progress regress advance growth attainment agreement negotiation gain bargain increase decrease output yield completion profit return
Evaluation	evaluation estimation result end summation summary victory defeat conclusion final outcome aftermath expiration termination record score

2.2.4. Complex-valued form. The latter procedure can be made intuitive by rendering it in complex-valued form, with $\vec{Y}$ and $\vec{X}$ becoming real and imaginary axes of the complex plane. Values (8) then become real and imaginary parts of a single complex-valued coordinate $c_i = -\exp \mathrm{i}\phi_i$, so that (9) transforms to a single complex-valued row:

$$A^{-1} = \frac{-1}{3}\left[e^{\mathrm{i}\Phi_1} \quad e^{\mathrm{i}\Phi_2} \quad \dots \quad e^{\mathrm{i}\Phi_6}\right]. \tag{10}$$

For arbitrary k, the product (7) then takes form:

$$\vec{\Omega} = -\frac{2}{k}\sum_{i=1}^{k} \vec{W}_i e^{\mathrm{i}\Phi_i}, \tag{11}$$

with $\vec{Y}$ and $\vec{X}$ being real and imaginary parts of this vector. Up to the normalization factor $2/k$, (11) is then recognized as a complex-valued generalization of (4), combining relevant prototypes weighted by their theoretically-expected positions in the corresponding subspace.

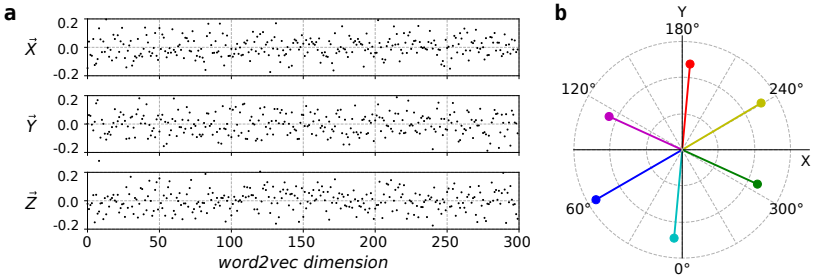


Fig. 3. a) components of Osgood's potency $\vec{X}$, activity $\vec{Y}$, and evaluation $\vec{Z}$ dimensions in 300-dimensional word2vec space, obtained from (4) and (11), b) projection (12) of six process-semantic prototypes (5) to the (uncorrected) XY plane. In the same color encoding, this layout closely aligns with the theoretical scheme in Figure 1(b). The difference is corrected by stretching the plane along the 150° - 330° direction as described in Section 3.1

2.3. Extraction of Osgood's semantics of arbitrary words. Any word of natural language is located in the XYZ space by projecting the corresponding word2vec representation $\vec{w}$ to its basis vectors:

$$\vec{s} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \vec{w} \cdot \vec{X} \\ \vec{w} \cdot \vec{Y} \\ \vec{w} \cdot \vec{Z} \end{bmatrix}, \tag{12}$$

where $\cdot$ denotes scalar product. This mapping of 300-dimensional vector $\vec{w}$ to 3-dimensional vector $\vec{s}$ finalizes the algorithm. The obtained components z, $-x$, and y are Osgood's semantic factors *evaluation*, *potency*, and *activity*, extracted from the word2vec data.

3. Experiment. The XYZ axes are obtained according to Sects. 2.1 and 2.2 are explicitly shown in Figure 3. Each axis appears to be a broad superposition of word2vec dimensions, indicating a non-trivial relation between word2vec model and Osgood's semantics.

To verify the above theory, it is necessary to check (i) if the prototypes (5) are actually located in the obtained XY plane according to the theoretical expectation shown in Figure 1(b), and (ii) if the Z prototypes (3), Figure 1(a), fall on different sides of this plane above and below the origin.

The first check is done by applying the first two lines of (12) to the prototypes (5). The obtained vectors $\vec{S}_{xy}$, shown in Figure 3(b), align with Figure 1(b) up to the following differences:

– prototypes *perception* (cyan) and *planning* (purple) deviate from their expected phases 0° and 120° towards the *analysis* prototype (blue) by 5.3° and 5.5°, respectively;

– prototypes *action* (red) and *evaluation* (green) deviate from their expected phases 180° and 300° towards the *progress* prototype (yellow) by 5.2° and 5.5°, respectively;

– *analysis* and *progress* prototypes have the largest vector lengths 0.25 and 0.23, compared to the other four being 0.21 on average.

3.1. Adjusting the X and Y axes. These differences might result from the asymmetry of the chosen prototypes, or of the word2vec model itself, possibly reflecting the process-semantic anisotropy of natural language. In any case, all of them are eliminated by squeezing the XY plane along the *analysis-progress* 60° - 240° direction by ≈ 1.2, or by stretching it along the orthogonal 150° - 330° direction by the same factor. The procedure also decreases the scalar product between the $\vec{X}$ and $\vec{Y}$ axes nearly twice, making them more orthogonal. Three pairwise scalar products become:

$$\vec{X} \cdot \vec{Y} = 0.036, \quad \vec{Y} \cdot \vec{Z} = 0.009, \quad \vec{Z} \cdot \vec{X} = 0.006,$$

showing nearly perfect orthogonality as expected for independent semantic dimensions. This correction is included in the following calculations.

3.2. Full map and features. The full semantic map of the prototypes $\vec{W}_i$ and their individual words $\vec{w}_i^{\,j}$ is obtained by projection (12) of the corresponding vectors to the adjusted XY plane:

$$\vec{S}_i = \vec{W}_i \cdot \vec{\Omega}, \qquad \vec{s}_i^{\,j} = \vec{w}_i^{\,j} \cdot \vec{\Omega}. \tag{13}$$

Figure 4 shows the resulting XY components of these vectors in the same color coding as in Figure 1(b). Unlike Figure 3(b), the prototypes are now centered at their theoretic angles (8) with a standard deviation of 0.6°. Individual terms of each prototype (Table 1) scatter around their means with standard angular deviations:

$$\Delta\phi_i = \sqrt{\frac{1}{N_i} \sum_{j=1}^{N_i} \left(\phi_i^j - \Phi_i\right)^2}, \tag{14}$$

amounting to 17° on average. The standard distance of individual words from each prototype center is indicated by semi-transparent circles of the corresponding radii and color.

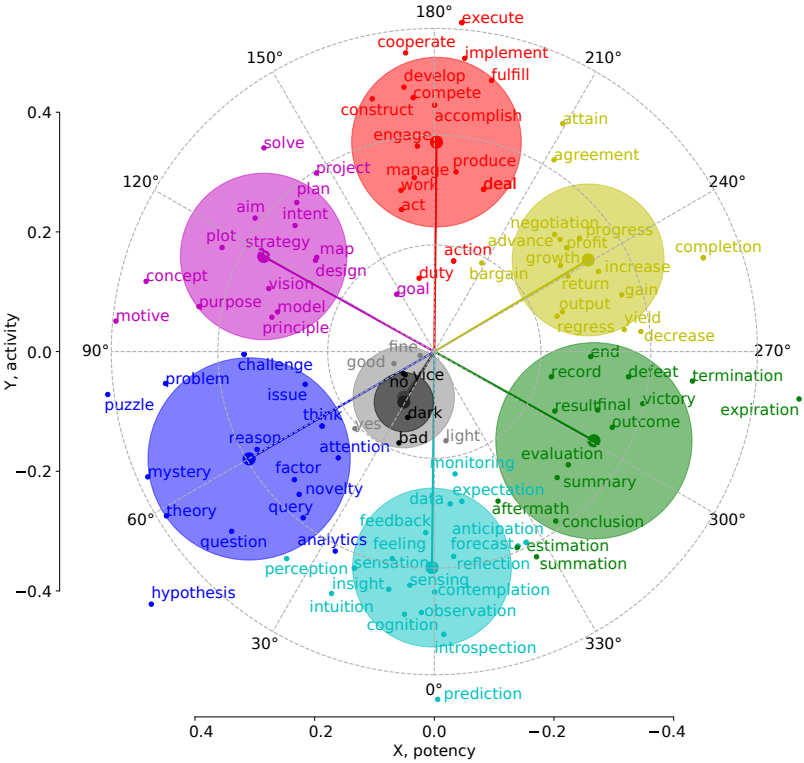


Fig. 4. Prototypes and their individual words (Table 1), projected to the potency-activity XY plane (11) via (12). Gray and black: “good” and “bad” prototypes (3). Process-stage prototypes (5): cyan - *perception*, blue - *analysis*, magenta - *planning*, red - *action*, yellow - *progress*, and green - *evaluation*, located in agreement with theoretical expectations shown in Figure 1(b). Semi-transparent circles indicate the scattering of individual terms in each prototype

With lengths in the XY plane 0.33 ± 0.025 and Z coordinates 0.024 ± 0.012, six XY prototypes are nearly coplanar as expected in theory, approximating the regular hexagon shown in Figure 1(b).

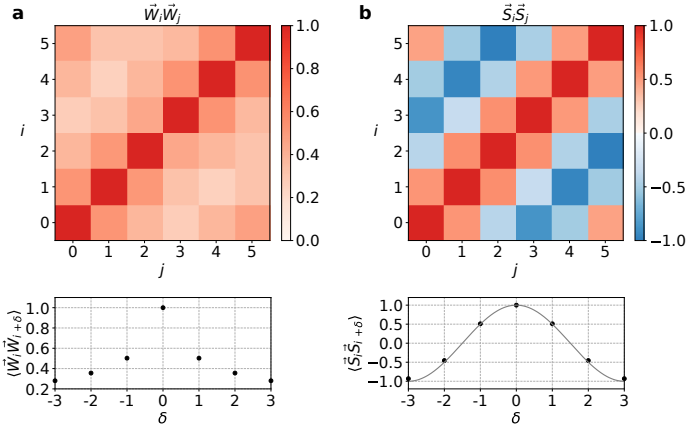


Fig. 5. Transformation of similarities among six XY prototypes (5) due to projection (12) from 300-dimensional word2vec space (a) to the semantic XYZ space (b). Top: pairwise scalar products of the prototype vectors. Bottom: the same as an averaged function of a stage-number difference $\delta = i - j$. All similarities among the original vectors (a) are positive, while in the process-semantic projection (b) similarities follow the harmonic function of the angular difference, as expected for circular layout (gray line)

"Good" and "bad" evaluation prototypes (3) are shown in gray and black. As expected, they are projected close to the origin of the XY plane with a displacement of ≈ 0.1 in the *perception* - *analysis* direction $\phi \approx 30°$. Z coordinates of these prototypes are 0.29 and -0.31, respectively. Both verification conditions, indicated at the beginning of this section, are thereby fulfilled.

Transformation of similarity Mapping of the XY prototypes from 300-dimensional word2vec representation $\vec{W}_k$ to the process plane (13) is illustrated by the transformation of their mutual similarity. According to their circular arrangement in the process plane, pairwise scalar products of three-dimensional prototype projections S_k follow the harmonic pattern shown in Figure 5(b). Initial prototype vectors $\vec{W}_k$ in 300-dimensional word2vec space, in contrast, all have positive scalar products shown in panel (a).

4. Testing. This section verifies the statistical significance of the obtained map and quantifies the ability of the method to predict semantic scores of individual words.

4.1. Randomization test. To verify the significance of the obtained map, the above procedure was performed for the same number of prototypes,

but composed of individual words sampled from Table 1 in a random way. The prototypes were ascribed with the same theoretical phases (8) and used to find the XY plane as before (11). The resulting map does not show the regularity demonstrated above. In contrast to Figure 4, prototypes are located at random phase angles and distances from the origin, while the scattering of individual terms (14) within prototypes is much larger than in the original map.

Finding a plane in which random word2vec vectors would take the prescribed angular positions Φ_i, therefore, does not seem to be possible. The regular structure appears only for semantically-coherent prototypes, composed in agreement with objective regularities of the word2vec data. The map shown in Figure 4 thereby reflects a non-trivial feature of natural language, rather than a mathematical artifact brought in by the construction method.

4.2. Mapping of novel words. In this test, 15 terms populating each context class according to Table 1 were divided into n "seed" and $15 - n$ "probe" items. The process-semantic plane (3) was then identified based on $6n$ seed terms, while the remaining $6(15 - n)$ probe terms were mapped to this plane by the procedure described above.

With n seed terms per semantic class randomly selected from Table 1, this procedure was repeated $m = 200$ times. For $n = 10$, the resulting scattering of $6m(15 - n) = 6000$ probe terms is shown in Figure 6(a). The mean of standard angular deviations (14) amounting to 35° indicates the ability of the method to correctly map novel words to the process-semantic plane, extracting the necessary information from the word2vec model.

Figure 6(b) shows the dependence of standard angular deviation (14) and mean length of the prototype vectors $\langle S_i \rangle$ (13) on the seed size n. With increasing n, semantic features of individual words average out more efficiently, suppressing angular noise $\Delta\phi$ and increasing process-semantic coherence of the prototypes $\langle S_i \rangle$. Angular discrimination threshold of $\Delta\phi \approx 30^\circ$ is reached near $n = 10$ when the mean scattering radius $\langle R \rangle$ drops below one-half on the mean length $\langle S_i \rangle$. The map in panel (a) is close to this borderline regime.

5. Discussion.

5.1. Natural semiotics in ML. The theoretical structure of the prototypes used in Section 2 is not constructed specifically for the purpose of this paper, but was previously identified from a cognitive-semiotic perspective. In particular, the process-stage prototypes, shown in Figure 1(b), form an abstract template for causal-predictive simulation of behavior, structurally isomorphic to periodic processes in Nature like year- and day-night cycles [21]. Along with the evaluation dimension, the resulting spherical space is considered as a core semiotic system of living Nature, underlying natural sense-making from single-cell organisms to humans (ibid.).

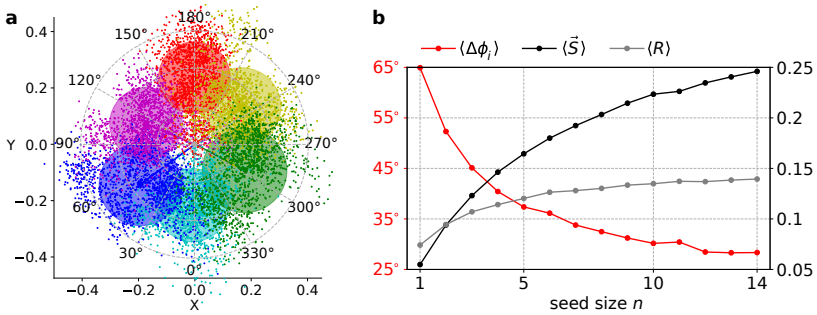


Fig. 6. Prediction of the process-semantics phase of individual words. Lists of individual terms for six XY prototypes in Table 1 are split into two parts so that the "seed" part of the words is used to find the XY plane, while the remaining "probe" part is used for testing. a: In each class of 15 individual words, 10 randomly selected ones are used as seed and the rest is used for testing. The result of 200 runs is shown. b: a mean of the standard angular deviations (14), a mean of the lengths of the prototype vectors $\langle\vec{S}\rangle$, and a mean of the scattering radii $\langle R\rangle$ as a function of the number of seed words

Observing the expected theoretical structure in Figure 4 provides insight into the general principles of ML and AI. Although aimed at a simple predictive task (recovery of skipped words based on their surroundings), efficiency required word2vec to accurately reflect the spatial structure of this highly abstract semiotic alphabet. Machine-learning image of natural language thus extends to the level of universally interpretable affective meaning, achieving deeper correspondence with cognitive semantics than was previously known [23]. Central features of the interpretable semantic map, recognized as a key element of next-generation biologically-inspired AI [18], thus appear to be already in place.

The established isomorphism can also be harnessed in novel architectures of AI and ML. If an efficient model is bound to reflect fundamental regularities of Nature, knowing the latter facilitates the design of the former. Instead of converging to such regularities in a blind search, they could be hardwired into the neural architectures from the start, reducing the dimensionality of the training optimization without loss in quality. The resulting economy of computational resources opens prospects for better replication of natural cognition and learning [24–26], as advocated e.g. in the field of information retrieval [27]. Generalization of the developed approach to other vector representations [14, 28–30] and machine-learning models [15, 31] allows going beyond the simplest case considered in this paper.

5.2. Affectively interpretable AI. The obtained mapping contributes to solving the interpretability problem noted in the Introduction, that is, the inability to express internal states of the black box-type algorithms in sensible categories like that of natural language. Evaluation, potency, and activity dimensions are such categories, unique in their affective nature and cross-linguistic universality [32–34]. Due to the centrality of affective meaning in human cognition [35, 36], remarkably overlooked in recent surveys on explainable AI [37–44], these dimensions are basic for making the internal workings of such algorithms available for human inspection. Further, this might be used to align AI and ML algorithms with the principles of commonsense reason, enabling computing in natural categories of human thought [45, 46].

By finding Osgood's semantics in the baseline machine-learning model, the reported approach also opens a prospect for developing explainability tests for other complex algorithms. In the decision-support systems, for example, it could be used to "look inside" the black box [47] and observe or correct the affective state it simulates towards a target entity[3]. The concurrence of such a state with human ethics and reason could supplement other certification criteria [53].

5.3. Use for cognitive modeling. In the reverse direction, the established link between Osgood's semantics and word2vec data allows using the latter for cognitive modeling and research. According to Osgood's original method [12] and its successors, 50 to 80 percent of judgment data variance, defined by *evaluation*, *potency*, and *activity* factors, could be extracted directly from word2vec representation of situations and things. Based on that, judgment and decision probabilities of interest could be predicted without performing real-world experiments as envisioned by [54–56].

Besides speed and cost advantages, this approach is also expected to be higher in precision, since word2vec models (trained on huge corpora of texts) accumulate much more information, than usually collected in old-style experiments. Through subtle regularities of natural language, this also includes implicit knowledge, hardly observable in laboratory conditions.

6. Conclusion. The possibility to retrieve Osgood's semantics from the word2vec data shows that the most agnostic models of data science converge to the basic principles of natural thinking, previously revealed in cognitive and semiotic studies. After such validation, these principles facilitate finding of nature-inspired solutions for hard problems in computer science. The interpretability problem of AI, for example, might be not as hard as seen

[3]This is not to be mistaken with an affective state of a machine itself, sometimes ascribed to it within a so-called intentional stance [48] exemplifying cognitive fallacy to ensoul complex systems [49–52].

from a brute-force computational paradigm, dominating the field today. If the reported method could be extended to other ML models, the expainability of the present black-box AI might be approached by a minor add-on, analogous to the projection procedure described in this paper.

References

1. Mikolov T., Yih W., Zweig G. Linguistic Regularities in Continuous Space Word Representations. Proceedings of NAACL-HLT. 2013. pp. 746–751.
2. Pennington J., Socher R., Manning C.D. Glove: Global vectors for word representation. Proceedings of the 2014 conference on empirical methods in natural language processing. 2014. pp. 1532–1543.
3. Radford A., Narasimhan K., Salimans T., Sutskever I. Improving Language Understanding by Generative Pre-Training. 2018.
4. Yang Z., Dai Z., Yang Y., Carbonell J., Salakhutdinov R., Le Q.V. XLNet: Generalized autoregressive pretraining for language understanding. Proceedings of 33rd Conference on Neural Information Processing Systems. 2019.
5. Mikolov T., Joulin A., Baroni M. A Roadmap Towards Machine Intelligence. *Computational Linguistics and Intelligent Text Processing.* Cham: Springer, 2018. pp. 29–61.
6. Al-Saqqa S., Awajan A. The Use of Word2vec Model in Sentiment Analysis: A Survey. ACM International Conference Proceeding Series. 2019. pp. 39–43.
7. Dhar A., Mukherjee H., Dash N.S., Roy K. Text categorization: past and present. *Artificial Intelligence Review.* 2021. vol. 54. no. 4. pp. 3007–3054.
8. Konstantinov A., Moshkin V., Yarushkina N. Approach to the Use of Language Models BERT and Word2vec in Sentiment Analysis of Social Network Texts. *Recent Research in Control Engineering and Decision Making.* Cham: Springer, 2021. pp. 462–473.
9. Gunning D., Stefik M., Choi J., Miller T., Stumpf S., Yang G.Z. XAI—Explainable artificial intelligence. *Science Robotics*. 2019. vol. 4. no. 37.
10. Suvorova A. Interpretable Machine Learning in Social Sciences: Use Cases and Limitations. *Proceedings of Digital Transformation and Global Society 2021. Communications in Computer and Information Science, vol. 1503*. Cham: Springer, 2022. pp. 319–331.
11. Osgood C.E. The nature and measurement of meaning. *Psychological Bulletin*. 1952. vol. 49. no. 3. pp. 197–237.
12. Osgood C.E. Studies on the generality of affective meaning systems. *American Psychologist*. 1962. vol.17. no.1. pp. 10–28.
13. Hollis G., Westbury C. The principals of meaning: Extracting semantic dimensions from co-occurrence models of semantics. *Psychonomic Bulletin and Review*. 2016. vol. 23. no. 6. pp. 1744–1756.
14. Lenci A. Distributional Models of Word Meaning. *Annual Review of Linguistics*. 2018. vol. 4. no. 1. pp. 151–171.
15. Günther F., Rinaldi L., Marelli M. Vector-Space Models of Semantic Representation From a Cognitive Perspective: A Discussion of Common Misconceptions. *Perspectives on Psychological Science*. 2019. vol. 14. no. 6. pp. 1006–1033.
16. Samsonovich A.V., Ascoli G.A. Principal Semantic Components of Language and the Measurement of Meaning. *PLoS ONE*. 2010. vol. 5. no. 6.
17. Eidlin A.A., Eidlina M.A., Samsonovich A.V. Analyzing weak semantic map of word senses. *Procedia Computer Science*. 2018. vol. 123. pp. 140–148.

18. Samsonovich A.V. On semantic map as a key component in socially-emotional BICA. *Biologically Inspired Cognitive Architectures*. 2018. vol. 23. pp. 1–6.
19. Pretrained word2vec model “GoogleNews-vectors-negative300.bin.gz”. Google Code Archive. https://code.google.com/archive/p/word2vec/. 2013.
20. Mikolov T., Sutskever I., Chen K., Corrado G., Dean J. Distributed representations of words and phrases and their compositionality. Advances in Neural Information Processing Systems. 2013.
21. Surov I.A. Natural Code of Subjective Experience. *Biosemiotics*. 2022. vol. 15. no. 1. pp. 109–139.
22. Siegel M. The sense-think-act paradigm revisited. Proceedings of the 1st International Workshop on Robotic Sensing. 2003.
23. Gastaldi J.L. Why Can Computers Understand Natural Language? *Philosophy & Technology*. 2021. vol. 34. no. 1. pp. 149–214.
24. Jensen A.R. The relationship between learning and intelligence. *Learning and Individual Differences*. 1989. vol. 1. no. 1. pp. 37–62.
25. Sowa J.F. The Cognitive Cycle. Proceedings of the 2015 Federated Conference on Computer Science and Information Systems. 2015. vol. 5, pp. 11–16.
26. Wang Y., Yao Q., Kwok J.T., Ni L.M.: Generalizing from a Few Examples: A Survey on Few-shot Learning. *ACM Computing Surveys*. 2021. vol. 53. no. 3. pp. 1–34.
27. Hoenkamp E. Why Information Retrieval Needs Cognitive Science: A Call to Arms. 2005.
28. Turney P.D., Pantel P. From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research*. 2010. vol. 37. pp. 141–188.
29. Wang B., Buccio E.D., Melucci M. Representing Words in Vector Space and Beyond. *Quantum-Like Models for Information Retrieval and Decision-Making* (eds: Aerts D., Khrennikov A., Melucci M., Toni B.). Cham, Springer. pp. 83–113. 2019.
30. beim Graben P., Huber M., Meyer W., Römer R., Wolff M. Vector Symbolic Architectures for Context-Free Grammars. *Cognitive Computation*. 2022. vol. 14. no. 2. pp. 733–748.
31. Coenen A., Reif E., Kim A.Y.B., Pearce A., Viégas F., Wattenberg M. Visualizing and measuring the geometry of BERT. Proceedings of the Advances in Neural Information Processing Systems. 2019.
32. Tanaka Y., Oyama T., Osgood C.E. A cross-culture and cross-concept study of the generality of semantic spaces. *Journal of Verbal Learning and Verbal Behavior*. 1963. vol. 2. no. 5-6. pp. 392–405.
33. Tanaka Y., Osgood C.E. Cross-culture, cross-concept, and cross-subject generality of affective meaning systems. *Journal of Personality and Social Psychology*. 1965. vol. 2. no. 2. pp. 143–153.
34. Osgood C.E., May W.H., Miron M.S. *Cross-cultural universals of affective meaning*. Champaign, University of Illinois Press. 1975.
35. Zajonc R.B. Feeling and thinking: Preferences need no inferences. *American Psychologist*. 1980. vol. 35. no. 2. pp. 151–175.
36. Duncan S., Barrett L.F. Affect is a form of cognition: A neurobiological analysis. *Cognition and Emotion*. 2007. vol. 21, no. 6. pp. 1184–1211.
37. Lipton Z.C. The Mythos of Model Interpretability. *Queue*. 2018. vol. 3. pp. 31–57.
38. Molnar C. Interpretable Machine Learning. A Guide for Making Black Box Models Explainable. 2019.
39. Guidotti R., Monreale A., Ruggieri S., Turini F., Giannotti F., Pedreschi D. A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys*. 2019. vol. 51. no. 5.

40. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*. 2019. vol. 1. no. 5. pp. 206–215.
41. Vassiliades A., Bassiliades N., Patkos T. Argumentation and explainable artificial intelligence: A survey // *Knowledge Engineering Review*. 2021. vol. 36. pp. 1–35.
42. Borrego-Díaz J., Galán-Páez J. Explainable Artificial Intelligence in Data Science. *Minds and Machines*. 2022.
43. Chou Y.L., Moreira C., Bruza P., Ouyang C., Jorge J. Counterfactuals and causability in explainable artificial intelligence: Theory, algorithms, and applications. *Information Fusion*. 2022. vol. 81. pp. 59–83.
44. Tian L., Oviatt S., Muszunski M., Chamberlain B.C., Healey J., Sano, A. Applied Affective Computing. ACM Books. 2022.
45. Michelucci P. (ed.) Handbook of Human Computation. New York, Springer. 2013.
46. Samsonovich A.V. (ed.) Biologically Inspired Cognitive Architectures. Advances in Intelligent Systems and Computing vol. 948. Cham, Springer. 2020.
47. Adadi A., Berrada M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*. 2018. vol. 6. no. 52. pp. 138–152.
48. Dennett D.C. The intentional stance. Cambridge, MIT Press. 1998.
49. Caporael L.R. Anthropomorphism and mechanomorphism: Two faces of the human machine. *Computers in Human Behavior*. 1986. vol. 2. no. 3. pp. 215–234.
50. Guthrie S.E. Anthropomorphism: A definition and a theory. *Anthropomorphism, anecdotes, and animals*. (eds. Mitchell R.W., Thomson N.S., Miles H.L.), chap. 5, pp. 50-58. State University of New York Press, New York. 1997.
51. Watson D. The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence. *Minds and Machines.* 2019. vol. 29, no. 3. pp. 417–440.
52. Salles A., Evers K., Farisco M. Anthropomorphism in AI. *AJOB Neurosci*. 2020. vol. 11. no. 2. pp. 88–95.
53. Maclure J. AI, Explainability and Public Reason: The Argument from the Limitations of the Human Mind. *Minds and Machines*. 2021.
54. Arnulf J.K., Larsen K.R., Martinsen Ø.L., Bong C.H. Predicting survey responses: How and why semantics shape survey statistics on Organizational Behaviour. *PLoS ONE*. 2014. vol. 9. no. 9.
55. Jones M.N., Gruenenfelder T.M., Recchia G. In defense of spatial models of semantic representation. *New Ideas in Psychology*. 2018. vol. 50. pp. 54–60.
56. Arnulf J.K. Wittgenstein's Revenge: How Semantic Algorithms Can Help Survey Research Escape Smedslund's Labyrinth. *Respect for Thought* (eds. Lindstad T.G., Stänicke E., Valsiner J.), chap. 17, pp. 285–307. Springer, Cham. 2020.

Surov Ilya — Ph.D., Associate Professor, Senior researcher, ITMO University. Research interests: cognitive-behavioral modeling, quantum semiotics and semantics. The number of publications — 25. ilya.a.surov@itmo.ru; 49A, Kronverksky Av., 197101, St. Petersburg, Russia; office phone: +7(812)480-0000.

Acknowledgements. This research is supported by RNF (grant № 20-71-00136).

УДК 004.8 **DOI** 10.15622/ia.21.5.3

И.А. Суров

ОТКРЫТИЕ ЧЁРНОГО ЯЩИКА: ИЗВЛЕЧЕНИЕ СЕМАНТИЧЕСКИХ ФАКТОРОВ ОСГУДА ИЗ ЯЗЫКОВОЙ МОДЕЛИ WORD2VEC

Суров И.А. **Открытие чёрного ящика: Извлечение семантических факторов Осгуда из языковой модели word2vec.**

Аннотация. Современные модели искусственного интеллекта развиваются в парадигме чёрного ящика, когда значима только информация на входе и выходе системы, тогда как внутренние представления интерпретации не имеют. Такие модели не обладают качествами объяснимости и прозрачности, необходимыми во многих задачах. Статья направлена на решение данной проблемы путём нахождения семантических факторов Ч. Осгуда в базовой модели машинного обучения word2vec, представляющей слова естественного языка в виде 300-мерных неинтерпретируемых векторов. Искомые факторы определяются на основе восьми семантических прототипов, составленных из отдельных слов. Ось оценки в пространстве word2vec находится как разность между положительным и отрицательным прототипами. Оси силы и активности находятся на основе шести процессно-семантических прототипов (восприятие, анализ, планирование, действие, прогресс, оценка), представляющих фазы обобщённого кругового процесса в данной плоскости. Направления всех трёх осей в пространстве word2vec найдены в простой аналитической форме, не требующей дополнительного обучения. Как и ожидается для независимых семантических факторов, полученные направления близки к попарной ортогональности. Значения семантических факторов для любого объекта word2vec находятся с помощью простой проективной операции на найденные направления. В соответствии с требованиями к объяснимому ИИ, представленный результат открывает возможность для интерпретации содержимого алгоритмов типа “чёрный ящик” в естественных эмоционально-смысловых категориях. В обратную стонону, разработанный подход позволяет использовать модели машинного обучения в качестве источника данных для когнитивно-поведенческого моделирования.

Ключевые слова: аффект, семантика, пространство, Осгуд, смысл, язык, word2vec, чёрный ящик, объяснимость, интерпретация

Литература

1. Mikolov T., Yih W., Zweig G. Linguistic Regularities in Continuous Space Word Representations. Proceedings of NAACL-HLT. 2013. pp. 746–751.
2. Pennington J., Socher R., Manning C.D. Glove: Global vectors for word representation. Proceedings of the 2014 conference on empirical methods in natural language processing. 2014. pp. 1532–1543.
3. Radford A., Narasimhan K., Salimans T., Sutskever I. Improving Language Understanding by Generative Pre-Training. 2018.
4. Yang Z., Dai Z., Yang Y., Carbonell J., Salakhutdinov R., Le Q.V. XLNet: Generalized autoregressive pretraining for language understanding. Proceedings of 33rd Conference on Neural Information Processing Systems. 2019.
5. Mikolov T., Joulin A., Baroni M. A Roadmap Towards Machine Intelligence. *Computational Linguistics and Intelligent Text Processing*. Cham: Springer, 2018. pp. 29–61.

6. Al-Saqqa S., Awajan A. The Use of Word2vec Model in Sentiment Analysis: A Survey. ACM International Conference Proceeding Series. 2019. pp. 39–43.
7. Dhar A., Mukherjee H., Dash N.S., Roy K. Text categorization: past and present. *Artificial Intelligence Review.* 2021. vol. 54. no. 4. pp. 3007–3054.
8. Konstantinov A., Moshkin V., Yarushkina N. Approach to the Use of Language Models BERT and Word2vec in Sentiment Analysis of Social Network Texts. *Recent Research in Control Engineering and Decision Making.* Cham: Springer, 2021. pp. 462–473.
9. Gunning D., Stefik M., Choi J., Miller T., Stumpf S., Yang G.Z. XAI—Explainable artificial intelligence. *Science Robotics*. 2019. vol. 4. no. 37.
10. Suvorova A. Interpretable Machine Learning in Social Sciences: Use Cases and Limitations. *Proceedings of Digital Transformation and Global Society 2021. Communications in Computer and Information Science, vol. 1503*. Cham: Springer, 2022. pp. 319–331.
11. Osgood C.E. The nature and measurement of meaning. *Psychological Bulletin*. 1952. vol. 49. no. 3. pp. 197–237.
12. Osgood C.E. Studies on the generality of affective meaning systems. *American Psychologist*. 1962. vol.17. no.1. pp. 10–28.
13. Hollis G., Westbury C. The principals of meaning: Extracting semantic dimensions from co-occurrence models of semantics. *Psychonomic Bulletin and Review*. 2016. vol. 23. no. 6. pp. 1744–1756.
14. Lenci A. Distributional Models of Word Meaning. *Annual Review of Linguistics*. 2018. vol. 4. no. 1. pp. 151–171.
15. Günther F., Rinaldi L., Marelli M. Vector-Space Models of Semantic Representation From a Cognitive Perspective: A Discussion of Common Misconceptions. *Perspectives on Psychological Science*. 2019. vol. 14. no. 6. pp. 1006–1033.
16. Samsonovich A.V., Ascoli G.A. Principal Semantic Components of Language and the Measurement of Meaning. *PLoS ONE*. 2010. vol. 5. no. 6.
17. Eidlin A.A., Eidlina M.A., Samsonovich A.V. Analyzing weak semantic map of word senses. *Procedia Computer Science*. 2018. vol. 123. pp. 140–148.
18. Samsonovich A.V. On semantic map as a key component in socially-emotional BICA. *Biologically Inspired Cognitive Architectures*. 2018. vol. 23. pp. 1–6.
19. Pretrained word2vec model “GoogleNews-vectors-negative300.bin.gz”. Google Code Archive. https://code.google.com/archive/p/word2vec/. 2013.
20. Mikolov T., Sutskever I., Chen K., Corrado G., Dean J. Distributed representations of words and phrases and their compositionality. Advances in Neural Information Processing Systems. 2013.
21. Surov I.A. Natural Code of Subjective Experience. *Biosemiotics*. 2022. vol. 15. no. 1. pp. 109–139.
22. Siegel M. The sense-think-act paradigm revisited. Proceedings of the 1st International Workshop on Robotic Sensing. 2003.
23. Gastaldi J.L. Why Can Computers Understand Natural Language? *Philosophy & Technology*. 2021. vol. 34. no. 1. pp. 149–214.
24. Jensen A.R. The relationship between learning and intelligence. *Learning and Individual Differences*. 1989. vol. 1. no. 1. pp. 37–62.
25. Sowa J.F. The Cognitive Cycle. Proceedings of the 2015 Federated Conference on Computer Science and Information Systems. 2015. vol. 5, pp. 11–16.
26. Wang Y., Yao Q., Kwok J.T., Ni L.M.: Generalizing from a Few Examples: A Survey on Few-shot Learning. *ACM Computing Surveys*. 2021. vol. 53. no. 3. pp. 1–34.

27. Hoenkamp E. Why Information Retrieval Needs Cognitive Science: A Call to Arms. 2005.
28. Turney P.D., Pantel P. From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research*. 2010. vol. 37. pp. 141–188.
29. Wang B., Buccio E.D., Melucci M. Representing Words in Vector Space and Beyond. *Quantum-Like Models for Information Retrieval and Decision-Making* (eds: Aerts D., Khrennikov A., Melucci M., Toni B.). Cham, Springer. pp. 83–113. 2019.
30. beim Graben P., Huber M., Meyer W., Römer R., Wolff M. Vector Symbolic Architectures for Context-Free Grammars. *Cognitive Computation*. 2022. vol. 14. no. 2. pp. 733–748.
31. Coenen A., Reif E., Kim A.Y.B., Pearce A., Viégas F., Wattenberg M. Visualizing and measuring the geometry of BERT. Proceedings of the Advances in Neural Information Processing Systems. 2019.
32. Tanaka Y., Oyama T., Osgood C.E. A cross-culture and cross-concept study of the generality of semantic spaces. *Journal of Verbal Learning and Verbal Behavior*. 1963. vol. 2. no. 5-6. pp. 392–405.
33. Tanaka Y., Osgood C.E. Cross-culture, cross-concept, and cross-subject generality of affective meaning systems. *Journal of Personality and Social Psychology*. 1965. vol. 2. no. 2. pp. 143–153.
34. Osgood C.E., May W.H., Miron M.S. *Cross-cultural universals of affective meaning*. Champaign, University of Illinois Press. 1975.
35. Zajonc R.B. Feeling and thinking: Preferences need no inferences. *American Psychologist*. 1980. vol. 35. no. 2. pp. 151–175.
36. Duncan S., Barrett L.F. Affect is a form of cognition: A neurobiological analysis. *Cognition and Emotion*. 2007. vol. 21, no. 6. pp. 1184–1211.
37. Lipton Z.C. The Mythos of Model Interpretability. *Queue*. 2018. vol. 3. pp. 31–57.
38. Molnar C. Interpretable Machine Learning. A Guide for Making Black Box Models Explainable. 2019.
39. Guidotti R., Monreale A., Ruggieri S., Turini F., Giannotti F., Pedreschi D. A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys*. 2019. vol. 51. no. 5.
40. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*. 2019. vol. 1. no. 5. pp. 206–215.
41. Vassiliades A., Bassiliades N., Patkos T. Argumentation and explainable artificial intelligence: A survey // *Knowledge Engineering Review*. 2021. vol. 36, pp. 1-35.
42. Borrego-Díaz J., Galán-Páez J. Explainable Artificial Intelligence in Data Science. *Minds and Machines*. 2022.
43. Chou Y.L., Moreira C., Bruza P., Ouyang C., Jorge J. Counterfactuals and causability in explainable artificial intelligence: Theory, algorithms, and applications. *Information Fusion*. 2022. vol. 81. pp. 59–83.
44. Tian L., Oviatt S., Muszunski M., Chamberlain B.C., Healey J., Sano, A. Applied Affective Computing. ACM Books. 2022.
45. Michelucci P. (ed.) Handbook of Human Computation. New York, Springer. 2013.
46. Samsonovich A.V. (ed.) Biologically Inspired Cognitive Architectures. Advances in Intelligent Systems and Computing vol. 948. Cham, Springer. 2020.
47. Adadi A., Berrada M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*. 2018. vol. 6. no. 52. pp. 138–152.
48. Dennett D.C. The intentional stance. Cambridge, MIT Press. 1998.

49. Caporael L.R. Anthropomorphism and mechanomorphism: Two faces of the human machine. *Computers in Human Behavior*. 1986. vol. 2. no. 3. pp. 215–234.
50. Guthrie S.E. Anthropomorphism: A definition and a theory. *Anthropomorphism, anecdotes, and animals*. (eds. Mitchell R.W., Thomson N.S., Miles H.L.), chap. 5, pp. 50-58. State University of New York Press, New York. 1997.
51. Watson D. The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence. *Minds and Machines*. 2019. vol. 29, no. 3. pp. 417–440.
52. Salles A., Evers K., Farisco M. Anthropomorphism in AI. *AJOB Neurosci*. 2020. vol. 11. no. 2. pp. 88–95.
53. Maclure J. AI, Explainability and Public Reason: The Argument from the Limitations of the Human Mind. *Minds and Machines*. 2021.
54. Arnulf J.K., Larsen K.R., Martinsen Ø.L., Bong C.H. Predicting survey responses: How and why semantics shape survey statistics on Organizational Behaviour. *PLoS ONE*. 2014. vol. 9. no. 9.
55. Jones M.N., Gruenenfelder T.M., Recchia G. In defense of spatial models of semantic representation. *New Ideas in Psychology*. 2018. vol. 50. pp. 54–60.
56. Arnulf J.K. Wittgenstein's Revenge: How Semantic Algorithms Can Help Survey Research Escape Smedslund's Labyrinth. *Respect for Thought* (eds. Lindstad T.G., Stänicke E., Valsiner J.), chap. 17, pp. 285–307. Springer, Cham. 2020.

Суров Илья Алексеевич — канд. физ.-мат. наук, доцент, старший научный сотрудник, Университет ИТМО. Область научных интересов: когнитивно-поведенческое моделирование, квантовая семиотика и семантика. Число научных публикаций — 25. surov.i.a@yandex.ru; Кронверкский проспект, 49А, 197101, Санкт-Петербург, Россия; р.т.: +7(812)480-0000.

Поддержка исследований. Работа выполнена при финансовой поддержке РНФ (проект № 20-71-00136).

УДК 004.093 **DOI** 10.15622/ia.21.5.4

М.Н. ФАВОРСКАЯ, Н. НИШЧХАЛ

ВЕРИФИКАЦИЯ РАЗЛИВОВ НЕФТИ НА ВОДНЫХ ПОВЕРХНОСТЯХ ПО АЭРОФОТОСНИМКАМ НА ОСНОВЕ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ

Фаворская М.Н., Нишчхал Н. **Верификация разливов нефти на водных поверхностях по аэрофотоснимкам на основе методов глубокого обучения.**

Аннотация. В статье решается задача верификации разливов нефти на водных поверхностях рек, морей и океанов по оптическим аэрофотоснимкам с использованием методов глубокого обучения. Особенностью данной задачи является наличие визуально похожих на разливы нефти областей на водных поверхностях, вызванных цветением водорослей, веществ, не приносящих экологический ущерб (например, пальмовое масло), бликов при съемке или природных явлений (так называемые «двойники»). Многие исследования в данной области основаны на анализе изображений, полученных от радаров с синтезированной апертурой (Synthetic Aperture Radar (SAR) images), которые не обеспечивают точной классификации и сегментации. Последующая верификация способствует сокращению экологического и материального ущерба, а мониторинг размеров площади нефтяного пятна используется для принятия дальнейших решений по устранению последствий. Предлагается новый подход к верификации оптических снимков как задачи бинарной классификации на основе сиамской сети, когда фрагмент исходного изображения многократно сравнивается с репрезентативными примерами из класса нефтяных пятен на водных поверхностях. Основой сиамской сети служит облегченная сеть VGG16. При превышении порогового значения выходной функции принимается решение о наличии разлива нефти. Для обучения сети был собран и размечен собственный набор данных из открытых интернет-ресурсов. Существенной проблемой является несбалансированность выборки данных по классам, что потребовало применения методов аугментации, основанных не только на геометрических и цветовых манипуляциях, но и на основе генеративной состязательной сети (Generative Adversarial Network, GAN). Эксперименты показали, что точность классификации разливов нефти и «двойников» на тестовой выборке достигает значений 0,91 и 0,834 соответственно. Далее решается дополнительная задача семантической сегментации нефтяного пятна с применением сверточных нейронных сетей (СНС) типа кодировщик-декодировщик. Для сегментации исследовались три архитектуры глубоких сетей, а именно U-Net, SegNet и Poly-YOLOv3. Лучшие результаты показала сеть Poly-YOLOv3, достигнув точности 0,97 при среднем времени обработки снимка 385 с веб-сервисом Google Colab. Также была спроектирована база данных для хранения исходных и верифицированных изображений с проблемными областями.

Ключевые слова: обнаружение разливов нефти, верификация, сегментация, глубокое обучение, аэрофотоснимки, дистанционное зондирование Земли.

1. Введение. Сырая нефть и нефтепродукты являются широко распространенными загрязнителями воды и почвы в результате разливов на водных поверхностях и на суше, динамика которых различна. Разливы нефти на водных поверхностях являются последствиями аварий на танкерах, судах, трубопроводах и нефтяных платформах, когда сырая нефть, бензин, топливо или побочные

продукты нефтепереработки сбрасываются в воду. Международная статистика по объемам разливов нефти свидетельствует о том, что большинство разливов нефти небольшие (менее 7 тонн), а крупные аварии составляют малую долю от общего количества нефти, попадающей в окружающую среду [1]. Однако даже небольшие аварийные выбросы сильно загрязняют локальные территории и не исключены пути, по которым нефть может попасть обратно к людям через накопления в рыбе или потребление загрязненных подземных вод.

Поскольку в настоящее время разведка нефти ведется в более глубоких водах и на более удаленных территориях, куда сложно попасть для своевременной очистки водных поверхностей, риск будущих аварий становится намного выше. Следует отметить, что из-за случайного характера разливов нефти их полное предотвращение невозможно. Атмосферные и водные условия (температура, ветер, течение, соленость) могут значительно увеличить перенос нефти и скорость выветривания. Следовательно, поведение и экологические последствия разливов в море или океане непредсказуемы и неопределенны [2], что приводит к необходимости построения модели предсказания поведения нефтяного пятна [3]. Воздушные, поверхностные и подповерхностные химические измерения показали, что около 5% массы выбрасываемых углеводородов испаряется в атмосферу, 10% составляет поверхностное пятно, а остаток растворяется или рассеивается в толще воды. Причем около одной трети углеводородов непосредственно находится в глубоких подводных течениях или накапливается в отложениях [4].

Процесс устранения последствий аварии может длиться долго, поэтому необходим мониторинг распространения разливов нефти для оценки влияния на экологическую безопасность окружающей среды. Общепринятым подходом является использование космических снимков или снимков, полученных от камер беспилотных летательных аппаратов или средств малой авиации, для обнаружения и мониторинга аварийной ситуации.

В данном исследовании решаются две связанные задачи – верификация разливов нефти на водных поверхностях и их сегментация с целью мониторинга последствий аварии. Результаты записываются в базу данных, что позволяет получать статистические данные, строить графики в указанный пользователем временной интервал.

2. Обзор существующих методов. Основными исходными данными для обнаружения и мониторинга разливов нефти

на водных поверхностях являются радарные и гиперспектральные / мультиспектральные космические снимки. При этом анализу радарных снимков посвящено большинство работ из-за основного преимущества – получение SAR снимков не подвержено метеорологическим факторам. Однако радарные снимки имеют существенные недостатки в виде спекл-шумов и ограничений в обнаружении относительно тонких нефтяных пленок. Также по радарным снимкам невозможно отличить биогенные пятна (например, цветение водорослей или другой растительности) и природные явления, вызванные ветром, течениями, водоворотами, дождями и кильватерными следами кораблей (называемых в литературе «двойниками»), от разливов нефти [5, 6]. Улучшенные оценки распознавания показывают методы, обрабатывающие поляризованные SAR изображения [7, 8]. Отметим, что объективно как традиционные методы машинного обучения, так и методы глубокого обучения, применяемые для обработки радарных снимков, характеризуются худшими показателями относительно других видов снимков.

Если гиперспектральная визуализация аккумулирует излучение всей спектральной полосы для каждого пиксела изображения, то мультиспектральная визуализация характеризуется коэффициентами отражениями ограниченного числа длин волн, связанных с видимой, инфракрасной и ультрафиолетовой частями гиперспектральных изображений. Их характеристики можно использовать как для обнаружения разливов нефти и для выявления «двойников» (эффекты флуоресценции, растительность и пр.). Перспективным подходом является совместное использование в алгоритмах как гиперспектральных, так и радарных космических снимков [9]. Оптические снимки предоставляют лучшие возможности для классификации и сегментации, однако существуют типичные ограничения в их получении средствами дистанционного зондирования Земли (ДЗЗ) (отсутствие облачности, дневное время суток). Поэтому оптические снимки, в основном, получают с использованием средств малой авиации, обеспечивая устойчивый мониторинг и относительно низкую стоимость их получения [10].

Эволюцию методов оценки разливов нефти можно сформулировать следующим образом. Классические методы машинного обучения, такие как линейный дискриминантный анализ, множественная линейная регрессия, кластеризация k-средних, методы логистической регрессии, деревья решений и т.д., применялись в 2000–2005 гг. В последующие годы стали использоваться усовершенствованные модели классификации (нечеткая

кластеризация, метаэвристическая оптимизация и модели на основе искусственные нейронных сетей). Быстрое распространение глубоких нейронных сетей, начиная с 2012 г., позволило объединить методы семантической и объектной сегментации SAR изображений с целью более точной классификации разливов нефти. В работе [7] многоуровневый автокодировщик и сеть глубокого доверия оптимизировали наборы признаков поляризованных SAR изображений за счет послойного неконтролируемого предварительного обучения. Глубокая сверточная нейронная сеть с архитектурой, подобной семейству VGG-сетей, предложена в работе [6]. При этом значения верности (accuracy), полноты (recall) и точности (precision) составили 94,01%, 83,51%, и 85,7% для сегментации разливов нефти на SAR изображениях. Сеть глубокого обучения для обнаружения и категоризации разливов нефти на SAR изображениях в крупном масштабе на основе архитектуры U-Net предложена в работе [8]. Недавно предложенная модель BO-DRNet [11] основана на байесовской оптимизации гиперпараметров и архитектурах ResNet-18 и DeepLabv3+. Тем не менее, исследователи не ограничиваются применением методов глубокого обучения. Так, в работе [12] использована комбинация многоцелевой оптимизации алгоритма серых волков и *k*-средних для поиска наилучшего порогового уровня для сегментации SAR изображений.

Методов, относящихся к обработке оптических снимков для решения данной задачи, намного меньше. В работе [13] разработан интеллектуальный алгоритм обнаружения признаков разлива нефти на поверхности моря в виде двунаправленной семантической сети Ghost-DABiSeNetV2 со слоями «призрачной» свертки [14], на вход которой подаются дневные снимки видимого диапазона и ночные снимки инфракрасного диапазона. Спектральные свойства разливов нефти при наличии солнечных бликов изучались в работе [15]. Результаты исследований оптических свойств (отражательной способности и поглощения) эмульсий нефти в воде и воды в нефти с различными концентрациями под воздействием выветривания с целью интерпретации аэрофотоснимков представлены в работах [16, 17].

Отметим, что в последние годы обнаружение и мониторинг разливов нефти в арктических и около арктических широтах привлекает пристальное внимание исследователей [18, 19].

3. Постановка задачи. Из обзора существующих методов следует, что для достоверного распознавания разливов нефти SAR-снимков не достаточно. Обработку таких снимков можно рассматривать как первый этап решения задачи распознавания, во

время которого локализуются все регионы, похожие на разлив нефти и нефтепродуктов. Тем самым минимизируется ошибка «пропуска цели». На втором этапе распознавания требуется подтвердить, что регионы, похожие на разлив нефти и нефтепродуктов, действительно являются таковыми, иными словами, минимизируется ошибка «ложной тревоги». Дополнительно для уточнения геометрических характеристик регионов интереса выполняется семантическая сегментация, во время которого вычисляются глобальные признаки и все полученные данные записываются в базу данных мониторинга проблемной территории. Таким образом, верификация регионов изображения, похожих на разлив нефти (второй этап), является важной задачей, решение которой необходимо для принятия правильного решения экологического характера.

Сформулируем проблему распознавания как задачу бинарной классификации, которая подразумевает, что имеется класс нефтяных пятен и класс «двойников», куда попадают как органические субстанции, так и артефакты съемки. Бинарная классификация является простейшим случаем классификации. Пусть $\mathbf{X}$ – множество всех возможных объектов в пространстве объектов, в данном случае нефтяных пятен и «двойников» различного происхождения. Пусть $\mathbf{L}$ – пространство меток, а $\mathbf{Y}$ – пространство выходов. Для решения задачи требуется построить модель, которая отображает пространство объектов в пространство выходов. При классификации пространством выходов является множество классов, в данном случае состоящее из двух элементов. Для обучения модели необходим обучающий набор помеченных объектов $(x, l(x))$, называемых примерами, где $l{:}\mathbf{X} \to \mathbf{L}$ – помечающая функция. При этом пространство меток совпадает с пространством выходов $\mathbf{L} = \mathbf{Y}$. Модель подвержена шуму, который может быть меточным (наблюдение искаженной метки l') или объектным (наблюдение искаженного объекта x'). Для предотвращения переобучения на шуме создается тестовый набор из помеченных данных. Отображение $\hat{c}(x) : \mathbf{X} \to [0,1]$ называется бинарным классификатором, где $\hat{c}(x)$ – оценка истиной, но неизвестной функции $c(x)$. Обучающие примеры имеют вид $(x, c(x))$, где $x \in \mathrm{X}$ – объект, а $c(x)$ – истинный класс этого объекта. Под обучением классификатора понимается построение функции $\hat{c}(x)$, которая как можно лучше аппроксимирует истинную функцию $c(x)$ не только на обучающем наборе данных, но и на всем пространстве объектов $\mathbf{X}$.

Семантическая сегментация верифицированных объектов позволяет получить точный вывод на основе плотных прогнозов, иными словами, для каждого пиксела определяется метка класса. Поскольку в данном случае имеются только два класса, то на выходе сети типа кодировщик-декодировщик будет построена бинарная маска, по которой легко вычислить геометрические признаки объекта интереса.

4. Верификация фрагментов изображений разливов нефти оптическим снимкам. В условиях быстрого развития моделей и методов глубокого обучения технологии анализа оптических снимков, включая снимки ДЗЗ, позволяют получать результаты, иногда лучшие, чем экспертные оценки. Можно определить четыре категории признаков разливов нефти на водных поверхностях:

– геометрические признаки, характеризующие форму нефтяных пятен (площадь и периметр пятен);

– физические признаки, определяющие обратное рассеяние волн в физическом пространстве (среднее значение и стандартное отклонение коэффициента контрастности пятен);

– топологические признаки в ближайшем окружении (количество темных пятен вблизи основного нефтяного пятна);

– текстурные признаки, характеризующие поверхность нефтяного пятна (на основе текстурных признаков оценивается непрерывность пятна).

Считается, что геометрические признаки используются наиболее часто из-за относительной простоты их извлечения на SAR снимках. Так, в работе [20] вычислялись 9 видов геометрических признаков, включая собственные значения моментов Ху и эллиптических дескрипторов Фурье, для классификации разлива нефти многослойной сетью прямого распространения.

В данном исследовании использовались, в основном, текстурные и цветовые признаки для верификации разливов нефти. Задачу можно свести к задаче бинарной классификации, когда к одному классу относятся изображения нефтяных пятен, а к другому классу – визуальные артефакты «двойников» биогенного и природного происхождения. Оригинальным подходом является использование особой модели СНС – сиамской сети. Изначально сиамская сеть была предложена в работе [21] в 1993 г. для верификации подписи. Позднее принцип сиамской сети применялся для решения различных задач, включая классификацию изображений, повторную идентификацию человека, отслеживание объектов и др. Сиамская состязательная сеть

(Siamese Adversarial Network) является одной из последних модификаций данного семейства глубоких сетей.

Сиамская сеть состоит из двух идентичных подсетей с одинаковыми наборами весов. Она позволяет сравнивать векторы признаков двух объектов, чтобы выделить в них сходство или различие. Сиамская сеть может определить, принадлежат ли два сравниваемых изображения одному и тому же классу, даже если эти изображения ни разу не использовались во время обучения. Для классификации текстур рекомендуется выбирать такие модели, как AlexNet, VGG19, ResNet50, ResNet101, InceptionResnetV2, DenseNet201, InceptionV3 [22] или семантическую сеть семейства DeepLab [23]. Каждое семейство глубоких моделей имеют свои преимущества и недостатки. Проведенные эксперименты показали, что для данной задачи достаточно использовать относительно простые модели AlexNet или VGGNet. Таким образом, для реализации сиамской сети была выбрана облегченная сеть VGG16 [24].

Обучение сиамской сети осуществлялось методом самообучения, в частности, с использованием функции контрастной потери (contrastive loss function), которая оценивает, насколько хорошо сиамская сеть различает пары изображений. Следует отметить, что недавно разработанные методы самообучения основаны на сочетании двух понятий: контрастные потери и множество преобразований изображений. Контрастные потери не используют понятие экземпляров классов, а сравнивают признаки изображений непосредственно. В то же время преобразования изображений предполагают поиск инвариантных признаков [25].

Идеальная разделимость достигается, когда расстояние между классами в пространстве признаков становится большим, однако хорошей сходимости сети достичь трудно. Обычно стараются увеличить зазор между классами таким образом, чтобы расстояние положительных примеров было больше некоторого значения m, а расстояние отрицательных примеров – меньше некоторого значения n. После обучения примеры разделены зазором $m–n$. Иными словами, векторы признаков вносят вклад в функцию потерь только в том случае, если их параметризованное расстояние находится в пределах этого зазора. Тогда для положительных *Pos* и отрицательных *Neg* примеров можно записать следующие выражения:

$$Pos = \max\left(m - D_{s_1,s_2}, 0\right), \tag{1}$$

$$Neg = \max\left(D_{s_1,s_2} - n, 0\right), \tag{2}$$

где D_{s_1,s_2} – метрика расстояния между классами s_1 и s_2. Для данной задачи выбрана Евклидова метрика:

$$D_{s_1,s_2}\left(\mathbf{y}_1, \mathbf{y}_2\right) = \left\|\mathbf{y}_1 - \mathbf{y}_2\right\|_2, \tag{3}$$

где $\mathbf{y}_1$ и $\mathbf{y}_2$ – выходные векторы признаков двух подсетей.

Из-за несбалансированности классов в обобщенной функции контрастной потери вводятся коэффициенты α и β, подбираемые эмпирически. Примем, что l – значение метки выходного класса, $l = [0, 1]$. Тогда функция контрастной потери $Loss_{CL}$ примет вид выражения (4):

$$\begin{aligned} Loss_{CL}\left(\mathbf{y}_1, \mathbf{y}_2, l\right) &= \alpha\left(1-l\right)\max\left(m - D_{s_1,s_2}\left(\mathbf{y}_1, \mathbf{y}_2\right), 0\right) + \\ &+ \beta l \max\left(D_{s_1,s_2}\left(\mathbf{y}_1, \mathbf{y}_2\right) - n, 0\right). \end{aligned} \tag{4}$$

На вход одной подсети поступает фрагмент изображения для классификации (патч, patch), в то время как на вход другой подсети подаются типовые фрагменты изображений нефтяных пятен. Обе подсети имеются одинаковую топологию СНС, включающую слои свертки, подвыборки (max-pooling) и полносвязные слои, формирующие векторы признаков. Значение контрастной потери рассчитывается как Евклидово расстояние между векторами признаков. Затем цикл повторяется, во время которого входной классифицируемый фрагмент сравнивается с другим типовым фрагментом. Иными словами, реализуется простейший ансамблевый метод бинарной классификации. Значения контрастных потерь суммируются, в результате чего выдается решение о принадлежности входного фрагмента классу нефтяных пятен. В ходе экспериментов было выявлено около 20 типовых фрагментов изображений нефтяных пятен, примеры которых представлены на рисунке 1. Количество типовых фрагментов, используемых для верификации входного фрагмента сиамской сети, выбирается исходя из обучающего набора данных. Если классификационный порог (50%) превышен, то считается, что входное изображение верифицировано как принадлежащее классу нефтяных пятен или не принадлежащее

данному классу. Процесс верификации фрагментов исходного изображения с помощью сиамской сети изображен на рисунке 2.

Рис. 1. Примеры фрагментов изображений нефтяных пятен

Вторая ветвь является обученной на нескольких десятков тысяч фрагментов (патчей) с разрешением 128×128 пикселов, вырезанных из аэрофотоснимков разливов нефти, полученных по всему миру. Для расширения обучающего набора данных применялись методы аугментации, о чем будет дополнительно сказано в разделе 6.2.

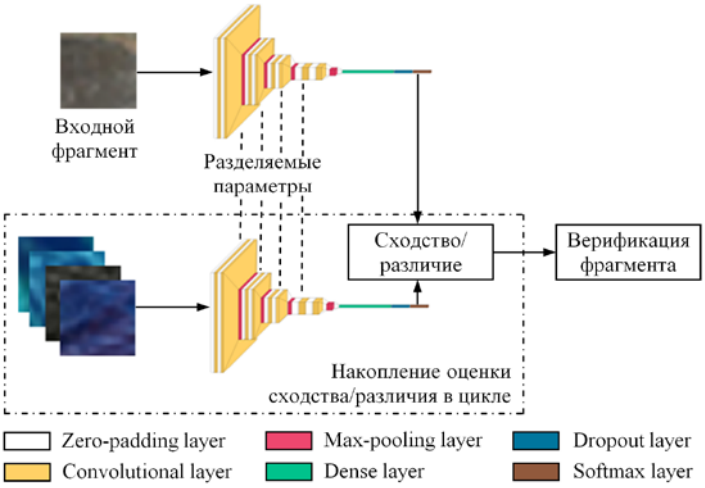


Рис. 2. Верификация фрагментов изображения на основе сиамской сети

5. Сегментация изображений разливов нефти. Помимо верификации фрагментов оптические снимки позволяют с высокой точностью оценить геометрические признаки, характеризующие форму нефтяных пятен (площадь и периметр пятен), как один из способов быстрого и эффективного мониторинга. В настоящее время общепринятыми моделями сегментации являются глубокие сети Faster R-CNN, Darknet-53, DeepLabv3+, LinkNet, SegNet и U-Net [26]. Отметим, что даже сеть U-Net, изначально спроектированная для

сегментации объектов на медицинских снимках, в настоящее время используется для сегментации разливов нефти на SAR-снимках.

Семантическая сегментация с использованием полносвязных СНС выполняется за счет того, что полносвязные выходные слои заменяются сверточными слоями с повышающей дискретизацией для предсказания выходных пикселов. Помимо этого, в сети используются соединения с пропусками (skip connections) для улучшения предсказания [27]. В свою очередь, архитектура сети U-net состоит из кодировщика и декодировщика с последовательными сверточными слоями понижающей и повышающей дискретизации соответственно [28]. При этом слои кодировщика соединены со слоями декодировщика. Такая архитектура позволяет точнее сегментировать изображения. Модель SegNet имеет аналогичную архитектуру кодировщика-декодировщика, но использует билинейную интерполяцию на уровне декодировщика, что позволяет уменьшить количество обучаемых параметров и тем самым ускорить обучение [29]. Основу кодировщика сети SegNet составляют сверточные слои сети VGG16. Декодировщик состоит из транспонированных сверточных слоев и слоев повышающей дискретизации, а его структура симметрична структуре кодировщика. Недавние исследования привели к созданию гибридной модели U-SegNet, использующей мультимасштабные признаки и сверточные слои, основанные на внимании [30].

В работе [31] была предложена модификация одноэтапного многофункционального детектора YOLO, названная как Poly-YOLOv3, которая способна выполнять сегментацию визуальных объектов. Однако следует отметить, что сегментация, выполняемая сетью Poly-YOLOv3, основана на переводе изображения объекта из декартовой системы в полярную систему координат с началом в центральной точке обрамляющего прямоугольника. Иными словами, сеть Poly-YOLOv3 предназначена для сегментации объектов выпуклых форм. В редких случаях при сильных вихревых течениях форма нефтяного пятна перестает быть выпуклой, что приводит к неточным результатам сегментации. Неоспоримым преимуществом сетей семейства YOLO является их высокое быстродействие.

Для сегментации изображений верифицированных нефтяных пятен использовались три архитектуры глубоких сетей, а именно U-Net, SegNet и Poly-YOLOv3. На рисунке 3 приведены примеры бинарных масок, построенных различными методами.

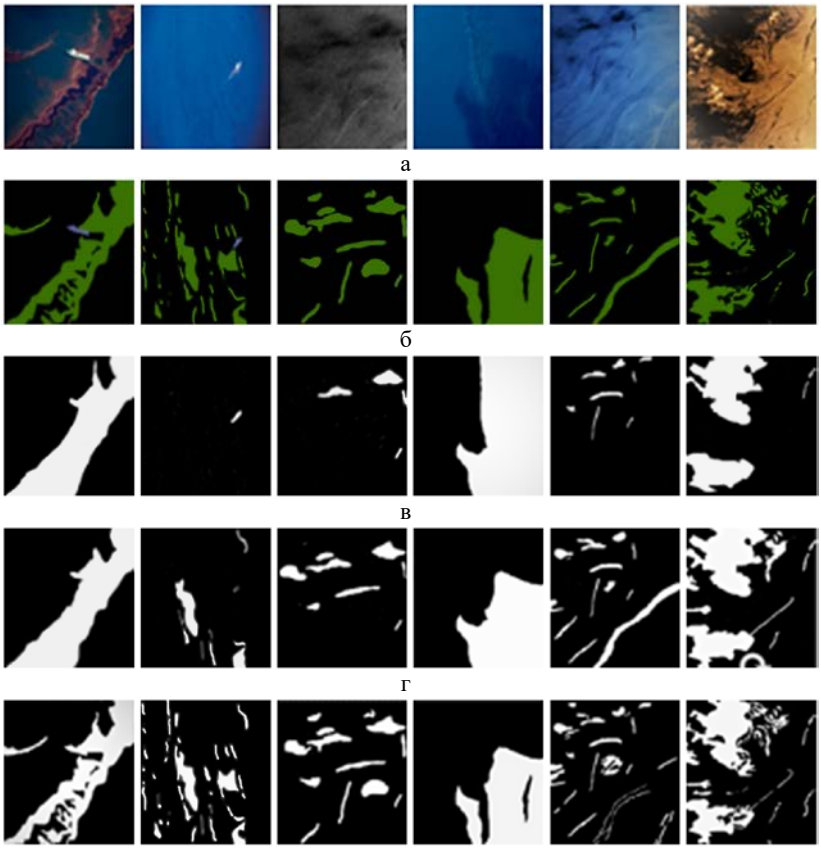


Рис. 3. Примеры сегментации изображений нефтяных пятен: а) исходные снимки, б) бинарные маски, построенные экспертом, в) бинарные маски, полученные сетью U-Net, г) бинарные маски, полученные сетью SegNet, д) бинарные маски, полученные сетью Poly-YOLOv3

Как видно из рисунка 3, бинарные маски, построенные сетями U-Net и SegNet, почти идентичны. Сеть Poly-YOLOv3 дает визуально лучшие результаты сегментации участков разливов нефти. Обобщенные оценки точности сегментации трем сетями приведены в разделе 6.4.

6. Экспериментальные исследования. В данном разделе представлено описание набора данных, собранного из нескольких открытых источников. Для увеличения собранного набора данных и

устранения несбалансированности классов применены методы аугментации. Приведены оценочные метрики, а также экспериментальные настройки и полученные результаты.

6.1. Набор данных. В отличие от доступных наборов космических SAR снимков с разливами нефти, аэрофотоснимки оптических и/или мультиспектральных изображений отсутствуют в открытом доступе. Был собран собственный набор данных из открытых источников, в частности, с сайтов Alamy [32], Getty Images [33] и др.

Собранный набор данных содержит 380 изображений морской поверхности с разливами нефти, 135 изображений «чистой» морской поверхности и 35 изображений «двойников», большинство из которых получено при съемке с беспилотных летательных аппаратов и средств малой авиации по всему миру. Съемка осуществлялась при различных высотах полета и погодных условиях, в результате чего снимки можно классифицировать по следующим показателям:

– виду местности (прибрежные районы, пляж, морские акватории);

– наличию дополнительных объектов (корабли, дым, лес, горы, люди);

– погодным условиям (облачно/солнечно);

– условиям освещенности (дневные/вечерние);

– цветам разливов нефти (красный, оранжевый, желтый, зеленый).

Разрешение изображений варьируется от нескольких сотен до тысяч пикселов. На рисунке 4 приведены примеры изображений собранного набора данных.

Изображения с разрешением 1250×650 были разбиты на фрагменты размером 128×128 путем кадрирования и сканирования изображений с использованием ядра размерностью 128×128 и шагом 1. Далее сгенерированные патчи проверялись на наличие участков с разливом нефти и участков «чистой» морской поверхности и «двойников», тем самым выполнялась ручная разметка фрагментов. В общей сложности было отобрано 199990 фрагментов, из которых около 70% относится к разливам нефти, а остальные – к «чистой» морской поверхности и «двойникам». В силу того, что изображений «двойников» слишком мало для решения задачи верификации, потребовалось выполнить аугментацию не только простыми методами, но и с использованием генеративных состязательных сетей.

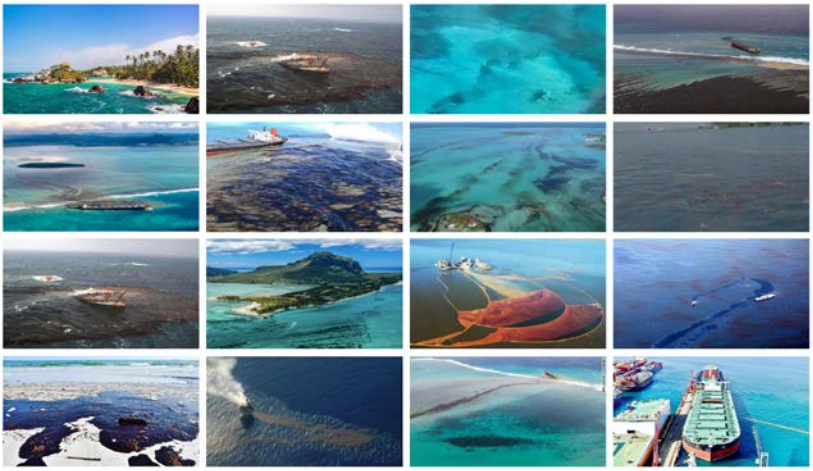


Рис. 4. Примеры изображений собранного набора данных

6.2. Аугментация. Аугментация как процесс контролируемого увеличения примеров для обучения глубоких сетей является широко распространенной практикой для частных задач, для которых отсутствуют свободно распространяемые наборы данных, количество примеров невелико или набор данных является несбалансированным. Задача верификации разливов нефти относится именно к таким частным задачам в силу ряда объективных обстоятельств экологического и материального характера.

Известны разные подходы к классификации методов аугментации. Наиболее полный обзор с целью выбора приемлемых методов аугментации для конкретного приложения можно найти в работе [34]. Методы аугментации изображений разделяются на три основных категории: методы, основанные на преобразованиях (их иногда называют простыми методами аугментации), модельные методы, создающие изображения с помощью обученной модели, и оптимальные методы на основе политик, предполагающие нахождение оптимальных стратегий аугментации.

В данной работе были использованы методы из первых двух категорий, предназначенные для обработки одиночных снимков (технологии перемешивания изображений не использовались):

– методы первой категории, основанные на простых преобразованиях, таких как геометрические преобразования, преобразования цветового пространства, ядерные фильтры и случайная обрезка;

– методы второй категории, основанные на состязательном обучении, а именно метод аугментации с использованием генеративно-состязательных сетей. Такие методы характеризуются тем, что теоретически распределение сгенерированных изображений аналогично распределению исходного набора данных, однако сгенерированные изображения не совпадают с исходными изображениями.

Приведем параметры аугментации методов, основанных на преобразованиях, которые были использованы для расширения набора данных и компенсации несбалансированности двух классов:

– сдвиги по осям OX и OY не более 10% от разрешения исходного снимка;
– повороты в диапазоне [–45°…45°];
– отражение по горизонтали или вертикали;
– растяжение/сжатие в диапазоне [0,75…1,25];
– масштабирование до 25%;
– равномерное наложение белого шума со случайным изменением яркости в цветовом канале в диапазоне [–30…30];
– изменение яркости в каждом цветовом канале не более 15%;
– размытие с помощью гауссова фильтра;
– изменение резкости с помощью фильтра краев.

Следует отметить, что исходные изображения имели разрешение, достаточное для того, чтобы разрешение аугментированных изображений соответствовало заранее установленному разрешению входных изображений, подаваемых на сиамскую сеть.

Дополнительно, был применен метод аугментации с использованием сети GAN. Изначально архитектура GAN основывалась на том, что генератор и дискриминатор представляют многослойные персептроны. Однако в настоящее время считается более перспективным способом аугментации применение таких GAN сетей, как DCGANs, Progressively Growing GANs, CycleGANs и Conditional GANs. В данной работе была применена глубокая сверточная сеть GAN (Deep Convolutional GANs, DCGAN) [35].

В результате аугментации количество фрагментов «двойников» возросло до 4600. Далее весь сформированный набор данных был разделен на обучающую и тестовую выборки в соотношении 70:30.

6.3. Оценочные метрики. Оценка точности сегментации нефтяных пятен проводилась с использованием пяти показателей, вычисляемых по матрице неточностей [36], а именно:

– верность *AC*, показывает отношение правильного прогноза к общему наблюдению;

– точность *PR*, указывает на соотношение правильно сегментированных пикселов;

– полнота *RE*, показывает, какую долю пикселов положительного класса из всех пикселов положительного класса нашел алгоритм;

– мера F_1, характеризует степень совпадения предсказанной границы с истиной границей классов;

– пересечение по объединению *IoU*, также называемое коэффициентом перекрытия Жаккара (Jaccard), указывает на отношение сходства между предсказанным регионом и истинным регионом класса.

6.4. Экспериментальные настройки и полученные результаты. В экспериментах была использована облегченная сеть VGG-16, являющаяся базовой для сиамской сети [24]. Данная модель концептуально повторяет архитектуру полной модели СНС, однако она учитывает особенности в связи с обработкой снимков ДЗЗ низкого разрешения, которые имеют относительно малое количество точечных особенностей. Архитектура облегченной сети VGG-16, адаптированная под обработку фрагментов изображений, представлена в таблице 1.

Слой Zero-padding (слой нулевого заполнения) повышает нелинейность сети, более точно извлекая характерные особенности снимков ДЗЗ. Размер фильтров в слое Convolutional составляет 3×3. Слой Max-pooling (слой подвыборки) нелинейно уплотняет карту признаков. Слой Dense – это плотно связанный слой. Слой Dropout обеспечивает регуляризацию с целью уменьшения переобучения сети. Слой Softmax применяет обобщенную логистическую функцию к выходу сети.

Для разработки архитектуры сети была использована открытая библиотека Keras, написанная на языке Python. Обучение облегченной сети VGG-16 проводилось со следующими параметрами: размер пакета – 32, значение импульса – 0,9, скорость обучения – 0,001, алгоритм оптимизации –Adam, количество эпох обучения – 100. Среднее время обучения составило 30 часов. На рисунках 5 и 6 приведены графики функций точности и функций потерь в процессе обучения (training) и тестирования (testing) сети VGG-16.

Таблица 1. Архитектура облегченной сети VGG-16

Слой	Входная размерность	Выходная размерность
Input image	128×128×3	
Zero-padding	128×128×3	128×128×64
Convolutional 3-64+ReLU	128×128×3	128×128×64
Zero-padding	128×128×3	128×128×64
Convolutional 3-64+ReLU	128×128×3	128×128×64
Max-pooling	2×2	
Zero-padding	128×128×64	64×64×128
Convolutional 3-128+ReLU	128×128×64	64×64×128
Zero-padding	128×128×64	64×64×128
Convolutional 3-128+ReLU	128×128×64	64×64×128
Max-pooling	2×2	
Zero-padding	64×64×128	32×32×256
Convolutional 3-256+ReLU	64×64×128	32×32×256
Zero-padding	64×64×128	32×32×256
Convolutional 3-256+ReLU	64×64×128	32×32×256
Max-pooling	2×2	
Zero-padding	16×16×512	16×16×512
Convolutional 3-512+ReLU	16×16×512	16×16×512
Zero-padding	16×16×512	16×16×512
Convolutional 3-512+ReLU	16×16×512	16×16×512
Max-pooling	2×2	
Dense	16×16×512	1×1×4096
Dropout	1×1×0,5	
Softmax	1×1×1000	1×1×1000

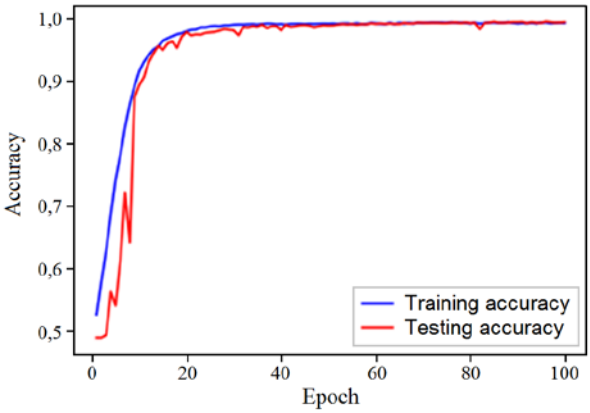


Рис. 5. Графики функций точности при обучении и тестировании сети VGG-16

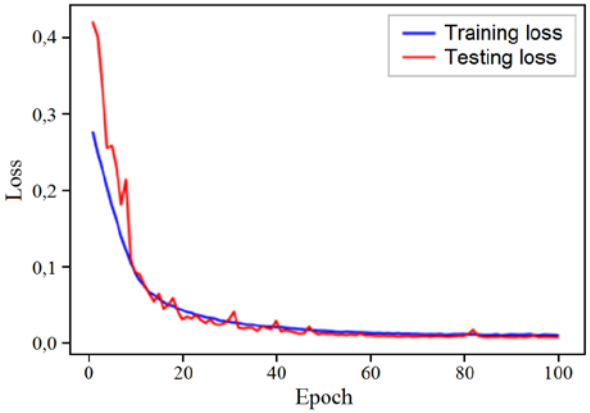


Рис. 6. Графики функций потерь при обучении и тестировании сети VGG-16

В таблице 2 приведены результаты классификации тестовых RGB изображений, относящихся к двум классам – разлив нефти и «двойники», предложенной сиамской нейронной сетью.

Таблица 2. Результаты классификации тестовых RGB изображений

Класс	*AC*	*PR*	*RE*	F_1	*IoU*
Разлив нефти	0,910	0,920	0,938	0,930	0,870
«Двойники»	0,834	0,864	0,855	0,867	0,751

После верификации областей разлива нефти требуется выполнить их сегментацию. Причем, как правило, рядом с основным нефтяным пятном имеются небольшие нефтяные пятна, что можно считать дополнительным признаком наличия разлива нефти. Для семантической сегментации были протестированы три сети U-Net, SegNet и Poly-YOLOv3 с небольшими модификациями слоев. Результаты семантической сегментации тестовых RGB изображений приведены в таблице 3.

Таблица 3. Результаты сегментации тестовых RGB изображений

Сеть	*AC*	*PR*	*RE*	F_1	Среднее время, с
U-Net	0,802	0,8695	0,8450	0,8571	458
SegNet	0,827	0,8529	0,8787	0,8656	625
Poly-YOLOv3	0,924	0,9705	0,9166	0,9428	385

Как видно из таблицы 3, лучшие результаты по точности сегментации и быстродействию показала сеть Poly-YOLOv3, позволяющая не только обнаруживать и распознавать объекты, но и определять их форму. Кроме того, сеть Poly-Yolov3 относится к одноэтапным детекторам, что приводит к сокращению времени вычислений. Тестирование проводилось с помощью веб-сервиса Google Colab.

Дополнительно была разработана база данных под управлением свободно распространяемой СУБД MySQL 5.7, которая хранит информацию об исходных снимках и результатах верификации и сегментации в виде визуальных и числовых данных. В описание исходных снимков входят следующие атрибуты:

- идентификатор снимка;
- наименование снимка;
- географические координаты;
- тип и вид прибора (при наличии);
- разрешение снимка;
- дата съемки;
- дата обработки;
- идентификатор пользователя.

В описание обработанных снимков входят следующие атрибуты:

- идентификатор снимка;
- результат верификации (1/0);
- площадь основного нефтяного пятна;

– количество дополнительных нефтяных пятен;
– средняя площадь дополнительных нефтяных пятен;
– дата обработки;
– идентификатор пользователя.

На рисунке 7 приведен пример экранной формы предварительного просмотра аэрофотоснимков.

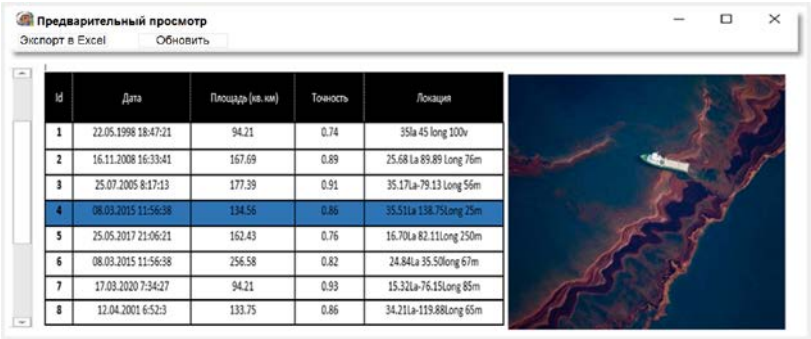


Рис. 7. Пример экранной формы «Предварительный просмотр»

Также имеются транзакционные таблицы, позволяющие хранить и обрабатывать несколько снимков одной и той же территории, но полученных в разное время с целью мониторинга разлива нефти. База данных содержит 14 связанных таблиц, 2 представления. Также реализованы 3 отчета и экспорт данных о снимках в Excel.

7. Заключение. В статье рассматривается задача верификации разливов нефти по оптическим аэрофотоснимкам как задача бинарной классификации. Впервые предложено решение на основе сиамской сети, когда поступающий фрагмент сравнивается с репрезентативными фрагментами нефтяных пятен в пространстве признаков. Для обучения сиамской сети используется контрастная функция потери. Основой сиамской сети служит облегченная сеть VGG-16. Решение о верификации снимка принимается, если количество фрагментов, классифицированных как разлив нефти, превысило установленное значение. Далее выполняется сегментация нефтяных пятен, основного и дополнительных, которые, как правило, возникают при разливах нефти. Для этого используются сети семантической сегментации. Проведено экспериментальное исследование трех сетей, а именно U-Net, SegNet и Poly-YOLOv3. Наилучшую точность и наилучшее быстродействие на собранном наборе данных показала сеть Poly-YOLOv3. Дополнительно разработано приложение для построения

бинарных масок и вычисления относительной площади разлива нефти. Как исходные снимки с соответствующими атрибутами, так и результаты сегментации записываются в разработанную базу данных, работающую под управлением СУБД MySQL 5.7 с использованием dbForge Studio 2020 for MySQL. Такое решение является удобным для конечного пользователя, который формирует запросы и отчеты для мониторинга экологической ситуации.

Литература

1. Ivshina I.B., Kuyukina M.S., Krivoruchko A.V., Elkin A.A., Makarov S.O., Cunningham C.J., Peshkur T.A., Atlas R.M., Philp J.C. Oil spill problems and sustainable response strategies through new technologies // Environmental Science: Processes & Impacts journal. 2015. vol. 17. no. 7. pp. 1201-1219.
2. Hackett B., Comerma E., Daniel P., Ichikawa H. Marine oil pollution prediction // Oceanography. 2009. vol. 22. pp. 168-175.
3. Wang R.; Zhu Z.; Zhu W.; Fu X., Xing S. A dynamic marine oil spill prediction model based on deep learning // Journal of Coastal Research. 2021. vol. 37. no. 4. pp. 716-725.
4. Lubchenco J., McNutt M.K., Dreyfus G., Murawski S.A., Kennedy D.M., Anastas P.T., Chu S., Hunter T. Science in support of the deepwater horizon response // PNAS. 2012. vol. 109. no. 50. pp. 20212-20221.
5. Mera D., Bolon-Canedo V., Cotos J.M., Alonso-Betanzos A. On the use of feature selection to improve the detection of sea oil spills in SAR images // Computers & Geosciences. 2017. vol. 100. pp. 166-178.
6. Zeng K., Wang Y. A deep convolutional neural network for oil spill detection from spaceborne SAR images // Remote Sensing. 2020. vol. 12. no. 6. pp. 1015.1-1015.23.
7. Chen G., Li Y., Sun G., Zhang Y. Application of deep networks to oil spill detection using polarimetric synthetic aperture radar images // Applied Sciences. 2017. vol. 7. no. 10. pp. 968.1-968.15.
8. Bianchi F.M., Espeseth M.M., Borch N. Large-scale detection and categorization of oil spills from SAR images with deep learning // Remote Sensing. 2020. vol. 12. pp. 2260.1-2260.27.
9. Angelliaume S., Ceamanos X., Viallefont-Robinet F., Baque R., Deliot P., Miegebielle V. Hyperspectral and radar airborne imagery over controlled release of oil at sea // Sensors. 2017. vol. 17. no. 8. pp. 1772.1-1772.21.
10. Huang H., Wang C., Liu S., Sun Z., Zhang D., Liu C., Jiang Y., Zhan S., Zhang H., Xu R. Single spectral imagery and Faster R-CNN to identify hazardous and noxious substances spills // Environmental Pollution. 2020. vol. 258. pp. 113688.1-113688.11.
11. Wang D., Wan J., Liu S., Chen Y., Yasir M., Xu M., Ren P. BO-DRNet: An improved deep learning model for oil spill detection by polarimetric features from SAR images // Remote Sensing. 2022. vol. 14. pp. 264.1-264.18.
12. Aghaei N., Akbarizadeh G., Kosarian A. GreyWolfLSM: An accurate oil spill detection method based on level set method from synthetic aperture radar imagery // European Journal of Remote Sensing. 2022. vol. 55. no. 1. pp. 181-198.
13. Chen Y., Sun Y., Yu W., Liu Y, Hu H. A novel lightweight bilateral segmentation network for detecting oil spills on the sea surface // Marine Pollution Bulletin. 2022. vol. 175. pp. 113343.1-113343.12.
14. Paoletti M.E., Haut J.M., Pereira N.S. Ghostnet for hyperspectral image classification // IEEE Transactions on Geoscience and Remote Sensing. 2021. vol. 59. no. 12. pp. 10378-10393.

15. Bulgarelli B., Djavidnia S. On MODIS Retrieval of oil spill spectral properties in the marine environment // IEEE Geoscience and Remote Sensing Letters. 2012. vol. 9. no. 3. pp. 398-402.
16. Lu Y., Shi J., Wen Y., Hu C., Zhou Y., Sun S., Zhang M., Mao Z., Liu Y. Optical interpretation of oil emulsions in the ocean – Part I: Laboratory measurements and proof-of-concept with AVIRIS observations // Remote Sensing of Environment. 2019. vol. 230. pp. 111183.1-111183.14.
17. Lu Y., Shi J., Hu C., Zhang M., Sun S., Liu Y. Optical interpretation of oil emulsions in the ocean – Part II: Applications to multi-band coarse-resolution imagery // Remote Sensing of Environment. 2020. vol. 242. pp. 111778.1-111778.14.
18. Yang Z., Chen Z., Lee K., Owens E., Boufadel M.C., C., Taylor E. Decision support tools for oil spill response (OSR-DSTs): Approaches, challenges, and future research perspectives // Marine Pollution Bulletin. 2021. vol. 167. pp. 112313.1-112313.16.
19. Mohammadiun S., Hu G., Gharahbagh A.A., Jianbing Li c, Hewage K., Sadiq R. Evaluation of machine learning techniques to select marine oil spill response methods under small-sized dataset conditions // Journal of Hazardous Materials. 2022. vol. 436. pp. 129282.1-129282.11.
20. Guo Y., Zhang H.Z. Oil spill detection using synthetic aperture radar images and feature selection in shape space // International Journal of Applied Earth Observation and Geoinformation. 2014. vol. 30, pp. 146-157.
21. Bromley J., Bentz J.W., Bottou L., Guyon I., Lecun Y., Moore C., Säckinger E., Shah R. Signature verification using a "siamese" time delay neural network // International Journal of Pattern Recognition and Artificial Intelligence. 1993. vol. 7. no. 4. pp. 669-688.
22. Yelchuri R., Dash J.K., Singh P., Mahapatro A., Sibarama S. Exploiting deep and hand-crafted features for texture image retrieval using class membership // Pattern Recognition Letters. 2022. vol. 160. pp. 163-171.
23. Chen L.C., Papandreou G., Kokkinos I., Murphy K., Yuille A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs // IEEE Transactions on Pattern Analysis & Machine Intelligence. 2018. vol. 40. no. 4. pp. 834-848.
24. Ye M., Ruiwen N., Chang Z., He G., Tianli H., Shijun L., Yu S., Tong Z., Ying G. A lightweight model of VGG-16 for remote sensing image classification // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2021. vol. 14. pp. 6916-6922.
25. Caron M., Misra I., Mairal J., Goyal P., Bojanowski P., Joulin A. Unsupervised learning of visual features by contrasting cluster assignments // Proceeding of the 34th International Conference on Neural Information Processing Systems (NIPS'20). 2020. Article no. 831. pp. 9912-9924.
26. de Moura N.V.A., de Carvalho O.L.F., Gomes R.A.T., Guimaraes R.F., de Carvalho Júnior O.A. Deep-water oil-spill monitoring and recurrence analysis in the brazilian territory using Sentinel-1 time series and deep learning // International Journal of Applied Earth Observations and Geoinformation. 2022. vol. 107. pp. 102695.1-102695.11.
27. Long J., Shelhamer E., Darrell T. Fully convolutional networks for semantic segmentation // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. pp. 3431-3440.
28. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. 2015. pp. 234-241.

29. Badrinarayanan V.; Kendall A.; Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. vol. 39. pp. 2481-2495.
30. Dayananda C., Choi J.-Y., Lee B. Multi-scale squeeze U-SegNet with multi global attention for brain MRI segmentation // Sensors. 2021. vol. 21. pp. 3363.1-3363.22.
31. Hurtik P., Molek V., Hula J., Vajgl M., Vlasanek P., Nejezchleba T. Poly-YOLO: Higher speed, more precise detection and instance segmentation for YOLOv3 // CoRR arXiv preprint, arXiv:2005.13243v2. 2020. pp. 1-18.
32. Alamy [Official web site Alamy Stock photography]. Available at: www.almy.com. (accessed 26.07.2022).
33. Getty Images [Official web site of Getty Images]. Available at: www.gettyimages.nl. (accessed 26.07.2022).
34. Xu M., Yoon S., Fuentes A., Park D.S. A comprehensive survey of image augmentation techniques for deep learning // CoRR arXiv preprint, arXiv:2205.01491v1. 2022. pp. 1-41.
35. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks // The 4th International Conference on Learning Representations (ICLR 2016). 2016. pp. 1–16.
36. Goodfellow I., Bengio Y., Courville A. Deep learning / Dietterich T. (ed.) // Cambridge, Massachusetts, London: The MIT Press. 2016. 800 p.

Фаворская Маргарита Николаевна — д-р техн. наук, профессор, заведующий кафедрой, кафедра информатики и вычислительной техники, Сибирский государственный университет науки и технологий имени М.Ф. Решетнева. Область научных интересов: компьютерное зрение, обработка изображений и видеопоследовательностей, глубокое обучение, распознавание образов. Число научных публикаций — 300. favorskaya@gmail.com; проспект им. газеты Красноярский рабочий, 31, 660037, Красноярск, Россия; р.т.: +7(3912)139-622.

Нишчхал Нишчхал — магистрант, кафедра информатики и вычислительной техники, Сибирский государственный университет науки и технологий имени акад. М.Ф. Решетнева. Область научных интересов: интеллектуальные системы, глубокое обучение. Число научных публикаций — 18. nik.321g@yandex.ru; проспект им. газеты "Красноярский рабочий", 31, 660037, Красноярск, Россия; р.т.: +7(3912)139-623.

DOI 10.15622/ia.21.5.4

M. FAVORSKAYA, N. NISHCHHAL

VERIFICATION OF MARINE OIL SPILLS USING AERIAL IMAGES BASED ON DEEP LEARNING METHODS

Favorskaya M., Nishchhal N. **Verification of Marine Oil Spills Using Aerial Images Based on Deep Learning Methods.**

Abstract. The article solves the problem of verifying oil spills on the water surfaces of rivers, seas and oceans using optical aerial photographs, which are obtained from cameras of unmanned aerial vehicles, based on deep learning methods. The specificity of this problem is the presence of areas visually similar to oil spills on water surfaces caused by blooms of specific algae, substances that do not cause environmental damage (for example, palm oil), or glare when shooting (so-called look-alikes). Many studies in this area are based on the analysis of synthetic aperture radars (SAR) images, which do not provide accurate classification and segmentation. Follow-up verification contributes to reducing environmental and property damage, and oil spill size monitoring is used to make further response decisions. A new approach to the verification of optical images as a binary classification problem based on the Siamese network is proposed, when a fragment of the original image is repeatedly compared with representative examples from the class of marine oil slicks. The Siamese network is based on the lightweight VGG16 network. When the threshold value of the output function is exceeded, a decision is made about the presence of an oil spill. To train the networks, we collected and labeled our own dataset from open Internet resources. A significant problem is an imbalance of classes in the dataset, which required the use of augmentation methods based not only on geometric and color manipulations, but also on the application of a Generative Adversarial Network (GAN). Experiments have shown that the classification accuracy of oil spills and look-alikes on the test set reaches values of 0.91 and 0.834, respectively. Further, an additional problem of accurate semantic segmentation of an oil spill is solved using convolutional neural networks (CNN) of the encoder-decoder type. Three deep network architectures U-Net, SegNet, and Poly-YOLOv3 have been explored for segmentation. The Poly-YOLOv3 network demonstrated the best results, reaching an accuracy of 0.97 and an average image processing time of 385 s with the Google Colab web service. A database was also designed to store both original and verified images with problem areas.

Keywords: oil spill detection, verification, segmentation, deep learning, aerial photographs, Earth remote sensing.

References

1. Ivshina I.B., Kuyukina M.S., Krivoruchko A.V., Elkin A.A., Makarov S.O., Cunningham C.J., Peshkur T.A., Atlas R.M., Philp J.C. Oil spill problems and sustainable response strategies through new technologies. Environmental Science: Processes & Impacts journal. 2015. vol. 17. no. 7. pp. 1201-1219.
2. Hackett B., Comerma E., Daniel P., Ichikawa H. Marine oil pollution prediction. Oceanography. 2009. vol. 22. pp. 168-175.
3. Wang R.; Zhu Z.; Zhu W.; Fu X., Xing S. A Dynamic marine oil spill prediction model based on deep learning. Journal of Coastal Research. 2021. vol. 37. no. 4. pp. 716-725.
4. Lubchenco J., McNutt M.K., Dreyfus G., Murawski S.A., Kennedy D.M., Anastas P.T., Chu S., Hunter T. Science in support of the deepwater horizon response. PNAS. 2012. vol. 109. no. 50. pp. 20212-20221.

5. Mera D., Bolon-Canedo V., Cotos J.M., Alonso-Betanzos A. On the use of feature selection to improve the detection of sea oil spills in SAR images. Computers & Geosciences. 2017. vol. 100. pp. 166-178.
6. Zeng K., Wang Y. A deep convolutional neural network for oil spill detection from spaceborne SAR images. Remote Sensing. 2020. vol. 12. no. 6. pp. 1015.1-1015.23.
7. Chen G., Li Y., Sun G., Zhang Y. Application of deep networks to oil spill detection using polarimetric synthetic aperture radar images. Applied Sciences. 2017. vol. 7. no. 10. pp. 968.1-968.15.
8. Bianchi F.M., Espeseth M.M., Borch N. Large-scale detection and categorization of oil spills from SAR images with deep learning. Remote Sensing. 2020. vol. 12. pp. 2260.1-2260.27.
9. Angelliaume S., Ceamanos X., Viallefont-Robinet F., Baque R., Deliot P., Miegebielle V. Hyperspectral and radar airborne imagery over controlled release of oil at sea. Sensors. 2017. vol. 17. no. 8. pp. 1772.1-1772.21.
10. Huang H., Wang C., Liu S., Sun Z., Zhang D., Liu C., Jiang Y., Zhan S., Zhang H., Xu R. Single spectral imagery and Faster R-CNN to identify hazardous and noxious substances spills. Environmental Pollution. 2020. vol. 258. pp. 113688.1-113688.11.
11. Wang D., Wan J., Liu S., Chen Y., Yasir M., Xu M., Ren P. BO-DRNet: An improved deep learning model for oil spill detection by polarimetric features from SAR images. Remote Sensing. 2022. vol. 14. pp. 264.1-264.18.
12. Aghaei N., Akbarizadeh G., Kosarian A. GreyWolfLSM: An accurate oil spill detection method based on level set method from synthetic aperture radar imagery. European Journal of Remote Sensing. 2022. vol. 55. no. 1. pp. 181-198.
13. Chen Y., Sun Y., Yu W., Liu Y, Hu H. A novel lightweight bilateral segmentation network for detecting oil spills on the sea surface. Marine Pollution Bulletin. 2022. vol. 175. pp. 113343.1-113343.12.
14. Paoletti M.E., Haut J.M., Pereira N.S. Ghostnet for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing. 2021. vol. 59. no. 12. pp. 10378-10393.
15. Bulgarelli B., Djavidnia S. On MODIS Retrieval of oil spill spectral properties in the marine environment. IEEE Geoscience and Remote Sensing Letters. 2012. vol. 9. no. 3. pp. 398-402.
16. Lu Y., Shi J., Wen Y., Hu C., Zhou Y., Sun S., Zhang M., Mao Z., Liu Y. Optical interpretation of oil emulsions in the ocean – Part I: Laboratory measurements and proof-of-concept with AVIRIS observations. Remote Sensing of Environment. 2019. vol. 230. pp. 111183.1-111183.14.
17. Lu Y., Shi J., Hu C., Zhang M., Sun S., Liu Y. Optical interpretation of oil emulsions in the ocean – Part II: Applications to multi-band coarse-resolution imagery. Remote Sensing of Environment. 2020. vol. 242. pp. 111778.1-111778.14.
18. Yang Z., Chen Z., Lee K., Owens E., Boufadel M.C., C., Taylor E. Decision support tools for oil spill response (OSR-DSTs): Approaches, challenges, and future research perspectives. Marine Pollution Bulletin. 2021. vol. 167. pp. 112313.1-112313.16.
19. Mohammadiun S., Hu G., Gharahbagh A.A., Jianbing Li c, Hewage K., Sadiq R. Evaluation of machine learning techniques to select marine oil spill response methods under small-sized dataset conditions. Journal of Hazardous Materials. 2022. vol. 436. pp. 129282.1-129282.11.
20. Guo Y., Zhang H.Z. Oil spill detection using synthetic aperture radar images and feature selection in shape space. International Journal of Applied Earth Observation and Geoinformation. 2014. vol. 30, pp. 146-157.
21. Bromley J., Bentz J.W., Bottou L., Guyon I., Lecun Y., Moore C., Säckinger E., Shah R.Signature verification using a "siamese" time delay neural network. International

Journal of Pattern Recognition and Artificial Intelligence. 1993. vol. 7. no. 4. pp. 669-688.
22. Yelchuri R., Dash J.K., Singh P., Mahapatro A., Sibarama S. Exploiting deep and hand-crafted features for texture image retrieval using class membership. Pattern Recognition Letters. 2022. vol. 160. pp. 163-171.
23. Chen L.C., Papandreou G., Kokkinos I., Murphy K., Yuille A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis & Machine Intelligence. 2018. vol. 40. no. 4. pp. 834-848.
24. Ye M., Ruiwen N., Chang Z., He G., Tianli H., Shijun L., Yu S., Tong Z., Ying G. A lightweight model of VGG-16 for remote sensing image classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2021. vol. 14. pp. 6916-6922.
25. Caron M., Misra I., Mairal J., Goyal P., Bojanowski P., Joulin A. Unsupervised learning of visual features by contrasting cluster assignments. Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS'20). 2020. Article no. 831. pp. 9912-9924.
26. de Moura N.V.A., de Carvalho O.L.F., Gomes R.A.T., Guimaraes R.F., de Carvalho Júnior O.A. Deep-water oil-spill monitoring and recurrence analysis in the brazilian territory using Sentinel-1 time series and deep learning. International Journal of Applied Earth Observations and Geoinformation. 2022. vol. 107. pp. 102695.1-102695.11.
27. Long J., Shelhamer E., Darrell T. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. pp. 3431-3440.
28. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. 2015. pp. 234-241.
29. Badrinarayanan V.; Kendall A.; Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. vol. 39. pp. 2481-2495.
30. Dayananda C., Choi J.-Y., Lee B. Multi-scale squeeze U-SegNet with multi global attention for brain MRI segmentation. Sensors. 2021. vol. 21. pp. 3363.1-3363.22.
31. Hurtik P., Molek V., Hula J., Vajgl M., Vlasanek P., Nejezchleba T. Poly-YOLO: Higher speed, more precise detection and instance segmentation for YOLOv3. CoRR arXiv preprint, arXiv:2005.13243v2. 2020. pp. 1-18.
32. Alamy [Official web site Alamy Stock photography]. Available at: www.almy.com. (accessed 26.07.2022).
33. Getty Images [Official web site of Getty Images]. Available at: www.gettyimages.nl. (accessed 26.07.2022).
34. Xu M., Yoon S., Fuentes A., Park D.S. A comprehensive survey of image augmentation techniques for deep learning. CoRR arXiv preprint, arXiv:2205.01491v1. 2022. pp. 1-41.
35. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. The 4th International Conference on Learning Representations (ICLR 2016). 2016. pp. 1–16.
36. Goodfellow I., Bengio Y., Courville A. Deep learning. Dietterich T. (ed.). The MIT Press. 2016. 800 p.

Favorskaya Margarita — Ph.D., Dr.Sci., Professor, Head of the department, Department of informatics and computer techniques, Reshetnev Siberian State University of Science and Technology. Research interests: computer vision, image and video sequence processing, deep

learning, pattern recognition. The number of publications — 300. favorskaya@gmail.com; 31, Krasnoyarsky Rabochy Av., 660037, Krasnoyarsk, Russia; office phone: +7(3912)139-622.

Nishchhal Nishchhal — Master student, Department of informatics and computer techniques, Reshetnev Siberian State University of Science and Technology. Research interests: intelligent systems, deep learning. The number of publications — 18. nik.321g@yandex.ru; 31, Krasnoyarsky Rabochy Av., 660037, Krasnoyarsk, Russia; office phone: +7(3912)139-623.

DOI 10.15622/ia.21.5.5

T. YIFTER, YU. RAZOUMNY, V. LOBANOV

DEEP TRANSFER LEARNING OF SATELLITE IMAGERY FOR LAND USE AND LAND COVER CLASSIFICATION

Yifter T., Razoumny Yu., Lobanov V. **Deep Transfer Learning of Satellite Imagery for Land Use and Land Cover Classification.**

Abstract. Deep learning has been instrumental in solving difficult problems by automatically learning, from sample data, the rules (algorithms) that map an input to its respective output. Purpose: Perform land use landcover (LULC) classification using the training data of satellite imagery for Moscow region and compare the accuracy attained from different models. Methods: The accuracy attained for LULC classification using deep learning algorithm and satellite imagery data is dependent on both the model and the training dataset used. We have used state-of-the-art deep learning models and transfer learning, together with dataset appropriate for the models. Different methods were applied to fine tuning the models with different parameters and preparing the right dataset for training, including using data augmentation. Results: Four models of deep learning from Residual Network (ResNet) and Visual Geometry Group (VGG) namely: ResNet50, ResNet152, VGG16 and VGG19 has been used with transfer learning. Further training of the models is performed with training data collected from Sentinel-2 for the Moscow region and it is found that ResNet50 has given the highest accuracy for LULC classification for this region. Practical relevance: We have developed code that train the 4 models and make classification of the input image patches into one of the 10 classes (Annual Crop, Forest, Herbaceous Vegetation, Highway, Industrial, Pasture, Permanent Crop, Residential, River, and Sea&Lake).

Keywords: neural networks, deep transfer learning, land use land cover classification, satellite imagery.

1. Introduction. There is no doubt that soon artificial intelligence (AI) will penetrate every task that requires intelligent decisions based on learning from the collected data. It will change the way things are done, from language translation to self-driving cars and plenty of other tasks essential for a human being. It will have a remarkable effect on our day-to-day life. In recent years, AI's success is accelerated by advancement in deep learning (DL), a subset of AI dealing with learning from experience or data [1].

There are several reasons why deep learning is spreading across many fields so fast now. The availability of big data collected so far allows preparing massive training datasets required for DL applications. The rapid improvement of hardware and software computer components has enabled scientists to solve problems that were not solvable a few years ago. Science and engineering topics that were taken by Ph.D. students to do their complete thesis now can be addressed with a few lines of code on a decent computing device. Datasets can be collected online, and state-of-the-art algorithms can be applied to the same data leading to different outputs [2].

Meanwhile, the advances in Earth observation (EO) technologies have significantly improved the spatial, spectral, and temporal resolution of remotely sensed imagery. These achievements have allowed satellites to collect big EO data on a global scale related to different application scenarios. Simultaneously, deep learning is outpacing the other machine learning techniques in LULC classification tasks on satellite images [2].

Preparing training datasets is arguably the most challenging step in any machine learning project. EO data's nature adds additional complexity to the process because of higher spectral and radiometric resolution and other specific issues such as cloud and atmospheric noise.

The emergence of Google Earth Engine (GEE), the first publicly available cloud platform for big EO data analysis, has enabled scientists to process and analyze geospatial data in a multi-petabyte catalog without solving technical issues related to big data [3, 4, 5].

The platform proved to be helpful in preparing experiment-ready images with masked clouds and cloud shadows, snow/ice, and low-quality pixels, which is necessary for almost all remote sensing-related studies [6]. The increasing attention to GEE from researchers is pictured in Figure 1. This data is retrieved from an abstract and citation database Scopus with a search query: (TITLE-ABS-KEY ("earth engine") AND TITLE-ABS-KEY (google) AND DOCTYPE (ar OR re) AND PUBYEAR < 2022. However, only 15 from 691 published works relate to deep learning, making DL-related applications with GEE an urgent research topic.

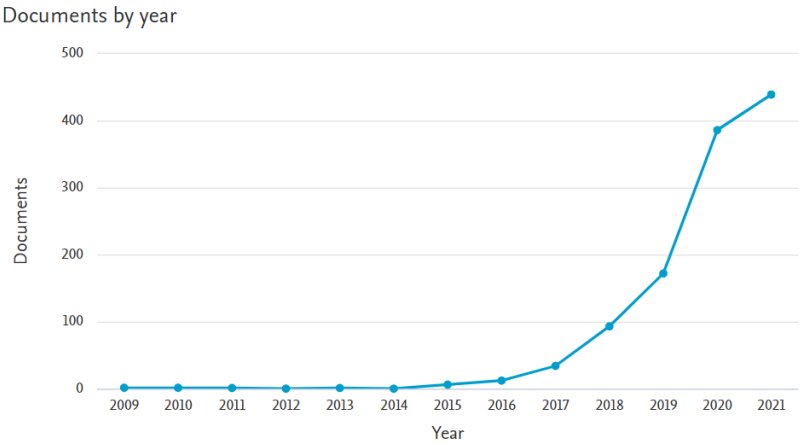


Fig. 1. Number of publications of the topic Deep Learning in Scopus

As it is complicated to construct large-scale well-annotated datasets due to the expense of data acquisition and labelling, a new DL approach called deep transfer learning (DTL) was developed allowing to solve the problem of insufficient training data. In DTL, the training data and test data are not required to be independent and identically distributed, and the target model does not need to be trained from scratch, which can significantly reduce the demand for training data and training time [7].

The technological advancements in both the data collection and processing have enabled a lot of new applications to pop up for solving real-world problems. The field of remote sensing has come a long way now to become very developed in collecting huge amounts of data that cover the entire earth and with continuous improvements in the quality of collected data. Even though different platforms can be used to collect remote sensing data, we are mainly focusing on data collected from satellites as it is the one which covers the entire earth. These huge continuously collected data have to be converted into insights to be useful for humanity. The complexity of the remotely sensed data hinders the extraction of insights out of it. Luckily, the advancement of technologies in the computing world has been a great motivation to process these huge data and extracting the right information at the right time. The recent development in Artificial Intelligence (AI) together with the progress made in the field of Computer Vision has given a lot of hope and hype for utilizing the remotely sensed data in different application areas. However, there is a gap between the continuously improving rate and quality of the collected data and its utilization for solving problems. This gap is a challenge, but at the same time, it is also an opportunity, that if the rate of utilization is improved, many difficult problems related to Agriculture, Climate, Environment, Economical and other areas will be tackled in an automated fashion.

The past decade has seen rapid progress in the capacity of DL especially for detecting objects from an image. In the state-of-the-art algorithms, the capability of DL has reached the human level, even better in some situations. But these DL algorithms cannot yield the same results when applied in remote sensing, and the reason for this is that there are some differences between these general image types and remote sensing images. So, there is a need to find the right approach for remote sensing data to benefit from the current achievement of DL in the general non-aerial image types. This approach requires understanding the spectral, spatial, and other characteristics of the remote sensing images, like texture.

With huge collected data and advancements in computational resources, remote sensing has become one of the beneficiaries of deep learning. At the same time, remote-sensing data presents some new

challenges for deep learning, because satellite image analysis raises unique issues that pose difficult new scientific questions. The authors in [8] discuss the unique characteristics of remote sensing data saying that it comes from geodetic measurements with quality controls that are completely dependent on the adequacy of a sensor, and they are geo-located, time variable and usually multi-modal, i.e., captured jointly by different sensors with different contents. These characteristics raise new challenges on how to deal with the data that comes with a variety of impacting variables and may require prior knowledge about how it has been acquired. In addition, despite the fast-growing data volume on a global scale that contains plenty of metadata, it is lacking adequate annotations for direct use of supervised machine learning-based approaches. Therefore, to effectively employ machine learning and deep learning techniques on such data, additional efforts are needed. Moreover, in many cases, remote sensing is to retrieve geophysical and geochemical quantities rather than land cover classification and object detection, which [8] indicates that expert-free use of deep learning techniques is still getting questioned. Further challenges include limited resolution, high dimensionality, redundancy within data, atmospheric and acquisition noise, calibration of spectral bands, and many other source-specific issues [2, 8].

The success of the methods of DL in applications of remote sensing imagery depends on both the data and the used DL model. DL is a data-hungry process that thrives when there is a lot of training data. But there are many situations where it is very hard to get enough training data. In situations like this, data augmentation has been used to compensate for the lack of enough training samples [9].

One thing that we observed from our experiments is the importance of understanding the data itself. Even though the enhancements in the algorithms have improved the accuracy of training and testing, at some point it becomes stagnant that the model almost makes no progress at all. As a result, we concentrated our focus on the collected training data and we improved its quality by carefully selecting representative data and its results with an improvement in the accuracy of the detection.

The remote sensing methods of the Earth together with the development of the algorithms for processing them are the best combination to tackle problems that are affecting the Earth we live on. The relevance of these studies is associated with such challenges as climate change and predictions, analysis, and recognition of technical objects, infrastructure, preservation and development of environmental systems that is important for the global system of Society and ecology as well as for many other practical tasks.

With the Petabyte-scale images collected from satellites, information extraction could be challenging. Deep learning methods, which thrive in data-driven applications, can be the right tool to create insights out of this huge collected data. However, the methods require the laborious preparation of training datasets. The quality of the collected training data is as important as developing the right deep learning model for attaining high accuracy with the inferences [10]. This data-centric approach helps to make some accuracy gains. With this study, we tried to classify land use land cover in the Moscow region, by using a publicly available EuroSAT dataset. This dataset is created from European urban areas. The reason that we used this dataset is that there is a lot of similarity between the Moscow region and the European urban areas where this dataset is collected from.

In this paper, to create a LULC map of the Moscow region, we prepared a high-quality annotated test and validation datasets using satellite imagery in the Google Earth Engine platform. Based on the existing Earth observation datasets, we train state-of-the-art deep learning algorithms through transfer learning to distinguish between ten LULC classes. And in doing so, we made the following contributions:

– We introduce Moscow patch-based LULC classification dataset (MoscowSAT) based on Sentinel-2 satellite images. Every image in the dataset is labeled and geo-referenced.

– Four models of CNNs namely ResNet-50, ResNet-152, VGG16 and VGG19 were used with the datasets EuroSAT and MoscowSAT for training and testing respectively. Both datasets are from sentinel-2 red-green-blue (RGB) images.

– We have streamlined the models by tweaking the learning rate hyperparameter for the highest accuracy.

– The performances of the four models for classifying the 10 classes are displayed in the confusion matrix.

2. Literature review. Lately, after the introduction of neural networks for solving general computer vision problems, neural networks for satellite imagery have been included in the hot research topics. The reason being satellite imagery hosts a lot of information that is very important for solving problems that span a wide area of fields. Neural networks have proven to be very important tools for extracting insights from these huge collected data. Here we have mentioned some of the researches conducted on the diverse fields of applications.

Before the introduction of deep learning, other machine learning algorithms were applied to solve different problems. The algorithms Support Vector Machine (SVM) and Random Forest (RF) were the most successful classifiers of these types of algorithms [2]. Deep learning is

different from these other machine learning algorithms in many ways. But we want to mention the two reasons that make it so popular. One of the reasons is that in the most problematic domains deep learning gives the highest classification accuracy, and the second reason is that it doesn't require manual feature extraction, and this enables end-to-end connections which automate feature extraction and perform the classification [2]. And as a result of its high accuracy and automation of almost the whole process of learning, it has been applied for solving problems in different application areas. Remote sensing has been applied successfully to a variety of classification and detection problems. The researchers in this study [11] presented an evaluation of fully convolutional neural networks (FCNNs) for road segmentation in satellite images. The authors' models show results, successfully extracting most of the roads in the test data set.

Detecting small objects such as vehicles in satellite images is a difficult problem. Deep convolutional neural networks (DNNs) can learn rich features from the training data automatically, and they have achieved state-of-the-art performance in many image classification databases. these authors [12] proposed a vehicle detection method in satellite images using a Deep Convolutional Neural Network(DNN). On a similar detection problem, these authors [13] present a hybrid DNN (HDNN) that HDNN significantly outperforms the traditional DNN for vehicle detection on satellite images. Still, with the problem of object detection, the authors of this paper [14] propose a new airport detection framework based on objectiveness detection techniques (e.g., BING) and Convolutional Neural Networks (CNN). Similarly, this work [15], proposed a method using convolutional neural networks (CNNs) for airport detection on optical satellite images.

Automatic object detection is a fundamental but challenging problem in the process of interpretation of satellite images. In this paper [16], an end-to-end multiscale convolutional neural network (MSCNN) is proposed, which is based on a unified multiscale backbone named EssNet for extracting features of diverse-scale objects in satellite images.

The paper on [17] describes a method for the effective semantic segmentation of satellite images and compares different object classifiers in terms of accuracy and execution time.

Very high resolution satellite imagery and image processing algorithms allow for the development of remote sensing applications. Recently, in addition to machine learning algorithms, deep learning methods have also been used to classify VHR images. In this paper [18], the authors compare the accuracy of the convolutional neural network (CNN)

algorithm with some machine learning methods, for the classification of a satellite image with 50 cm spatial resolution.

Hyperspectral Satellite Images (HSI) present a rich spectrum of input data. In this paper [19], the authors propose an approach to the reduction and classification of HSI using dual Convolutional Neural Networks (DCNN).

Deep learning with satellite images provides a way to make predictions about the distribution of poverty. The work in this paper [20] results in a test accuracy of 76% for the three countries, whose satellite imagery is used in the research.

The work in this paper [21] aims to perform individual tree recognition on the basis of satellite images using deep learning approaches for northern temperate mixed forests in the Primorsky Region of the Russian Far East. Using U-Net-like CNN, they obtained a mean accuracy score of up to 0.96. A similar treetop detection proposed in this paper [22] introduces a framework using the automatically generated pseudo labels from unsupervised treetop detectors to train the CNNs, which saves manual labelling efforts.

A deep learning algorithm, the convolution neural network (CNN), was applied in this research [23] to rapidly extract road blockage information. The kappa coefficient and the F1 score of the results were 77.60% and 87.95%, respectively.

The success of applying a deep learning algorithm for solving problems is also dependent on the dataset collected for the purpose of training. This article [24] focuses on evaluating the available and public remote-sensing datasets and different techniques for satellite image classification, using Convolution Neural Networks (CNNs), precisely, AlexNet architecture with SVM classifier.

The authors of this paper [25] proposed an approach that can be used to extend the footprint of the high-resolution images to generate new time frames or to downscale the remote sensing imagery based on a distant but structurally similar training image. And in another domain, this paper presents [26], the technique of processing satellite images with their subsequent placement in cartographic services. There are many application areas where deep learning makes a big impact, and it keeps growing. It can be said that any problem domain that has a lot of data is suitable to be solved using deep learning.

3. Materials and Methods. This work has utilized the result achieved by the authors [27], in which they have created a dataset based on Sentinel-2 satellite images covering 13 spectral bands constituting (Table 1) 10 classes with in total of 27,000 labeled and geo-referenced images. They

provided benchmarks for this dataset with its spectral bands using state-of-the-art deep Convolutional Neural Networks (CNNs) and with the proposed dataset, they achieved an overall classification accuracy of 98.57%.

Table 1. Sentinel-2 RGB bands and parameters

Sentinel-2 Band	Center Wavelength(nm)	Spectral Width(nm)	Spatial Resolution(m)
Band 2-Blue	490	65	10
Band 3-Green	560	35	10
Band 4-Red	665	30	10

The free availability of continuous Earth observation satellite data has motivated the processing of the data for some insights. This data processing has resulted in devising solutions to wide area domains including agriculture, environment, urban planning and disaster recovery. One part of the processing involves creating structured semantics out of the data, which is fundamental to creating LULC Classification [28].

LULC classification is an important task that enables us to understand the relationship between humans and the environment. Land Cover refers to the physical characteristics of the Earth's surface, such as vegetation, water, and soil, while Land Use refers to the purposes for which humans exploit the Land Cover such as changes made by anthropogenic activities [29, 30]. Therefore, the aim of LULC classification is to automatically provide labels describing the represented physical land type or how a land area is used. LULC changes (LULCC) are a very important measurement for monitoring man-made changes (e.g., deforestation, urbanization, and agriculture intensification) or natural phenomenon (e.g., droughts, floods, and natural fires). Therefore, LULC data enables the creation of detailed mappings to enable sustainable development [29].

A. Study Area and Data. The test data for this study is collected from the Moscow region, Russia (Figure 2). This area is selected for the reason of applying the EuroSAT dataset as training data, which is created from European urban areas and using the MoscowSAT as a test area. Since the Moscow region is on the European side of Russia, we want to investigate how well the training dataset of EuroSAT fits to the collected testing dataset from the Moscow region.

Fig. 2. Region of Interest (ROI) of the study – Moscow region, Russia

To perform this task, we utilize the results achieved by [27] and apply them to the case of the Moscow region. And to keep everything similar between the 2 datasets, we have used the same classes for the classification. The LULC classes used in this study are Annual Crop, Forest, Herbaceous Vegetation, Highway, Industrial, Pasture, Permanent Crop, Residential, River, and Sea&Lake.

Generating labelled datasets requires manual work. We created the test dataset first by gathering satellite images of the Moscow region from Sentinel-2 and then, we created a dataset of 2,000 georeferenced and labeled image patches, representing the 10 classes that we sought to classify. The image patches measure 64x64 pixels and have been manually checked.

The Sentinel-2 is one of the satellites under the European Space Agency (ESA) Earth observation mission. ESA has made the continuously collected satellite images freely available within its Copernicus program. Besides this, sentinel-2 imagery is preferable for our task of LULC classification for the following reasons: (1) This satellite image has 13 bands obtained from the MSI (Multispectral Imager) instrument, (2) Temporal resolution of Sentinel-2 is 10 days performed by one satellite and 5 days performed with two satellites that will make large amounts of observational data available. This satellite has a spatial resolution from 10 to 60 m. The satellite image is composed of the following bands 4 (red), bands 3 (green), bands 2 (blue), bands 8 (Near-Infrared) and bands 11 (SWIR, Short-Wave Infrared). Band 4 is useful for identifying types of vegetation, soil and urban features; band 3 provides excellent contrast between clear and turbid (muddy) water; band 2 is useful for land and

vegetation identification, forest type mapping, and identifying human-made features; while band 11 is useful for measuring soil moisture and vegetation, and it provides good contrast between various types of vegetation.

B. Deep transfer learning for image classification. In this study, we have used state-of-the-art deep learning algorithms with transfer learning (Figure 3) for finding the optimal algorithm that gives high accuracy. This is implemented using a jupyter notebook equipped with a TensorFlow GPU computing environment. We have used "Differential learning rates", which means different learning rates for different parts of the network during our training. The purpose of this is to divide the layers into various layer groups and set different learning rates for each group so that we get the best results.

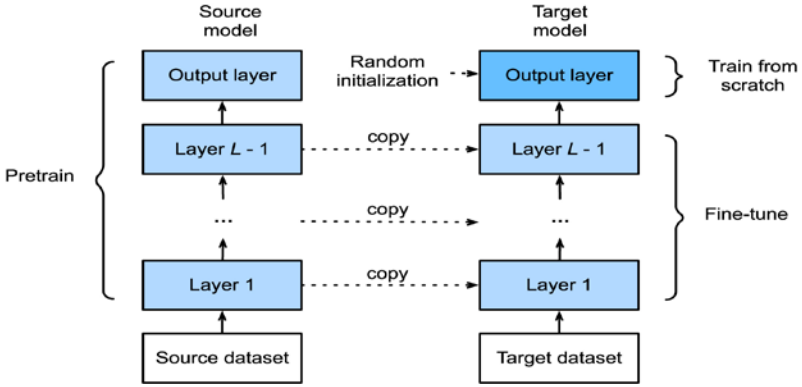


Fig. 3. Transfer Learning

This study used a supervised approach – creating training samples from input images as an automatic classification process. The final result of this process is a labeled dataset for training a classifier. The process starts by collecting data and ends by making predictions of the data into its respective classes. This process is depicted in Figure 4. And from the process implementation side, it can be summarized as: loading data from hard disk, data splitting into training and testing dataset, training CNN model, and evaluating model performance.

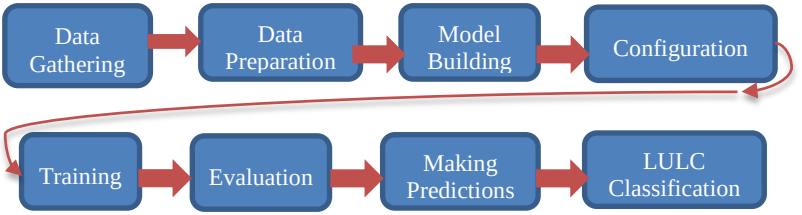


Fig. 4. Training process of the LULC classification

We divided the dataset into a training set (80%), validation set (10%) and test set (10%), in order to perform the training, validation and testing, respectively. We have used ResaNet 50 with transfer learning to train the model. This artificial neural network is basically composed of a 7×7 convolutional layer with 64 output channels and a stride of 2 followed by the 3×3 maximum pooling layer with a stride of 2. The batch normalization layer is added after each convolutional layer. Finally, the layers are preceded by one Flatten Layer, two Dense Layers and a Softmax Layer. The model architecture is depicted in Figure 5.

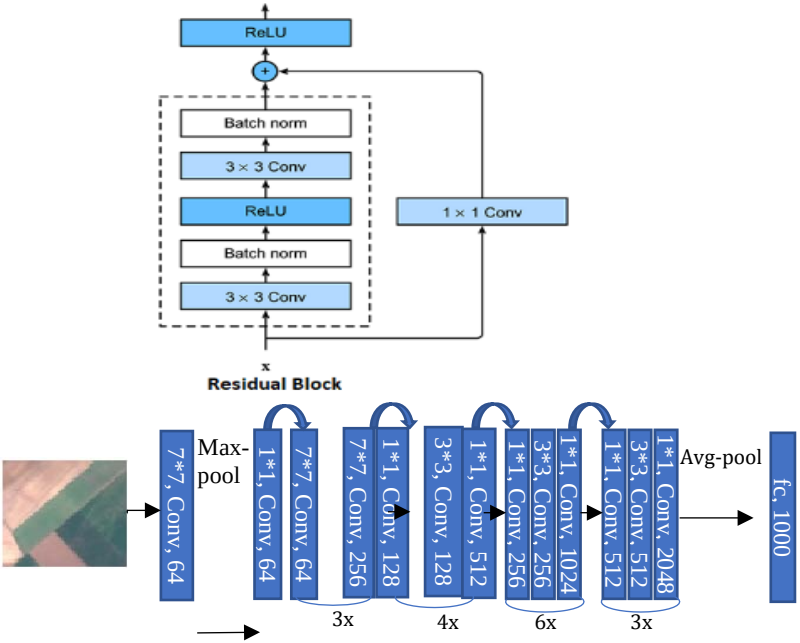
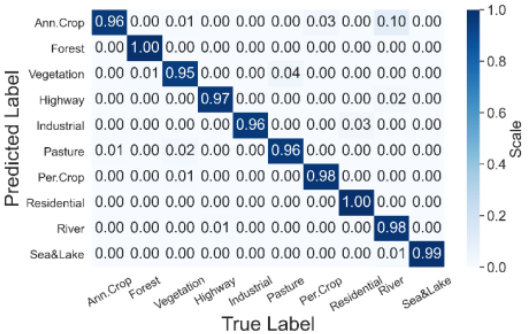


Fig. 5. ResNet 50 Model Architecture

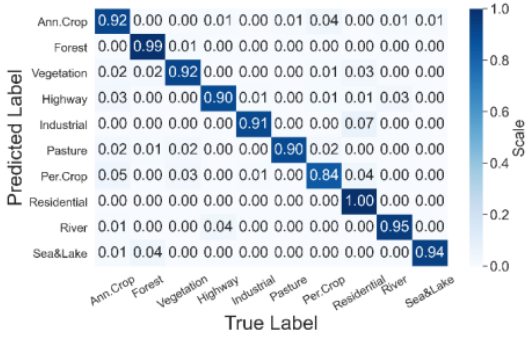
4. Results and Discussion. We run the classification algorithms ResNet 50, ResNet 152, VGG 16 and VGG 19 with transfer learning using the EuroSAT dataset and then continue training models with MoscowSAT dataset for the specifics related to the Moscow region. The used model and the dataset are crucial to achieving high accuracy. We have carefully created the MoscowSAT dataset to reflect the ground truth so that this part of the learning will be specific for the region. The results of our code analysis in the confusion matrix are displayed in Figure 6 from *a* to *d*. The confusion matrix is the results of the fine tunned training on the test dataset MoscowSAT using satellite images in the RGB color space. We discovered that ResNet 50 gives the highest accuracy, which is 98% in the training and 97.5% in the testing. As can be seen in the confusion matrix, there was high precision for almost all of the classes except there was some miss classification between permanent and annual crops (Figure 7). The comparison of the results of the three NN algorithms used is displayed in Table 2.

Table 2. Comparison of accuracy

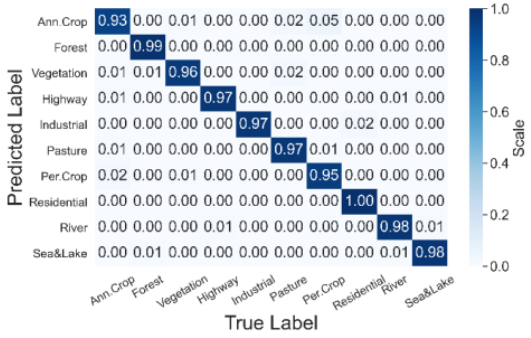
Accuracy	ResNet50	ResNet152	VGG16	VGG19
	0.975292588	0.932193944	0.952134497	0.968265



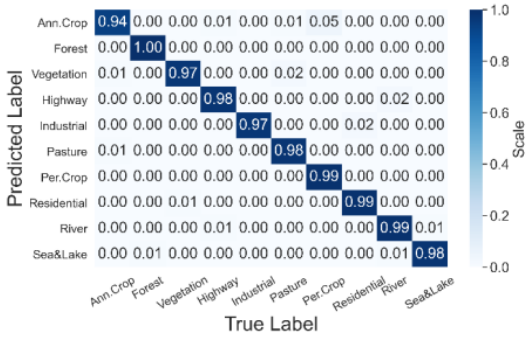
a) ResNet-50



b) ResNet-152



c) VGG 16



d) VGG 19

Fig. 6. Results of the fine tunned training on the test dataset MoscowSAT using satellite images in the RGB color space

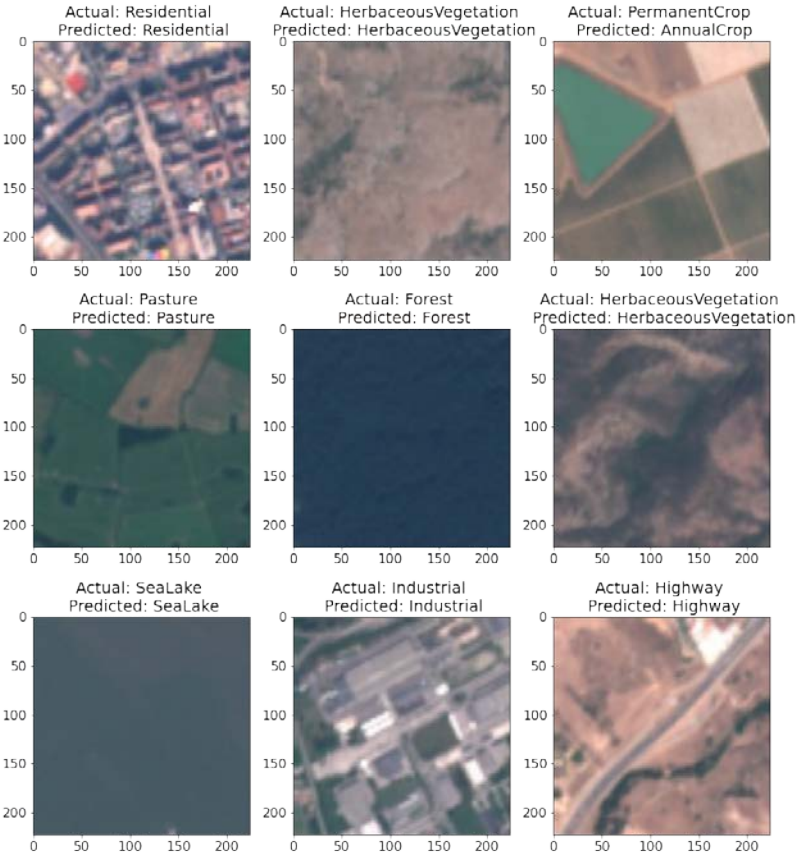


Fig. 7. ResNet-50 sample prediction output. The text displays the actual and predicted levels of each image. The permanent crop is confused with annual crop in this sample output

5. Conclusion. Regional land use planning and monitoring remain to be an important process that enables appropriate and optimal usage of resources. Despite many proposed models in the studies, the task remained a challenging problem. In this paper, we have used transfer learning of deep learning models for creating LULC of the Moscow region. We have fine-tuned 4 models of CNN, namely ResNet50, ResNet152, VGG16 and VGG19 with transfer learning and adjusting the learning rate for the optimal accuracy.

In general, our study is tailored toward creating the right tools to automate the extraction of information from satellite imagery. We have demonstrated the capability of remote sensing data together with deep learning for solving real-world problems. We achieved this by working towards creating the LULC Classification of the Moscow region. This study presented the results of LULC classification using Sentinel-2 satellite RGB imagery as input, the CNN model as the classifier, and the Moscow region area as the location for this study. The results showed that the ResNet50 CNN model together with transfer learning achieved high accuracy in classifying the 10 classes.

References

1. Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, Mar. 2015. doi: 10.1038/nature14539.
2. A. Vali, S. Comai, and M. Matteucci, “Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: A review,” Remote Sensing, vol. 12, no. 15. Multidisciplinary Digital Publishing Institute, p. 2495, Aug. 03, 2020. doi: 10.3390/RS12152495.
3. N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, “Google Earth Engine: Planetary-scale geospatial analysis for everyone,” Remote Sens. Environ., vol. 202, pp. 18–27, Mar. 2017, doi: 10.1016/j.rse.2017.06.031.
4. L. Kumar and O. Mutanga, “Google Earth Engine applications since inception: Usage, trends, and potential,” Remote Sens., vol. 10, no. 10, 2018, doi: 10.3390/rs10101509.
5. L. Parente, E. Taquary, A.P. Silva, C. Souza, and L. Ferreira, “Next generation mapping: Combining deep learning, cloud computing, and big remote sensing data,” Remote Sens., vol. 11, no. 23, 2019, doi: 10.3390/rs11232881.
6. H. Li et al., “A Google Earth Engine-enabled software for efficiently generating high-quality user-ready Landsat mosaic images,” Environ. Model. Softw., vol. 112, pp. 16–22, 2019, doi: 10.1016/j.envsoft.2018.11.004.
7. C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, “A survey on deep transfer learning,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 11141 LNCS, pp. 270–279, 2018, doi: 10.1007/978-3-030-01424-7_27.
8. X.X. Zhu et al., “Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources,” IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 4. Institute of Electrical and Electronics Engineers Inc., pp. 8–36, Mar. 2017. doi: 10.1109/MGRS.2017.2762307.
9. J. Song, S. Gao, Y. Zhu, and C. Ma, “A survey of remote sensing image classification based on CNNs,” Big Earth Data, vol. 3, no. 3, pp. 232–254, Mar. 2019, doi: 10.1080/20964471.2019.1657720.
10. S.E. Whang et al., “Data Collection and Quality Challenges in Deep Learning: A Data-Centric AI Perspective,” Dec. 2021, doi: 10.48550/arxiv.2112.06409.
11. C. Henry, S.M. Azimi, and N. Merkle, “Road segmentation in SAR satellite images with deep fully convolutional neural networks,” IEEE Geosci. Remote Sens. Lett., vol. 15, no. 12, pp. 1867–1871, 2018, doi: 10.1109/LGRS.2018.2864342.
12. Q.J.C. Cheng, “Deep neural networks-based vehicle detection in satellite images”.
13. X. Chen, S. Xiang, C.L. Liu, and C.H. Pan, “Vehicle detection in satellite images by hybrid deep convolutional neural networks,” IEEE Geosci. Remote Sens. Lett., vol. 11, no. 10, pp. 1797–1801, 2014, doi: 10.1109/LGRS.2014.2309695.

14. H. Wu, H. Zhang, J. Zhang, and F. Xu, “Fast aircraft detection in satellite images based on convolutional neural networks,” Proc. - Int. Conf. Image Process. ICIP, vol. 2015-December, pp. 4210–4214, Mar. 2015, doi: 10.1109/ICIP.2015.7351599.
15. P. Zhang, X. Niu, Y. Dou, and F. Xia, “Airport detection on optical satellite images using deep convolutional neural networks,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 8, pp. 1183–1187, Mar. 2017, doi: 10.1109/LGRS.2017.2673118.
16. Q. Yao, X. Hu, and H. Lei, “Multiscale Convolutional Neural Networks for Geospatial Object Detection in VHR Satellite Images,” IEEE Geosci. Remote Sens. Lett., vol. 18, no. 1, pp. 23–27, Mar. 2021, doi: 10.1109/LGRS.2020.2967819.
17. Y.H. Robinson, S. Vimal, M. Khari, F.C.L. Hernández, and R.G. Crespo, “Tree-based convolutional neural networks for object classification in segmented satellite images:,” https://doi.org/10.1177/1094342020945026, Mar. 2020, doi: 10.1177/1094342020945026.
18. M. Mohammadi and A. Sharifi, “Evaluation of Convolutional Neural Networks for Urban Mapping Using Satellite Images,” J. Indian Soc. Remote Sens. 2021 499, vol. 49, no. 9, pp. 2125–2131, Mar. 2021, doi: 10.1007/S12524-021-01382-X.
19. M. Hamouda and M.S. Bouhlel, “Dual Convolutional Neural Networks for Hyperspectral Satellite Images Classification (DCNN-HSI),” Commun. Comput. Inf. Sci., vol. 1332, pp. 369–376, 2020, doi: 10.1007/978-3-030-63820-7_42.
20. A. Okaidat, S. Melhem, H. Alenezi, and R. Duwairi, “Using Convolutional Neural Networks on Satellite Images to Predict Poverty,” 2021 12th Int. Conf. Inf. Commun. Syst. ICICS 2021, pp. 164–170, Mar. 2021, doi: 10.1109/ICICS52457.2021.9464598.
21. K.A. Korznikov, D.E. Kislov, J. Altman, J. Doležal, A.S. Vozmishcheva, and P.V. Krestov, “Using U-Net-Like Deep Convolutional Neural Networks for Precise Tree Recognition in Very High Resolution RGB (Red, Green, Blue) Satellite Images,” For. 2021, Vol. 12, Page 66, vol. 12, no. 1, p. 66, Mar. 2021, doi: 10.3390/F12010066.
22. C. Xiao, R. Qin, and X. Huang, “Treetop detection using convolutional neural networks trained through automatically generated pseudo labels,” https://doi.org/10.1080/01431161.2019.1698075, vol. 41, no. 8, pp. 3010–3030, Mar. 2019, doi: 10.1080/01431161.2019.1698075.
23. B. Yang et al., “Extraction of road blockage information for the Jiuzhaigou earthquake based on a convolution neural network and very-high-resolution satellite images,” Earth Sci. Informatics, vol. 13, no. 1, pp. 115–127, Mar. 2020, doi: 10.1007/S12145-019-00413-Z.
24. M.A. Shafaey, M.A.-M. Salem, M.N. Al-Berry, H.M. Ebied, and M.F. Tolba, “Remote Sensing Image Classification Based on Convolutional Neural Networks,” Adv. Intell. Syst. Comput., vol. 1153 AISC, pp. 353–361, Mar. 2020, doi: 10.1007/978-3-030-44289-7_33.
25. F. Oriani, M.F. McCabe, and G. Mariethoz, “Downscaling Multispectral Satellite Images without Colocated High-Resolution Data: A Stochastic Approach Based on Training Images,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 4, pp. 3209–3225, Mar. 2021, doi: 10.1109/TGRS.2020.3008015.
26. S.S. Dymkova, “Conjunction and synchronization methods of earth satellite images with local cartographic data,” 2020 Syst. Signals Gener. Process. F. Board Commun., Mar. 2020, doi: 10.1109/IEEECONF48371.2020.9078561.
27. P. Helber, B. Bischke, A. Dengel, and D. Borth, “Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 12, no. 7, pp. 2217–2226, 2019, doi: 10.1109/JSTARS.2019.2918242.
28. Y. Yang and S. Newsam, “Bag-of-visual-words and spatial extensions for land-use classification,” GIS Proc. ACM Int. Symp. Adv. Geogr. Inf. Syst., no. January 2010, pp. 270–279, 2010, doi: 10.1145/1869790.1869829.

29. M.E.D. Chaves, M.C.A. Picoli, and I.D. Sanches, “Recent Applications of Landsat 8/OLI and Sentinel-2/MSI for Land Use and Land Cover Mapping: A Systematic Review,” Remote Sens., vol. 12, no. 18, p. 3062, 2020, doi: 10.3390/rs12183062.
30. C. Gómez, J.C. White, and M.A. Wulder, “Optical remotely sensed time series data for land cover classification: A review,” ISPRS J. Photogramm. Remote Sens., vol. 116, pp. 55–72, 2016, doi: 10.1016/j.isprsjprs.2016.03.008.

Yifter Teklay — Graduate student, Peoples' Friendship University of Russia. Research interests: artificial Intelligence, Remote sensing. The number of publications — 4. teklay_ty@pfur.ru; 6, Miklukho-Maklay St., 117198, Moscow, Russia; office phone: +7(977)436-4584.

Razoumny Yury — Professor, Director, Engineering academy, Peoples' Friendship University of Russia. Research interests: dynamics, ballistics and motion control of aircraft. razoumny-yun@rudn.ru; 6, Miklukho-Maklay St., 117198, Moscow, Russia; office phone: +7(495)952-0829.

Lobanov Vasiliy — Senior lecturer, Peoples' Friendship University of Russia. Research interests: spatial analysis, satellite image processing, geoinformation, geospatial science, Earth observation. The number of publications — 3. lobanov-vk@rudn.ru; 6, Miklukho-Maklay St., 117198, Moscow, Russia; office phone: +7(926)899-9812.

УДК 528.854 **DOI** 10.15622/ia.21.5.5

Т.Т. Уифтер, Ю.Н. Разумный, В.К. Лобанов

ГЛУБОКОЕ ТРАНСФЕРНОЕ ОБУЧЕНИЕ НА ОСНОВЕ СПУТНИКОВЫХ ИЗОБРАЖЕНИЙ ДЛЯ КЛАССИФИКАЦИИ ЗЕМЛЕПОЛЬЗОВАНИЯ И ЗЕМНОГО ПОКРОВА

Уифтер Т.Т., Разумный Ю.Н., Лобанов В.К. **Глубокое трансферное обучение на основе спутниковых изображений для классификации землепользования и земного покрова.**

Аннотация. Алгоритмы глубокого обучения сыграли важную роль в решении многих комплексных задач, за счет автоматического изучения правил (алгоритмов) на основе выборочных данных, которые затем сопоставляют входные данные с соответствующими выходными данными. Цель работы: выполнить классификацию земных покровов (LULC) спутниковых снимков Московской области на основе обучающих данных и сравнить точность классификации, полученной с применением ряда моделей глубокого обучения. Методы: точность, достигаемая при классификации земных покровов с использованием алгоритмов глубокого обучения и данных космической съёмки, зависит как от конкретной модели глубокого обучения, так и от используемой обучающей выборки. Мы использовали наиболее современные модели глубокого обучения и обучения с подкреплением вкупе с релевантным набором обучающих данных. Для тонкой корректировки параметров моделей и подготовки обучающего набора данных применялись разливные методы, в том числе аугментация данных. Результаты: Применены четыре модели глубокого обучения на основе архитектур Residual Network (ResNet) и Visual Geometry Group (VGG) на основе обучения с подкреплением: ResNet50, ResNet152, VGG16 и VGG19. Последующее до-обучение моделей выполнялось с использованием обучающих данных, собранных спутником ДЗЗ Sentinel-2 на территории Московской области. На основе оценки результатов, архитектура ResNet50 дала наиболее высокую точность классификации земных покровов на территории выбранного региона. Практическая значимость: авторы разработали алгоритм обучения четырёх моделей глубокого обучения с последующей классификацией фрагментов входного космического снимка с присвоением одного из 10 классов (однолетние культуры, лесной покров, травянистая растительность, автодороги и шоссе, промышленная застройка, пастбища, многолетние культуры, жилая застройка, реки и озера).

Ключевые слова: нейронные сети, глубокое трансферное обучение, классификация землепользования, спутниковые снимки.

Литература

1. Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, Mar. 2015. doi: 10.1038/nature14539.
2. A. Vali, S. Comai, and M. Matteucci, “Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: A review,” Remote Sensing, vol. 12, no. 15. Multidisciplinary Digital Publishing Institute, p. 2495, Aug. 03, 2020. doi: 10.3390/RS12152495.
3. N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, “Google Earth Engine: Planetary-scale geospatial analysis for everyone,” Remote Sens. Environ., vol. 202, pp. 18–27, Mar. 2017, doi: 10.1016/j.rse.2017.06.031.
4. L. Kumar and O. Mutanga, “Google Earth Engine applications since inception: Usage, trends, and potential,” Remote Sens., vol. 10, no. 10, 2018, doi: 10.3390/rs10101509.

5. L. Parente, E. Taquary, A.P. Silva, C. Souza, and L. Ferreira, “Next generation mapping: Combining deep learning, cloud computing, and big remote sensing data,” Remote Sens., vol. 11, no. 23, 2019, doi: 10.3390/rs11232881.
6. H. Li et al., “A Google Earth Engine-enabled software for efficiently generating high-quality user-ready Landsat mosaic images,” Environ. Model. Softw., vol. 112, pp. 16–22, 2019, doi: 10.1016/j.envsoft.2018.11.004.
7. C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, “A survey on deep transfer learning,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 11141 LNCS, pp. 270–279, 2018, doi: 10.1007/978-3-030-01424-7_27.
8. X.X. Zhu et al., “Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources,” IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 4. Institute of Electrical and Electronics Engineers Inc., pp. 8–36, Mar. 2017. doi: 10.1109/MGRS.2017.2762307.
9. J. Song, S. Gao, Y. Zhu, and C. Ma, “A survey of remote sensing image classification based on CNNs,” Big Earth Data, vol. 3, no. 3, pp. 232–254, Mar. 2019, doi: 10.1080/20964471.2019.1657720.
10. S.E. Whang et al., “Data Collection and Quality Challenges in Deep Learning: A Data-Centric AI Perspective,” Dec. 2021, doi: 10.48550/arxiv.2112.06409.
11. C. Henry, S.M. Azimi, and N. Merkle, “Road segmentation in SAR satellite images with deep fully convolutional neural networks,” IEEE Geosci. Remote Sens. Lett., vol. 15, no. 12, pp. 1867–1871, 2018, doi: 10.1109/LGRS.2018.2864342.
12. Q.J.C. Cheng, “Deep neural networks-based vehicle detection in satellite images”.
13. X. Chen, S. Xiang, C.L. Liu, and C.H. Pan, “Vehicle detection in satellite images by hybrid deep convolutional neural networks,” IEEE Geosci. Remote Sens. Lett., vol. 11, no. 10, pp. 1797–1801, 2014, doi: 10.1109/LGRS.2014.2309695.
14. H. Wu, H. Zhang, J. Zhang, and F. Xu, “Fast aircraft detection in satellite images based on convolutional neural networks,” Proc. - Int. Conf. Image Process. ICIP, vol. 2015-December, pp. 4210–4214, Mar. 2015, doi: 10.1109/ICIP.2015.7351599.
15. P. Zhang, X. Niu, Y. Dou, and F. Xia, “Airport detection on optical satellite images using deep convolutional neural networks,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 8, pp. 1183–1187, Mar. 2017, doi: 10.1109/LGRS.2017.2673118.
16. Q. Yao, X. Hu, and H. Lei, “Multiscale Convolutional Neural Networks for Geospatial Object Detection in VHR Satellite Images,” IEEE Geosci. Remote Sens. Lett., vol. 18, no. 1, pp. 23–27, Mar. 2021, doi: 10.1109/LGRS.2020.2967819.
17. Y.H. Robinson, S. Vimal, M. Khari, F.C.L. Hernández, and R.G. Crespo, “Tree-based convolutional neural networks for object classification in segmented satellite images:,” https://doi.org/10.1177/1094342020945026, Mar. 2020, doi: 10.1177/1094342020945026.
18. M. Mohammadi and A. Sharifi, “Evaluation of Convolutional Neural Networks for Urban Mapping Using Satellite Images,” J. Indian Soc. Remote Sens. 2021 499, vol. 49, no. 9, pp. 2125–2131, Mar. 2021, doi: 10.1007/S12524-021-01382-X.
19. M. Hamouda and M.S. Bouhlel, “Dual Convolutional Neural Networks for Hyperspectral Satellite Images Classification (DCNN-HSI),” Commun. Comput. Inf. Sci., vol. 1332, pp. 369–376, 2020, doi: 10.1007/978-3-030-63820-7_42.
20. A. Okaidat, S. Melhem, H. Alenezi, and R. Duwairi, “Using Convolutional Neural Networks on Satellite Images to Predict Poverty,” 2021 12th Int. Conf. Inf. Commun. Syst. ICICS 2021, pp. 164–170, Mar. 2021, doi: 10.1109/ICICS52457.2021.9464598.
21. K.A. Korznikov, D.E. Kislov, J. Altman, J. Doležal, A.S. Vozmishcheva, and P.V. Krestov, “Using U-Net-Like Deep Convolutional Neural Networks for Precise Tree Recognition in Very High Resolution RGB (Red, Green, Blue) Satellite Images,” For. 2021, Vol. 12, Page 66, vol. 12, no. 1, p. 66, Mar. 2021, doi: 10.3390/F12010066.

22. C. Xiao, R. Qin, and X. Huang, “Treetop detection using convolutional neural networks trained through automatically generated pseudo labels,” https://doi.org/10.1080/01431161.2019.1698075, vol. 41, no. 8, pp. 3010–3030, Mar. 2019, doi: 10.1080/01431161.2019.1698075.
23. B. Yang et al., “Extraction of road blockage information for the Jiuzhaigou earthquake based on a convolution neural network and very-high-resolution satellite images,” Earth Sci. Informatics, vol. 13, no. 1, pp. 115–127, Mar. 2020, doi: 10.1007/S12145-019-00413-Z.
24. M.A. Shafaey, M.A.-M. Salem, M.N. Al-Berry, H.M. Ebied, and M.F. Tolba, “Remote Sensing Image Classification Based on Convolutional Neural Networks,” Adv. Intell. Syst. Comput., vol. 1153 AISC, pp. 353–361, Mar. 2020, doi: 10.1007/978-3-030-44289-7_33.
25. F. Oriani, M.F. McCabe, and G. Mariethoz, “Downscaling Multispectral Satellite Images without Colocated High-Resolution Data: A Stochastic Approach Based on Training Images,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 4, pp. 3209–3225, Mar. 2021, doi: 10.1109/TGRS.2020.3008015.
26. S.S. Dymkova, “Conjunction and synchronization methods of earth satellite images with local cartographic data,” 2020 Syst. Signals Gener. Process. F. Board Commun., Mar. 2020, doi: 10.1109/IEEECONF48371.2020.9078561.
27. P. Helber, B. Bischke, A. Dengel, and D. Borth, “Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 12, no. 7, pp. 2217–2226, 2019, doi: 10.1109/JSTARS.2019.2918242.
28. Y. Yang and S. Newsam, “Bag-of-visual-words and spatial extensions for land-use classification,” GIS Proc. ACM Int. Symp. Adv. Geogr. Inf. Syst., no. January 2010, pp. 270–279, 2010, doi: 10.1145/1869790.1869829.
29. M.E.D. Chaves, M.C.A. Picoli, and I.D. Sanches, “Recent Applications of Landsat 8/OLI and Sentinel-2/MSI for Land Use and Land Cover Mapping: A Systematic Review,” Remote Sens., vol. 12, no. 18, p. 3062, 2020, doi: 10.3390/rs12183062.
30. C. Gómez, J.C. White, and M.A. Wulder, “Optical remotely sensed time series data for land cover classification: A review,” ISPRS J. Photogramm. Remote Sens., vol. 116, pp. 55–72, 2016, doi: 10.1016/j.isprsjprs.2016.03.008.

Уифтер Теклай Тесфазги — аспирант, Российский университет дружбы народов. Область научных интересов: mechanical and mechatronics, computer engineering. Число научных публикаций — 4. teklay_ty@pfur.ru; улица Миклухо-Маклая, 6, 117198, Москва, Россия; р.т.: +7(977)436-4584.

Разумный Юрий Николаевич — профессор, директор, инженерная академия, Российский университет дружбы народов. Область научных интересов: динамика, баллистика и управление движением летательных аппаратов. razoumny-yun@rudn.ru; улица Миклухо-Маклая, 6, 117198, Москва, Россия; р.т.: +7(495)952-0829.

Лобанов Василий Константинович — старший преподаватель, Российский университет дружбы народов. Область научных интересов: пространственный анализ, обработка спутниковых изображений, геоинформация, геопространственная наука, наблюдение за Землей. Число научных публикаций — 3. lobanov-vk@rudn.ru; улица Миклухо-Маклая, 6, 117198, Москва, Россия; р.т.: +7(926)899-9812.

УДК 621.396 **DOI** 10.15622/ia.21.5.6

С.В. Дворников, С.С. Дворников, А.А. Устинов

КОРРЕЛЯЦИОННЫЕ СВОЙСТВА КОЭФФИЦИЕНТОВ КРАТНОМАСШТАБНОГО ПРЕОБРАЗОВАНИЯ ТИПОВЫХ ИЗОБРАЖЕНИЙ

Дворников С.В., Дворников С.С., Устинов А.А. **Корреляционные свойства коэффициентов кратномасштабного преобразования типовых изображений.**

Аннотация. Увеличивающийся поток фото и видеоинформации, передаваемый по каналам инфокоммуникационных систем и комплексов, стимулирует к поиску эффективных алгоритмов сжатия, позволяющих существенно снизить объем передаваемого трафика, при сохранении его качества. В общем случае, в основе алгоритмов сжатия лежат операции преобразования коррелированных значений яркостей пикселов матрицы изображения в их некоррелированные параметры, с последующим кодированием полученных коэффициентов преобразования. Поскольку основные известные декоррелирующие преобразования являются квазиоптимальными, то задача поиска преобразований, учитывающих изменения статистических характеристик сжимаемых видеоданных, все еще является актуальной. Указанные обстоятельства определили направление проведенного исследования, связанного с анализом декоррелирующих свойств формируемых вейвлет-коэффициентов, получаемых в результате кратномасштабного преобразования изображений. Основным результатом проведенного исследования явилось установления факта того, что вейвлет-коэффициенты кратномасштабного преобразования имеют структуру вложенных матриц, определенных как субматрицы. Поэтому, корреляционный анализ вейвлет-коэффициентов преобразования, целесообразно проводить раздельно для элементов каждой субматрицы на каждом уровне декомпозиции (разложения). Основным теоретическим результатом явилось доказательство того, что ядром каждого последующего уровня кратномасштабного преобразования является матрица, состоящая из вейвлет-коэффициентов предшествующего уровня декомпозиции. Именно данный факт и позволяет сделать вывод о зависимости соответствующих элементов соседних уровней. Кроме того, установлено, что между вейвлет-коэффициентами внутри локальной области изображения размером 8×8 пикселей существует линейная зависимость. При этом максимальная корреляция элементами субматриц непосредственно определяется формой их представления, и наблюдается между соседними элементами находящимися, соответственно в строке, столбце или по диагонали, что подтверждается характером рассеивания. Полученные результаты подтверждены на проведенном анализе выборок, более чем из двухсот типовых изображений. При этом обосновано, что между низкочастотными вейвлет-коэффициентами кратномасштабного преобразования верхнего уровня разложения сохраняются примерно одинаковые зависимости равномерно по всем направлениям. Практическая значимость исследования определяется тем, что все полученные в ходе его проведения результаты подтверждают наличие характерных зависимостей между вейвлет-коэффициентами преобразования на различных уровнях разложения изображений. Данный факт указывает на возможность достижения более высоких значений коэффициентов сжатия видеоданных в процессе их кодирования. Дальнейшие исследования авторы связывают с разработкой математической модели адаптивного арифметического кодирования видеоданных и изображений, учитывающей корреляционные свойства вейвлет-коэффициентов кратномасштабного преобразования.

Ключевые слова: сжатие изображений, вейвлет-коэффициенты кратномасштабного преобразования, преобразование коррелированных значений, декоррелирующие преобразования.

1. Введение. В настоящее время в основе всех известных алгоритмов сжатия видеоданных лежит выполнение двух основных операций [1-4]: преобразование коррелированных значений яркостей пикселов матрицы изображения в их некоррелированные параметры – коэффициенты преобразования; кодирование коэффициентов преобразования [5, 6].

Основной задачей первой операции является получение коэффициентов преобразования, характеризующихся неравномерным распределением значений их абсолютных значений. При этом значения коэффициентов перераспределяется так, что при кодировании существенной их части, в ходе выполнения второй операции, используют незначительное число битов, либо вообще отказываются от кодирования коэффициентов преобразования, приравнивая их к нулю, что обеспечивает сжатие изображений [7, 8].

Основные декоррелирующие преобразования, используемые в настоящее время в различных стандартах сжатия, известны и достаточно хорошо исследованы. К таким преобразованиям относятся дискретное косинусное преобразование [9] и различные виды вейвлет-преобразований [3]. В работах [9, 10] показано, что данные преобразования являются квазиоптимальными с позиций перераспределения энергии, характерных для большинства типовых изображений. Кроме того, задача построения нового класса преобразований, учитывающих локальные изменения статистических характеристик сжимаемых изображений, является сложнейшей задачей и в настоящее время не решена.

С другой стороны, существующие методы кодирования, в том числе и арифметическое кодирование [4, 11], разрабатывались как методы кодирования последовательностей символов, то есть без учета того, каким образом эти последовательности были получены. Поэтому такое преобразование не учитывает их структурные особенности. Так, например, в большинстве случаев реализация арифметического кодирования учитывает только значения вероятностей кодируемых символов. При этом предполагается, что последующее построение кодеков, ориентированных на кодирование коэффициентов преобразования заданного вида, как раз и позволит получить большие значения коэффициентов сжатия [12].

В связи с этим основным содержанием данной статьи является представление результатов исследования свойств вейвлет-

коэффициентов, получаемых в результате кратномасштабного преобразования изображений, получившего широкое практическое применение в современных кодеках и рекомендациях [13-15].

2. Кратномасштабные преобразования на основе аналитического аппарата линейной алгебры. В общем случае, кратномасштабный анализ основан на последовательной декомпозиции сигнала (применительно к рассматриваемой тематике, далее по тексту – изображений) в ортогональном базисе вейвлетов, образованном путем временных сдвигов и кратномасштабных преобразований исходной порождающей функции [16]. В результате открывается возможность выделить характерные особенности исследуемых изображений в области локализации формирующих вейвлетов [17].

Теоретические основы кратномасштабного анализа как инструмента построения вейвлет-базисов в интересах анализа тонкой структуры изображений, были разработаны С. Малла и И. Мейром [18]. В теории вейвлетов понятие кратномасштабного анализа является фундаментальным [19], поскольку оно определило развитие нескольких направлений в радиотехнике и телекоммуникациях, связанных с обработкой и сжатием изображений.

Сущность кратномасштабного анализа основана на иерархическом представлении исходного пространства изображения $L^2(\mathbf{R})$ в виде упорядоченной системы вложенных субпространств $V_j \subset L^2(\mathbf{R})$ [18], где $j \in \mathbf{Z}$, которые отличаются друг от друга значением масштабируемой переменной, что позволяет формируемые на каждом уровне декомпозиции пространства представлять следующим образом:

$$V_j(\mathbf{R}),\ j \in \mathbf{Z}.$$

В общем случае правомерность реализация кратномасштабного преобразования определяется выполнением шести базовых условий [18]:

вложенности, когда постановка $V_j \subset V_{j+1}$ справедлива для всех значений $j \in \mathbf{Z}$;

полноты и плотности разбиения, при котором $\bigcup_{j \in \mathbf{Z}} V_j$ плотно в $L^2(\mathbf{R})$;

ортогональности, т.е. $\bigcap_{j \in \mathbf{Z}} V_j = 0$;

сохранения в пределах субпространства при любых сдвигах функций, $f(t) \in V_j \Leftrightarrow f(t+1) \in V_j$;

масштабирование любой функции $f(t) \in V_j$ *по аргументу в два раза, перемещает функцию в соседнее субпространство*, т.е. $f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1}$ и $f(t) \in V_j \Leftrightarrow f(t/2) \in V_{j-1}$;

существования такой скейлинг-функции $r(t) \in V_0$, *целочисленные сдвиги которой по аргументу образуют ортонормированный базис пространства* V_0: $f_{0,k}(t) = f(t-k)$, $k \in \mathbf{Z}$.

Тогда любую функцию $s(t)$ из пространства $L^2(\mathbf{R})$ можно представить совокупностью ее отображений $s_j(t)$ в каждом из вложенных субпространств таким образом, что:

$$s(t) = \lim_{j \to \infty} s_j(t), \quad j \in \mathbf{Z},$$

где параметр j фактически определяет размерность сформированного функционального базиса, т.е. глубину или точность детализации исходной функции $s(t)$ с использованием ее масштабированных копий.

Очевидно, что устремляя $j \to \infty$ можно обеспечить фактически полное восстановление исходного сигнала по его копиям. Однако такой подход не прагматичен, поскольку связан с большими объемами проводимых вычислений. Следовательно, необходим поиск компромиссных решений, позволяющих с одной стороны обеспечить требуемую точность восстановления (ошибку аппроксимации), а с другой – не превысить объем требуемого вычислительного ресурса.

В общем случае, процедура наращивания размерности функционального пространства является итерационной, и ее можно представить следующим выражением:

$$s_j(t) = \sum_j C(j,k) f_{j,k}(t), \quad j \in \mathbf{Z}. \tag{1}$$

Причем:

$$f_{0,0}(t) = 2\sum_k h_k f(2t-k),$$

где h_k представляет собой некоторую последовательность, определяющую выбор требуемой (необходимой) масштабирующей функции из семейства всей совокупности базисных функций.

Таким образом, задача кратномасштабного анализа сводится к поиску значений совокупности $\{ h_k \}_K$, где K – размерность, определяющая количество используемых коэффициентов, которые обеспечат аппроксимацию $s(t)$ с требуемой (допустимой) погрешностью.

Условиям кратномасшатбного анализа в полной мере удовлетворяют вейвлет-функции [20], которые на базе исходного (материнского) вейвлета позволяют формировать иерархическое пространство путем последовательных сдвигов и масштабирования его аргумента.

Теоретическим обоснованием такого подхода послужили двуполярные вейвлеты Хаара [10], позволившие наглядно продемонстрировать возможность реализации кратномасштабных преобразований.

Действительно, реализация прямого и обратного дискретного вейвлет-преобразования на основе вейвлета Хаара обеспечивает полное восстановление сигнала при условии, что для целых значений k существуют коэффициенты $\{ h_k \}_K$, которые удовлетворяют условию:

$$f(t/2) = \sqrt{2} \sum_k h_k f(t-k) . \qquad (2)$$

Уравнение (2) является основополагающим в теории вейвлет-преобразования [20], поскольку определяет механизм масштабирования.

Для вейвлета Хаара, решение уравнения (2) позволяет получить точные значения, характеризующие взаимосвязь между двумя соседними коэффициентами $h_{k=0}$ и h_{k+1}:

$$h_0 = h_1 = 1/\sqrt{2} .$$

Поскольку вейвлет Хаара является не единственно удовлетворяющим условию построения кратномасштабных пространств, то масштабирующее уравнение (2) может быть трансформировано к более удобному для алгоритмизации виду:

$$\frac{1}{2} f\left(x/2\right) = \sum_{n \in Z} v_n f(x-n) . \qquad (3)$$

В уравнении (3) вместо коэффициентов $\{h_k\}_K$ введены коэффициенты вида v_n, что позволяет получить более значительно удобную нормировку формируемых вейвлетов, при условии, что переменная x интерпретирует переменную времени t.

Кратномасштабный подход лег в основу алгоритмов сжатия сигналов изображений, предполагающих последовательное представление образов с различной степенью их детализации. Это позволило для описания изображения любого вида использовать аппроксимирующие коэффициенты $a_{j+1,\,k}$ и детализирующие $d_{j+1,\,k}$.

Тогда вейвлет-представление (совокупность вейвлет-коэффициентов) некоторой функции $s(t)$, с глубиной разложения J можно формализовать в терминах аппроксимирующих $a_{j+1,\,k}$ и детализирующих $d_{j+1,\,k}$ коэффициентов следующим образом:

$$\hat{s}(t) = \sum_{k \in Z} a_{j_0+J,k}\, f_{j_0+J,k}(t) + \sum_{j=j_0+1}^{J} \sum_{k \in N} a_{j,k} \psi_{j,k}(t)\,. \qquad (4)$$

В уравнении (4) обозначение $\hat{s}(t)$ указывает на то, что данная функция (сигнал) рассчитана по коэффициентам разложения исходного сигнала $s(t)$.

Прагматичность уравнения (4) определяет наличие быстрого алгоритма расчета его вейвлет-коэффициентов, определяемого в [20] как алгоритм Малла, согласно которого значения аппроксимирующих и детализирующих коэффициентов могут быть рассчитаны в соответствии с:

$$a_{j+1,k} = \sum_n h_n a_{j,n+2k}\ ,\ d_{j+1,k} = \sum_n g_n a_{j,n+2k}\,. \qquad (5)$$

В формуле (6) характер коэффициентов g_n аналогичен h_n, определяемых в выражении (2), т.е. это последовательность бинарных значений, используемая для выбора необходимых для восстановления сигналов аппроксимирующих коэффициентов.

Алгоритм восстановления сигнала по его вейвлет-коэффициентам определяется рекуррентным выражением вида:

$$a_{j,n} = \sum_k h_{n-2k} a_{j+1,k} + \sum_k g_{n-2k} d_{j+1,k}\ . \qquad (6)$$

В [10] определены условия реализации прямого и обратного вейвлет-преобразования, положенных в основу алгоритмов сжатия изображений:

1. Общее количество N аппроксимирующих коэффициентов $a_{j,k}$ должно быть кратным значению два, т.е.:

$$N / 2^m = E, \tag{7}$$

где m и E – целые числа, причем $m > 0$.

Условие (7) позволяет определить начальное значение, так называемого, «нулевого» уровня декомпозиции:

$$N = 2^{j_0}, \text{ или } j_0 = \log_2 N .$$

2. В ходе реализации прямого вейвлет-преобразования, при каждом переходе с уровня j на вышестоящий уровень $j + 1$, происходит уменьшение в два раза общего числа аппроксимирующих $a_{j,k}$ и детализирующих $d_{j,k}$ коэффициентов.

3. В ходе реализации обратного вейвлет-преобразования, т.е. при переходе на более низкий уровень декомпозиции, т.е. с уровня j на вышестоящий уровень $j – 1$, наоборот, количество аппроксимирующих $a_{j,k}$ и детализирующих $d_{j,k}$ коэффициентов возрастает в два раза.

3. Кратномасштабное вейвлет-представление изображений. Отличительной особенностью изображений является то, что они описываются посредством двумерных сигналов. Поэтому всякое изображение следует рассматривать как функцию двух переменных, т.е. $s(x, y)$. Следовательно, используемые для их обработки пространства, также должны быть двумерными, т.е. $L_2(R)$.

Так, базисные $f(x)$ и вейвлет $\psi(x)$ функции в результате преобразований порождают двумерное пространство:

$$L_2(R) : \left\{ f_{j,n}(x) \right\}, \left\{ \psi_{j,n}(x) \right\} .$$

Отметим, что тензорное произведение функций $f(x)$ и $\psi(x)$ приводит к формированию следующей комбинации базисных функций в пространстве $L_2(R^2)$:

$$\left\{ ff_{j,n,m}(x, y) = f_{j,n}(x) f_{j,m}(y) \right\};$$

$$\left\{ f\psi_{j,n,m}(x,y) = f_{j,n}(x)\psi_{j,m}(y) \right\};$$

$$\left\{ \psi f_{j,n,m}(x,y) = \psi_{j,n}(x) f_{j,m}(y) \right\};$$

$$\left\{ \psi\psi_{j,n,m}(x,y) = \psi_{j,n}(x)\psi_{j,m}(y) \right\}.$$

Для преобразованного указанным образом пространства на первом уровне декомпозиции, аппроксимирующие коэффициенты $\mathbf{A}_j = \{aa_{j,n,m}\}$ получают как коэффициенты, сформированные в результате разложения в базисе $\{ff_{j,n,m}(x,y)\}$. Горизонтальные детализирующие коэффициенты $\mathbf{H}_j = \{ad_{j,n,m}\}$ получают как коэффициенты, сформированные в результате разложения в базисе $\{f\psi_{j,n,m}(x,y)\}$. Соответственно, вертикальные детализирующие коэффициенты $\mathbf{V}_j = \{da_{j,n,m}\}$ получают как коэффициенты, сформированные в результате разложения в базисе $\{\psi f_{j,n,m}(x,y)\}$. А диагональные детализирующие коэффициенты $\mathbf{D}_j = \{dd_{j,n,m}\}$ получают как коэффициенты, сформированные в результате разложения в базисе $\{\psi\psi_{j,n,m}(x,y)\}$.

Соответственно, для второго уровня декомпозиции, в качестве исходных данных будет выступать матрица коэффициентов:

$$\mathbf{A}_1 \to (\mathbf{A}_2, \mathbf{H}_2, \mathbf{V}_2, \mathbf{D}_2).$$

Таким образом, следуя указанной аналогии, можно получить следующий алгоритм декомпозиции матрицы исходного изображения **S**.

$$\mathbf{S} \to (\mathbf{A}_1, \mathbf{H}_1, \mathbf{V}_1, \mathbf{D}_1) \to (\mathbf{A}_2, \mathbf{H}_2, \mathbf{V}_2, \mathbf{D}_2, \mathbf{H}_1, \mathbf{V}_1, \mathbf{D}_1) \to \ldots.$$

Закономерность изменения размеров формируемых двумерных массивов коэффициентов разложения заключается в следующем: на каждом уровне декомпозиции размеры получаемых массивов новых коэффициентов будут уменьшаться ровно в два раза по сравнению с предыдущими массивами. При этом общая сумма размеров всех массивов коэффициентов всегда будет равна размеру исходной матрицы **S**, что указывает на сохранение «объема» информации, и служит индикатором правильности выполнения преобразований.

В [3] указанная особенность была использована для разработки уникального алгоритма, позволившего повысить степень сжатия файлов графического изображения при заданной величине пикового отношения сигнал/шум (*PSNR*).

Такой алгоритм основан на возможности восстановления трех цветового кадра изображения $\hat{s}(l,h)$, l и h – вертикальная и горизонтальная координаты пикселя кадра, с компонентами *R, G, B,* с точностью до заданного значения *PSNR*:

$$PSNR = 10\lg\left[\frac{(2^B-1)^2 \times L \times H \times K}{\sum_{k=1}^{K}\sum_{l=1}^{L}\sum_{h=1}^{H}({}_k s(l,h) - {}_k\hat{s}(l,h))^2}\right]. \quad (8)$$

В алгоритме (8) L – горизонтальный размер кадра в пикселях; H – вертикальный размер; ${}_k s(l,h)$ – l-тое, h-тое значение пикселя k-й компоненты исходного кадра графического изображения; ${}_k\hat{s}(l,h)$ – l-тое, h-тое значение пикселя k-й компоненты декомпрессированного после сжатия восстановленного кадра; B – количество битов, отводимых на точку (в зависимости от количества представляемых цветов, на каждую точку отводится от 1 до 48 битов); $K = 3$ – число компонентов *R, G, B.*

Сущность алгоритма (8) базируется на совокупности последовательно выполняемых процедур дискретного вейвлет-преобразования, формирования матрицы вейвлет-коэффициентов, сжатия матрицы, декомпрессии, восстановления исходного кадра графического изображения путем формирования нулевой матрицы с декомпрессированными вейвлет-коэффициентами и выполнения обратного дискретного вейвлет-преобразования.

В соответствии с алгоритмом (8), чем меньше различий в числовых значениях пикселей между исходным кадром $s(l,h)$ и восстановленным $\hat{s}(l,h)$, тем выше получаемое значение показателя *PSNR*. А чем выше получаемое значение показателя *PSNR,* тем меньше претерпевает изменений файл после процедур прямого и обратного преобразований.

Представленный алгоритм сжатия посредством вейвлет-преобразования не единственный в своем роде. В настоящее время разработано несколько подходов к сжатию изображений на основе кратномасштабного преобразования их элементов, в том числе, и на

основе дискретно-косинусного преобразования [10]. Но при этом во всех алгоритмах сжатия используются одни и те же, с точки зрения технологического предназначения, технические процедуры. Их взаимосвязь показана на рисунке 1.

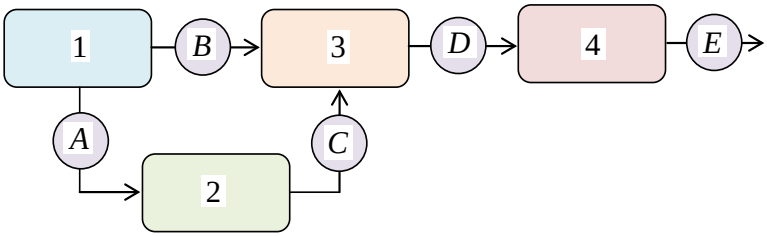


Рис. 1. Схема основных этапов преобразования изображения в алгоритмах сжатия на основе кратномаштабного преобразования

На рисунке 1 изображено:

1 – источник изображения (матрица дискретных данных);

2 – коэффициенты преобразования дискретных отсчетов изображения;

3 – массив двоичных видеоданных { $\mathbf{V}_1$ };

4 – массив двоичных видеоданных { $\mathbf{V}_2$ };

A – процедура преобразования дискретных отсчетов изображения;

B – процедура кодирования дискретных отсчетов изображения;

C – процедура кодирования коэффициентов преобразования;

D – процедура статического кодирования двоичных последовательностей;

E – цифровые коэффициенты сжатия.

Рассмотрим особенности основных процедур преобразования. Так, преобразование дискретных отсчётов изображения, блок *A*, позволяет минимизировать количество битов, используемых для кодирования результатов преобразования.

Операции кодирования преобразованных отсчетов изображения, блок *B*, и статистического кодирования результирующих двоичных последовательностей, блок *D*, обеспечивают приведение коэффициентов преобразования к двоичному виду, в интересах последующей их передачи по цифровым каналам связи.

Учитывая, что вся совокупность рассмотренных технологических операций характеризует сущность каждого из известных методов сжатия изображений, то уникальность и

возможности последних определяются не столько указанными операциями, сколько соответствием проводимых преобразований статистическим свойствам кодируемых изображений.

В качестве базовых методов сжатия изображений в настоящее время определены: адаптивное ортогональное преобразование, обеспечивающее наилучшее согласование со статистическими свойствами исходных дискретных данных; дискретно-косинусное преобразование и вейвлет-преобразование [8].

4. Обоснование вложенной структуры субматриц изображений, получаемых на основе вейвлет-коэффициентов. В рамках рассматриваемой области исследования, кратномасштабное вейвлет-преобразование исходного изображения, представленного матрицей яркостей пикселей $\mathbf{A}_{P\times N}$, где P – количество строк изображения; N – количество столбцов изображения, будет представлять собой многоуровневое иерархическое преобразование, в котором вейвлет-коэффициенты первого уровня разложения матрицы $\mathbf{A}_{P\times N}$ можно представить в матричной форме [3, 10]:

$$\mathbf{B}_1 = \begin{bmatrix} \mathbf{L}_{1\frac{P}{2}\times P} \\ \cdots \\ \mathbf{H}_{1\frac{P}{2}\times P} \end{bmatrix} \mathbf{A}_{P\times N} \left[\mathbf{L}^{\mathrm{T}}_{1\,N\times N/2} \vdots \mathbf{H}^{\mathrm{T}}_{1\,N\times N/2} \right],$$

где $\mathbf{L}_{1\,(P/2)\times P}$, $\mathbf{H}_{1\,(P/2)\times P}$, – матрицы преобразования первого уровня разложения размером $(P/2) \times P$ элементов; $\mathbf{L}^{\mathrm{T}}{}_{1\,N\times N/2}$, $\mathbf{H}^{\mathrm{T}}{}_{1\,N\times N/2}$ – матрицы преобразования первого уровня разложения размером $N\times N/2$; T – здесь и далее знак транспонирования.

Матрицы $\mathbf{L}_{1\,(P/2)\times P}$, $\mathbf{H}_{1\,(P/2)\times P}$, имеют вид:

$$\mathbf{L}_{1\frac{P}{2}\times P} = \begin{bmatrix} \mathbf{l}^{\mathrm{T}}(1) \\ \mathbf{l}^{\mathrm{T}}(2) \\ \vdots \\ \mathbf{l}^{\mathrm{T}}(P/2) \end{bmatrix},$$

$$\mathbf{H}_{1\frac{P}{2}\times P}=\begin{bmatrix}\mathbf{h}^{\mathrm{T}}(1)\\ \mathbf{h}^{\mathrm{T}}(2)\\ \vdots\\ \mathbf{h}^{\mathrm{T}}(P/2)\end{bmatrix},$$

$$\mathbf{L}^{\mathrm{T}}_{1\,N\times N/2}=\begin{bmatrix}\mathbf{l}(1) & \mathbf{l}(2) & \cdots & \mathbf{l}(N/2)\end{bmatrix},$$

$$\mathbf{H}^{\mathrm{T}}_{1\,N\times N/2}=\mathbf{h}\begin{bmatrix}\mathbf{l}(1) & \mathbf{h}(2) & \cdots & \mathbf{h}(N/2)\end{bmatrix},$$

где $\mathbf{l}^{\mathrm{T}}(p)$, $\mathbf{h}^{\mathrm{T}}(p)$ – векторы-строки, $p = 1, 2, \ldots, P/2$, состоящие из низкочастотных и высокочастотных коэффициентов, дополненных нулевыми коэффициентами, соответственно; $\mathbf{l}^{\mathrm{T}}(n)$, $\mathbf{h}^{\mathrm{T}}(n)$ – векторы-строки, $n = 1, 2, \ldots, N/2$, состоящие из низкочастотных и высокочастотных коэффициентов, дополненных нулевыми коэффициентами, соответственно.

Более наглядно, структура матриц $\mathbf{L}_{1\,(P/2)\times P}$, $\mathbf{H}_{1\,(P/2)\times P}$ представлена выражениями (9) и (10), соответственно:

$$\mathbf{L}_{1\frac{P}{2}\times P}=\begin{bmatrix}l_1 & l_2 & \cdots & l_n & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0\\ 0 & 0 & l_1 & l_2 & \cdots & l_n & 0 & 0 & 0 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ l_{n-1} & l_n & 0 & 0 & 0 & 0 & 0 & \cdots & l_1 & l_2 & \cdots & l_{n-2}\end{bmatrix}, \tag{9}$$

$$\mathbf{H}_{1\frac{P}{2}\times P}=\begin{bmatrix}h_1 & h_2 & \cdots & h_n & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0\\ 0 & 0 & h_1 & h_2 & \cdots & h_n & 0 & 0 & 0 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ h_{n-1} & h_n & 0 & 0 & 0 & 0 & 0 & \cdots & h_1 & h_2 & \cdots & h_{n-2}\end{bmatrix}. \tag{10}$$

Следует отметить, что структура матриц $\mathbf{L}^{\mathrm{T}}{}_{1\,N\times N/2}$ и $\mathbf{H}^{\mathrm{T}}{}_{1\,N\times N/2}$ является очевидной, поскольку они представляют собой результат транспонирования матриц, представленных выражениями (9) и (10), соответственно.

Таким образом, используя далее метод блочного умножения матриц [21], можно получить матрицу вейвлет-коэффициентов кратномасштабного преобразования первого уровня разложения, которую представим в виде четырех субматриц:

$$\mathbf{B}_1 = \begin{bmatrix} \mathbf{L}_1\mathbf{A}\mathbf{L}_1^{\mathrm{T}} & \mathbf{L}_1\mathbf{A}\mathbf{H}_1^{\mathrm{T}} \\ \mathbf{H}_1\mathbf{A}\mathbf{L}_1^{\mathrm{T}} & \mathbf{H}_1\mathbf{A}\mathbf{H}_1^{\mathrm{T}} \end{bmatrix}. \quad (11)$$

Так, на рисунке 2 показан пример визуализации матрицы $\mathbf{B}_1$, представляющий собой типовое изображение, рекомендованное для тестирования алгоритмов сжатия [22].

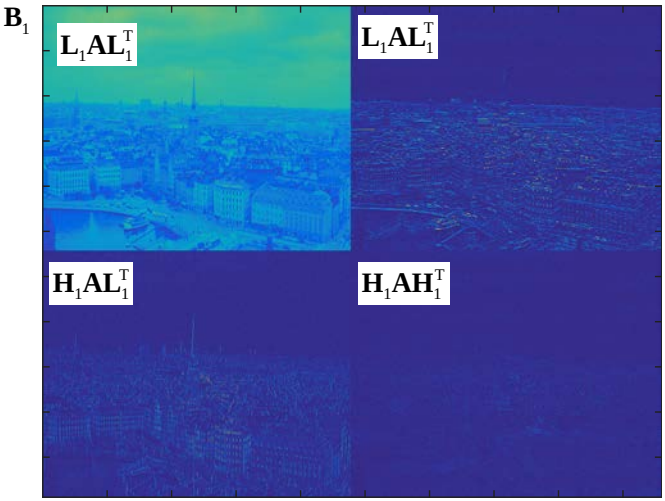


Рис. 2. Представление вейвлет-коэффициентов кратномаштабного разложения изображения на первом уровне

Матрица вейвлет-коэффициентов $\mathbf{B}_1$, представленная на рисунке 2, получена в соответствии с выражением (11).

Следует отметить, что на полученном изображении четко прослеживается 4 явно выраженных блока вейвлет-коэффициентов:

– левый верхний блок, образованный элементами субматрицы $\mathbf{L}_1\mathbf{A}\mathbf{L}^{\mathrm{T}}{}_1$;

– правый верхний блок, образованный элементами субматрицы $\mathbf{L}_1\mathbf{A}\mathbf{H}^{\mathrm{T}}{}_1$;

– левый нижний блок, образованный элементами субматрицы $\mathbf{H}_1\mathbf{A}\mathbf{L}^{\mathrm{T}}{}_1$;

– правый нижний блок, образованный элементами субматрицы $\mathbf{H}_1\mathbf{A}\mathbf{H}^{\mathrm{T}}{}_1$.

Отметим, что используемые в преобразовании (2) матрицы $\begin{bmatrix} \mathbf{L}_{1\frac{P}{2}\times P} \\ \cdots \\ \mathbf{H}_{1\frac{P}{2}\times P} \end{bmatrix}$ и $\left[\mathbf{L}^{\mathrm{T}}_{1N\times N/2} \vdots \mathbf{H}^{\mathrm{T}}_{1N\times N/2}\right]$ являются ортонормальными [21], т.е. для этих матриц справедливы следующие равенства:

$$\left[\mathbf{L}^{\mathrm{T}}_{1P\times P/2} \vdots \mathbf{H}^{\mathrm{T}}_{1P\times P/2}\right] \begin{bmatrix} \mathbf{L}_{1\frac{P}{2}\times P} \\ \cdots \\ \mathbf{H}_{1\frac{P}{2}\times P} \end{bmatrix} = \mathbf{E}_{P\times P}, \tag{12}$$

$$\left[\mathbf{L}^{\mathrm{T}}_{1N\times N/2} \vdots \mathbf{H}^{\mathrm{T}}_{1N\times N/2}\right] \begin{bmatrix} \mathbf{L}_{1\frac{N}{2}\times N} \\ \cdots \\ \mathbf{H}_{1\frac{N}{2}\times N} \end{bmatrix} = \mathbf{E}_{N\times N}, \tag{13}$$

где $\mathbf{E}_{P\times P}$ и $\mathbf{E}_{N\times N}$ – диагональные матрицы с единицами в главной диагонали размером $P\times P$ и $N\times N$ элементов, соответственно.

Отметим, что справедливость свойства ортонормальности сохраняется не только для матриц первого уровня преобразования, но и всех последующих уровней кратномасштабной декомпозиции.

Далее, используя правила блочного умножения матриц, выражения (12) и (13) можно записать в следующем виде:

$$\mathbf{L}^{\mathrm{T}}_{1P\times P/2}\,\mathbf{L}_{1(P/2)\times P} + \mathbf{H}^{\mathrm{T}}_{1P\times P/2}\,\mathbf{H}_{1(P/2)\times P} = 0,5\mathbf{E}_{P\times P} + 0,5\mathbf{E}_{P\times P} = \mathbf{E}_{P\times P}, \tag{14}$$

$$\mathbf{L}^{\mathrm{T}}_{1N\times N/2}\,\mathbf{L}_{1(N/2)\times N} + \mathbf{H}^{\mathrm{T}}_{1N\times N/2}\,\mathbf{H}_{1(N/2)\times N} = 0,5\mathbf{E}_{N\times N} + 0,5\mathbf{E}_{N\times N} = \mathbf{E}_{N\times N}. \tag{15}$$

Так, матричная форма коэффициентов второго уровня разложения будет иметь вид:

$$\mathbf{B}_2 = \begin{bmatrix} \mathbf{L}_{2\frac{P}{4}\times\frac{P}{2}} \\ \cdots \\ \mathbf{H}_{2\frac{P}{4}\times\frac{P}{2}} \end{bmatrix} \mathbf{B}_1(1{:}P/2, 1{:}N/2)\left[\mathbf{L}^{\mathrm{T}}_{1(N/2)\times N/4} \vdots \mathbf{H}^{\mathrm{T}}_{1(N/2)\times N/4}\right], \quad (16)$$

где $\mathbf{B}_1(1{:}P/2, 1{:}N/2)$ – левый верхний квадрант матрицы $\mathbf{B}_1$ размером $P/2 \times N/2$ элементов; $\mathbf{L}_{2(P/4)\times(P/2)}$, $\mathbf{H}_{2(P/4)\times(P/2)}$, $\mathbf{L}^{\mathrm{T}}{}_{1(N/2)\times(N/4)}$, $\mathbf{H}^{\mathrm{T}}{}_{1(N/2)\times(N/4)}$, – матрицы, по своей структуре, аналогичные матрицам $\mathbf{L}_{1(P/2)\times P}$, $\mathbf{H}_{1(P/2)\times P)}$, и $\mathbf{L}^{\mathrm{T}}{}_{1N\times(N/2)}$, $\mathbf{H}^{\mathrm{T}}{}_{1N\times(N/2)}$, соответственно.

Следует обратить внимание на то, что размер субматриц второго уровня кратномасштабного преобразования в два раза меньше. Это объясняется уменьшением в два раза числа строк и столбцов.

По аналогии с выражением (11) коэффициенты второго уровня разложения можно также представить в блочном виде:

$$\mathbf{B}_2 = \begin{bmatrix} \mathbf{L}_2\mathbf{B}_1(1{:}P/2, 1{:}N/2)\mathbf{L}_2^{\mathrm{T}} & \mathbf{L}_2\mathbf{B}_1(1{:}P/2, 1{:}N/2)\mathbf{H}_2^{\mathrm{T}} \\ \mathbf{H}_2\mathbf{B}_1(1{:}P/2, 1{:}N/2)\mathbf{L}_2^{\mathrm{T}} & \mathbf{H}_2\mathbf{B}_1(1{:}P/2, 1{:}N/2)\mathbf{H}_2^{\mathrm{T}} \end{bmatrix}. \quad (17)$$

Отметим, что если, в матрице $\mathbf{B}_1$ левый верхний блок, образованный элементами матрицы $\mathbf{L}_1\mathrm{A}\mathbf{L}^{\mathrm{T}}{}_1$, заменить матрицей $\mathbf{B}_2$, то в результате получим матрицу коэффициентов первого и второго уровней разложения исходного изображения, которая будет характеризоваться наличием субматриц первого и второго уровней разложения.

В общем случае, исходную матрицу $\mathbf{B}_1$ можно рассматривать как ядро для вейвлет-коэффициентов второго уровня кратномаштабного преобразования анализируемого изображения.

Пример визуального представления элементов матрицы вейвлет-коэффициентов второго уровня разложения $\mathbf{B}_2$ демонстрируется на рисунке 3.

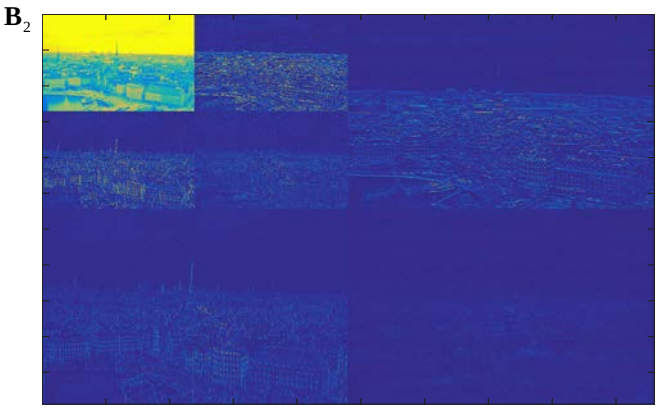


Рис. 3. Представление вейвлет-коэффициентов первого и второго уровней кратномасштабного разложения

С учетом выражений (9 – 17) вейвлет коэффициенты матрицы $\mathbf{A}_{P\times N}$ J-го уровня разложения в матричной форме можно записать в следующем виде:

$$\mathbf{B}_J = \begin{bmatrix} \mathbf{L}_{J\,\frac{P}{2^J}\times\frac{P}{2^{J-1}}} \\ \cdots \\ \mathbf{H}_{J\,\frac{P}{2^J}\times\frac{P}{2^{J-1}}} \end{bmatrix} \mathbf{B}_{J-1}(1{:}P/2^{J-1}, 1{:}N/2^{J-1})\times \\ \times\left[\mathbf{L}^{\mathrm{T}}_{J\,(N/2^{J-1})\times N/2^J} \vdots \mathbf{H}^{\mathrm{T}}_{J\,(N/2^{J-1})\times N/2^J}\right] . \quad (18)$$

Таким образом, выражение (18) можно рассматривать как обобщенную аналитическую форму алгоритма кратномасшабной декомпозиции изображений.

В качестве иллюстрации на рисунке 4 показаны вейвлет-коэффициенты на шести уровнях кратномасштабного преобразования тестируемого изображения.

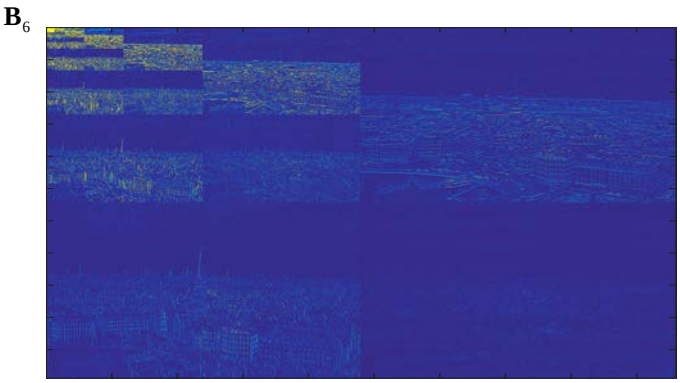


Рис. 4. Представление вейвлет-коэффициентов с первого по шестой уровни кратномасштабного разложения

На основе приведенных математический рассуждений можно сделать вывод о том, что коэффициенты вейвлет-преобразования имеют структуру вложенных субматриц. При этом каждый уровень разложения характеризуется своими субматрицами. В связи с этим корреляционный анализ целесообразно проводить раздельно для элементов субматриц соответствующих уровней разложения. Кроме того, поскольку ядром i-го уровня является матрица коэффициентов предшествующего, $i-1$-го, уровня, то можно сделать обоснованное предположение о зависимости соответствующих элементов соседних уровней. Поэтому далее проанализируем также зависимости между соответствующими коэффициентами различных уровней преобразования.

5. Анализ корреляционных зависимостей вейвлет-коэффициентов различных уровней кратномасштабного преобразования. В интересах исследования корреляционных зависимостей между вейвлет-коэффициентами, получаемых на различных уровнях кратномасштабного преобразования, определим прямоугольную область размером $p \times n$ элементов.

Затем последовательно будем смещать исследуемую область внутри анализируемой субматрицы от крайнего верхнего положения до крайнего нижнего. Общая последовательность указанных процедур представлена на рисунке 5.

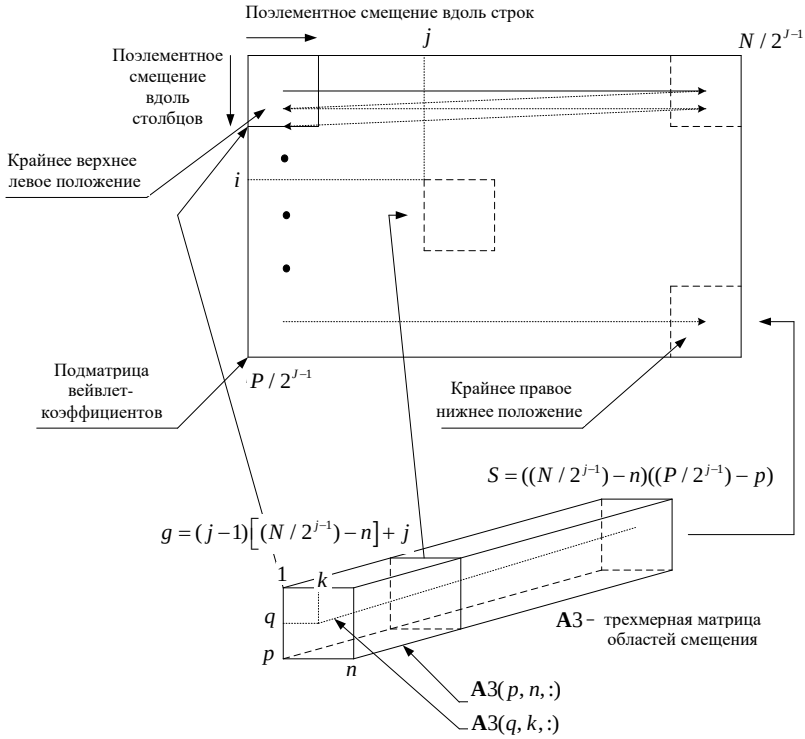


Рис. 5. Последовательность и направление процедур смещения исследуемой области вейвлет-коэффициентов внутри анализируемой субматрицы

Поскольку субматрица вейвлет-коэффициентов J-го уровня кратномасштабного разложения имеет $P/2^{J-1}$ строк и $N/2^{J-1}$ столбцов, то тогда общее количество поэлементных смещений в исследуемой области размером $p \times n$ элементов, непосредственно внутри субматрицы размером $P/2^{J-1} \times N/2^{J-1}$, составит:

$$S = \left(\frac{P}{2^{J-1}} - p\right)\left(\frac{N}{2^{J-1}} - n\right). \qquad (19)$$

С учетом полученных результатов, каждую матрицу смещения предлагается рассматривать как соответствующий слой трехмерной матрицы $\mathbf{A}3_{p \times n \times S}$. Принцип ее формирования представлен на рисунке 5.

Так, например, первый слой трехмерной матрицы образован из матрицы смещения, занимающей крайнее левое верхнее положение. В свою очередь, последний, *S*-ный слой матрицы $\mathbf{A}З_{p\times n\times S}$ образован из матрицы смещения, занимающей крайнее правое нижнее положение.

Таким образом, матрица смещения, левый верхний элемент которой, соответствует элементу субматрицы вейвлет-коэффициентов кратномасштабного преобразования с индексами *i*, *j*, образует *q*-ый слой трехмерной матрицы $\mathbf{A}З_{p\times n\times S}$. Причем процедуру формирования *q*-ного слоя трехмерной матрицы $\mathbf{A}З_{p\times n\times S}$ можно формализовать следующим образом:

$$q = (i-1)\left(\frac{N}{2^{J-1}} - n\right) + j\,. \tag{20}$$

В интересах оценки наличия взаимосвязей между элементами матриц смещений, имеющих одинаковую индексацию, будем исходить из того, что данные элементы, в свою очередь, образуют соответствующие вектора трехмерной матрицы $\mathbf{A}З_{p\times n\times S}$.

Например, элементы матриц смещений с индексами (*q*, *k*) образуют трехмерный вектор **А**З(*q*, *k*, :), как это показано на рисунке 4.

С целью выявления наличия взаимосвязей между элементами матриц смещения вычислим корреляцию между векторами **А**З(*q*, *k*, :), $q = 1, 2, \ldots, p$; $k = 1, 2, \ldots, n$ и вектором **А**З(*p*, *n*, :). Для этого воспользуемся выражением выборочного парного коэффициента корреляции К. Пирсона [23]:

$$r_{q,k} = \frac{\sum_{i}^{S}\left(\mathbf{A}З(q,k,i) - \frac{1}{S}\sum_{i=1}^{S}\mathbf{A}З(q,k,i)\right)}{\sqrt{\sum_{i=1}^{S}\left(\mathbf{A}З(q,k,i) - \frac{1}{S}\sum_{i=1}^{S}\mathbf{A}З(q,k,i)\right)^2}}\times \\ \times\frac{\left(\mathbf{A}З(p,k,i) - \frac{1}{S}\sum_{i=1}^{S}\mathbf{A}З(p,k,i)\right)}{\sqrt{\sum_{i=1}^{S}\left(\mathbf{A}З(p,k,i) - \frac{1}{S}\sum_{i=1}^{S}\mathbf{A}З(p,k,i)\right)^2}}\,. \tag{21}$$

В выражении (21) где $q = 1, 2, \ldots, p$; $k = 1, 2, \ldots, n$.

Отметим, что величины, рассчитанные в соответствии с выражением (21), образуют матрицу парных корреляций между

элементами с индексами (q, k) и элементами с индексами (p, n). При этом величины $r_{q,k}$, где $q = 1, 2, \ldots, p$; $k = 1, 2, \ldots, n$ изменяются в пределах:

$$-1 \le r_{q,\,k} \le 1. \qquad (22)$$

Физический смысл коэффициентов $r_{q,k}$, удовлетворяющий условию (22) состоит в том, что если значение $r_{q,k}$ по величине близко к 1, т.е. $r_{q,k} \approx 1$, то между элементами множеств: $\mathbf{A}3(q, k, i)$ и $\mathbf{A}3(p, n, i)$, $i = 1, 2, \ldots, S$ однозначно существует зависимость вида:

$$\mathbf{A}3(q, k, i) \approx a\mathbf{A}3(p, n, i) + b, \qquad (23)$$

где a, b – вещественные числа, причем $a > 0$.

А если выполняется приближенное равенство (22), то тогда между элементами множеств $\mathbf{A}3(q, k, i)$ и $\mathbf{A}3(p, n, i)$, $i = 1, 2, \ldots, S$ существует линейная зависимость. Причем возрастание элементов множества $\mathbf{A}3(q, k, i)$ приводит к возрастанию соответствующих элементов множества $\mathbf{A}3(p, n)$.

В свою очередь, если величина $r_{q,k}$ по своему значению близка к –1, т.е. $r_{q,k} \approx -1$, то тогда можно утверждать, что между элементами множеств $\mathbf{A}3(q, k, i)$ и $\mathbf{A}3(p, n, i)$, $i = 1, 2, \ldots, S$ существует зависимость в виде (23), но при этом $a < 0$. Это означает, что возрастание элементов множества $\mathbf{A}3(q, k, i)$ приводит к уменьшению соответствующих элементов множества $\mathbf{A}3(p, n, i)$. Промежуточные значения $r_{q,k}$, близкие к 0, указывают на слабую корреляцию между элементами и, соответственно, низкую зависимость.

6. Результаты практического эксперимента. На рисунках 6, 7 и 8 показаны результаты анализа зависимостей между вейвлет-коэффициентами кратномасштабного преобразования для $J = 1$ для $\mathbf{L}_1\mathbf{A}\mathbf{H}_1^{\mathrm{T}}$, $\mathbf{H}_1\mathbf{A}\mathbf{L}_1^{\mathrm{T}}$ и $\mathbf{H}_1\mathbf{A}\mathbf{H}_1^{\mathrm{T}}$, полученных в ходе практического эксперимента. Так на рисунках 6а, 7а, 8а показаны коэффициенты корреляции, вычисленные в соответствии с выражением (21) для субматриц первого уровня разложения.

В качестве локальной области смещения использовалась область размером 8 × 8 элементов. Расчеты проведены для исходного тестового изображения, представленного на рисунке 2.

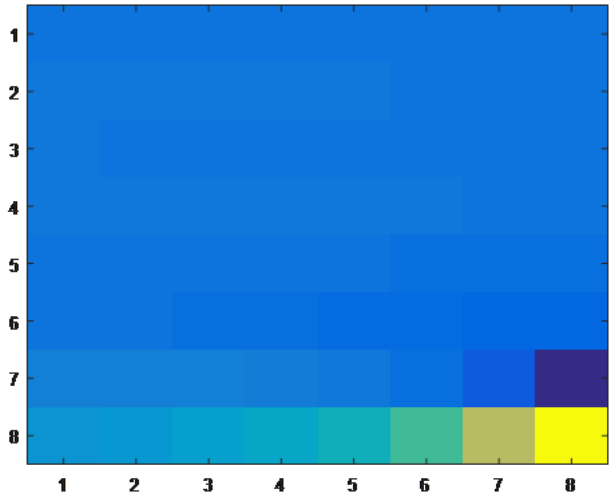


Рис. 6(а)

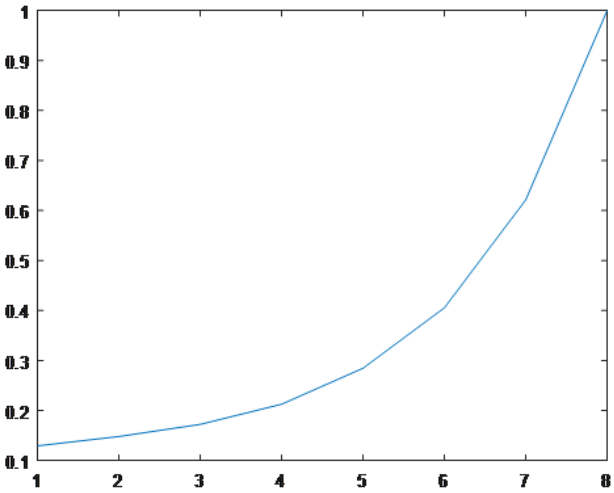


Рис. 6(б)

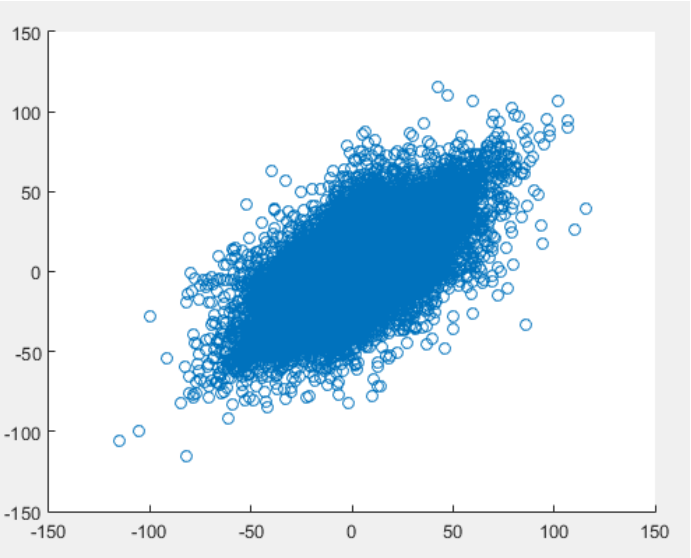


Рис. 6(в)

Рис. 6. Результаты анализа зависимостей между вейвлет-коэффициентами кратномасштабного преобразования $\mathbf{L}_1\mathbf{A}\mathbf{H}^{\mathrm{T}}{}_1$ для $J = 1$

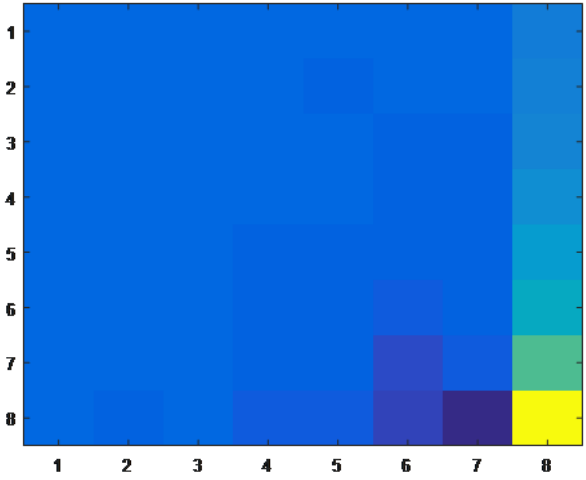


Рис. 7(а)

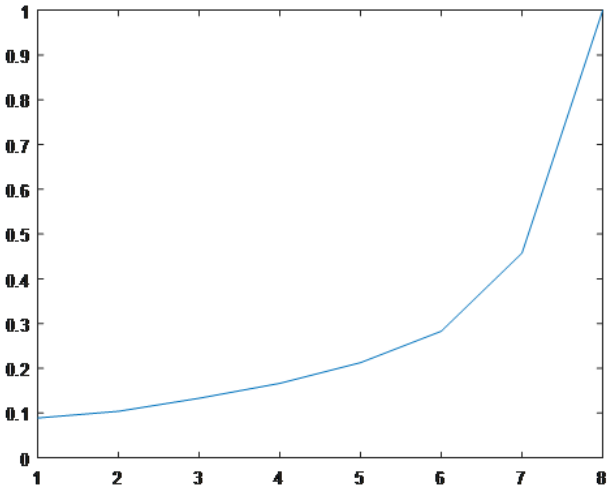


Рис. 7(б)

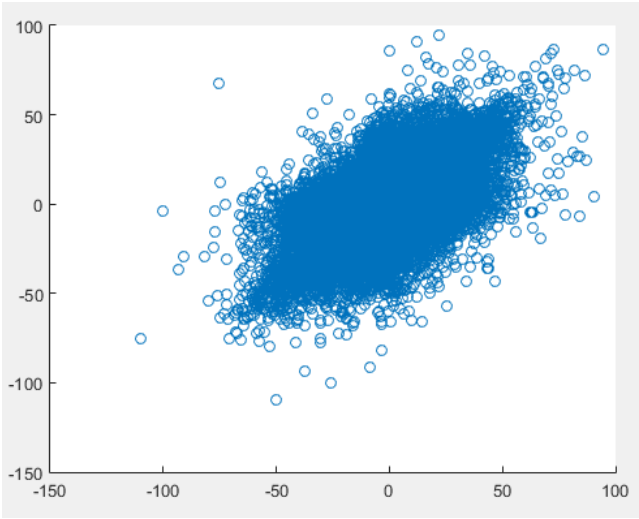


Рис. 7(в)

Рис. 7. Результаты анализа зависимостей между вейвлет-коэффициентами кратномасштабного преобразования $\mathbf{H}_1\mathbf{A}\mathbf{L}^{\mathrm{T}}{}_1$ для $J = 1$

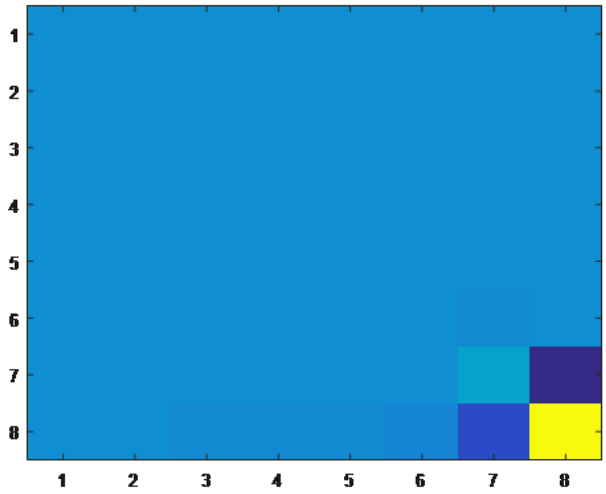


Рис. 8(а)

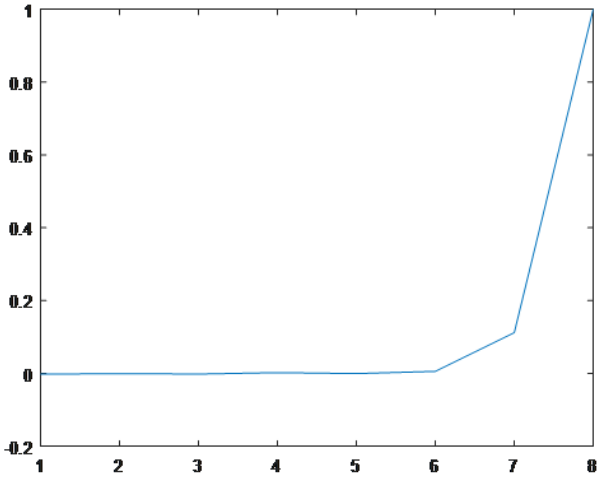


Рис. 8(б)

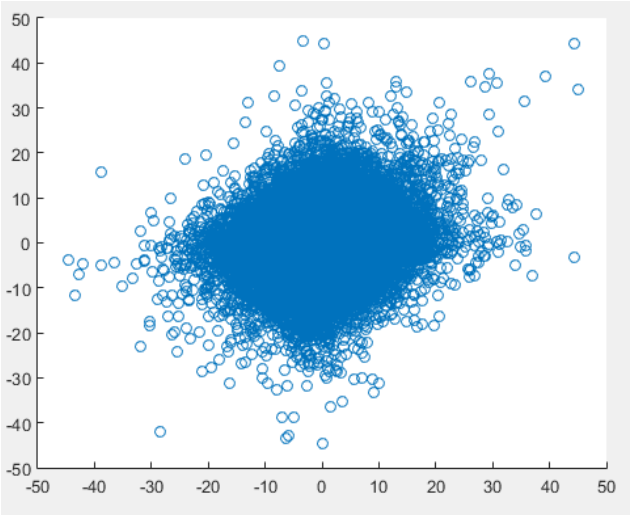


Рис. 8(в)

Рис. 8. Результаты анализа зависимостей между вейвлет-коэффициентами кратномасштабного преобразования $\mathbf{H}_1\mathbf{A}\mathbf{H}^{\mathrm{T}}{}_1$ для $J = 1$

Для лучшей наглядности сущности полученных результатов на рисунках 6(а), 7(а) и 8(а), на рисунке 9 представлена цветовая шкала, с помощью которой можно оценить степень корреляции. Дополнительно, над цветовой шкалой представлены и количественные значения.

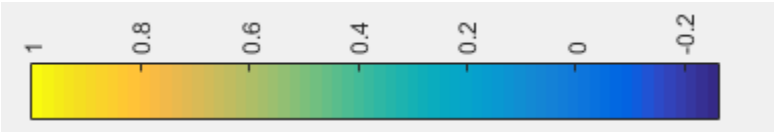


Рис. 9. Цветовая шкала соответствия значений коэффициентов корреляции в матрицах разложения

На рисунках 6(б), 7(б) и 8(б) изображены графики зависимости коэффициентов корреляции: в восьмой строке для субматрицы $\mathbf{L}_1\mathbf{A}\mathbf{H}^{\mathrm{T}}{}_1$; в восьмом столбце для субматрицы $\mathbf{H}_1\mathbf{A}\mathbf{L}^{\mathrm{T}}{}_1$; для диагональных элементов субматрицы $\mathbf{H}_1\mathbf{A}\mathbf{H}^{\mathrm{T}}{}_1$.

Наконец, на рисунках 6(в), 7(в) и 8(в) показаны графики рассеяния для вейвлет-коэффициентов с максимальной корреляцией.

Анализируя данные, полученные в ходе экспериментального исследования, было определено, что между вейвлет-коэффициентами внутри локальной области размером 8×8 существует однозначная линейная зависимость.

При этом установлено следующее:

– для субматрицы $\mathbf{L}_1\mathbf{A}\mathbf{H}^{\mathrm{T}}{}_1$ характерно наличие максимальной корреляции между соседними элементами, находящимися в строке;

– для субматрицы $\mathbf{H}_1\mathbf{A}\mathbf{L}^{\mathrm{T}}{}_1$ характерно наличие максимальной корреляции между соседними элементами, находящимися в столбце;

– для субматрицы $\mathbf{H}_1\mathbf{A}\mathbf{H}^{\mathrm{T}}{}_1$ характерно наличие максимальной корреляции между соседними диагональными элементами.

– корреляция между элементами субматрицы $\mathbf{H}_1\mathbf{A}\mathbf{H}^{\mathrm{T}}{}_1$ существенно ниже, чем между элементами субматриц $\mathbf{L}_1\mathbf{A}\mathbf{H}^{\mathrm{T}}{}_1$ и $\mathbf{H}_1\mathbf{A}\mathbf{L}^{\mathrm{T}}{}_1$, о чем свидетельствуют, как низкое значение максимального коэффициента корреляции (0,1 по сравнению с 0,6 и 0,5), так и более равномерный характер рассеяния.

Полученные закономерности сохраняются и для субматриц остальных уровней разложения. Для устойчивого подтверждения данного факта, описанные вычисления проводились на выборке из более двухсот типовых изображений.

Следует отметить, что на практике больший интерес представляет подход, основанный на реализации алгоритма (8). Это обусловлено рядом причин. Прежде всего – возможностью работы с различными форматами представления данных изображения. Если для исходного состояния изображения использовать графический файл формата *BMP* (от англ. *Bitmap Picture* − формат хранения растровых изображений, разработанный компанией *Microsoft*), то после сжатия, алгоритм позволяет перейти к более компактному формату сжатия – *JPEG* (от англ. *Joint Photographic Experts Group*, по названию организации-разработчика). Так, для первой итерации выбирают значение переменной $iter = 0$. Затем над исходным изображением $A(l,h)$, имеющим размер $\frac{L}{2^{iter}} \times \frac{H}{2^{iter}}$ (первая итерация $iter = 0$: $L \times H$) элементов производят вычисление матрицы вейвлет-преобразования размером $\frac{L}{2^{iter}} \times \frac{H}{2^{iter}}$. И уже из этой матрицы формируют матрицу вейвлет-коэффициентов $Y(l,h)$, где $l = 1,...,\frac{L}{2^{iter+1}}$ и $h = 1,...,\frac{H}{2^{iter+1}}$ (для $iter = 0$: $Y_1(l,h)$, размером $(L/2) \times (H/2)$). Следует отметить, что

значения вейвлет-коэффициентов в матрице на позициях $l = 1,..., \frac{L}{2^{iter+1}}$ и $h = 1,..., \frac{H}{2^{iter+1}}$ по своей величине равны нулю, или очень близки к этому значению, т.е. обладают очень низкой энергией.

Далее, полученные вейвлет-коэффициенты $Y(l,h)$ сжимают алгоритмом *JPEG*, после чего декомпрессируют и формируют нулевую матрицу $O(l,h)$ размером $\frac{L}{2^{iter}} \times \frac{H}{2^{iter}}$ элементов и вместо элементов с индексами $l = 1,..., \frac{L}{2^{iter+1}}$ и $h = 1,..., \frac{H}{2^{iter+1}}$ помещают соответствующие элементы декомпрессированной матрицы $\hat{Y}(l,h)$.

В результате, сформированная нулевая матрица $O_1(l,h)$ представляет собой матрицу размером $L \times H$, значения элементов которой равны нолю.

После этого, элементы, находящиеся на позициях $l = 1,..., \frac{L}{2^{iter+1}}$ и $h = 1,..., \frac{H}{2^{iter+1}}$, заменяют вейвлет-коэффициентами декомпрессированной матрицы $\hat{Y}_1(l,h)$.

На следующем этапе формируют восстановленное изображение $\hat{A}(l,h)$ посредством реализации процедур обратного вейвлет-преобразования над нулевой матрицей $O_1(l,h)$ размером $\frac{L}{2^{iter}} \times \frac{H}{2^{iter}}$, в позициях индексов $l = 1,..., \frac{L}{2^{iter+1}}$ и $h = 1,..., \frac{H}{2^{iter+1}}$ которой размещены вейвлет-коэффициенты декомпрессированной матрицы $\hat{Y}_1(l,h)$.

И на последнем этапе определяют величину *PSNR*, характеризующую качество восстановленного изображения $\hat{Y}_1(l,h)$ по отношению к исходному $A(l,h)$, и сравнивают ее заданной величиной $PSNR_{\text{доп}}$. А величину $PSNR_{iter}$ (для $iter = 0$: $PSNR_0$) оценивают по формуле (8), сравнивая ее с предварительно заданной величиной $PSNR_{\text{доп}}$. Выполнение условия $PSNR_0 > PSNR_{\text{доп}}$ указывает на то, что значение переменной $iter$ следует увеличить на единицу и повторить рассмотренные выше процедуры, при этом помещая каждый раз вместо $A(l,h)$ значения вейвлет-коэффициентов $Y(l,h)$,

сформированных на предыдущем этапе выполнения итераций. Указанные процедуры следует выполнять до тех пор, пока не будет достигнуто требование $PSNR_{iter} \le PSNR_{\text{доп}}$, указывающее на то, что сжатый файл изображения может быть использован для хранения.

7. Заключение. Представленные результаты теоретических исследований и данные проведенного практического эксперимента позволяют утверждать, что между низкочастотными коэффициентами кратномасштабного вейвлет-преобразования верхнего уровня разложения сохраняются примерно одинаковые зависимости равномерно по всем направлениям.

Таким образом, можно полагать о правомерности сделанных заключений о наличии характерных зависимостей между коэффициентами кратномасштабного вейвлет-преобразования различных уровней разложения матриц изображений.

Следовательно, учет указанных зависимостей между коэффициентами в процессе их кодирования позволит получить большие значения коэффициентов сжатия.

Авторы полагают, что совместная реализация полученных результатов с алгоритмом сжатия, основанном на учете заданной величины пикового отношения сигнал/шум, определяемого последовательностью выполняемых процедур кратномасштабного вейвлет-преобразования, формирования матрицы вейвлет-коэффициентов, ее сжатия, декомпрессии и восстановления исходного графического изображения путем формирования нулевой матрицы с декомпрессированными вейвлет-коэффициентами, и выполнением обратного дискретного вейвлет-преобразования, позволит существенно расширить арсенал средств в практике сжатия изображений.

Математическая модель адаптивного арифметического кодирования, учитывающая корреляционные свойства коэффициентов кратномасштабного вейвлет-преобразования, будет подробно рассмотрена в следующей статье по данной тематике.

Литература

1. Shevchuk B., Brayko Y., Geraimchuk M., Ivakhiv O. Highly information and energy-efficient monitoring data transmission in iot networks. Journal of Mobile Multimedia. 2021. Т. 17. № 4. С. 465-498.
2. Mizdos T., Uhrina M., Pocta P., Barkowsky M. How to reuse existing annotated image quality datasets to enlarge available training data with new distortion types. Multimedia Tools and Applications. 2021. Т. 80. № 18. С. 28137-28159.
3. Умбиталиев А.А., Дворников С.В., Оков И.Н., Устинов А.А. Способ сжатия графических файлов методами вейвлет-преобразований // Вопросы радиоэлектроники. Серия: Техника телевидения. 2015. № 3. С. 100-106.

4. Дворников С.В., Устинов А.А., Оков И.Н., Царелунго А.Б., Дворовой М.О., Цветков В.В. Способ сжатия графических файлов // Вопросы радиоэлектроники. Серия: Техника телевидения. 2017. № 4. С. 77-86.
5. Kamatar V.S., Baligar V.P., Savanur S.S. Wo phase image compression algorithm using diagonal pixels of image blocks. В сборнике: 2021 2nd International Conference for Emerging Technology, INCET 2021. 2. 2021. С. 9456290.
6. Tellez D., Litjens G., Van Der Laak J., Ciompi F. Neural image compression for gigapixel histopathology image analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021. Т. 43. № 2. С. 567-578.
7. Sokolova E.A., Nyrkov A.P., Ivanovskii A.N. 3D image compression with variable fragments. В сборнике: Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2021. 2021. С. 699-702.
8. Дворников С.В., Устинов А.А., Цветков В.В. Компенсация движения в видеокодеках, использующих трёхмерные ортогональные преобразования, на основе оптимальных разбиений кодируемых блоков во временной области // Вопросы радиоэлектроники. Серия: Техника телевидения. 2013. № 2. С. 98-111.
9. Messaoudi A., Benchabane F., Srairi K. DCT-based color image compression algorithm using adaptive block scanning. Signal, Image and Video Processing. 2019. Т. 13. № 7. С. 1441-1449.
10. Mander K., Jindal H. An improved image compression-decompression technique using block truncation and wavelets. International Journal of Image, Graphics and Signal Processing. 2017. Т. 9. № 8. С. 17-29.
11. Kumar G., Brar Er.S.S., Kumar R., Kumar A. Review: dwt-dct technique and arithmetic-huffman coding based image compression. International Journal of Engineering and Manufacturing. 2015. Т. 5. № 3. С. 20-33.
12. Pertsau D.Yu., Doudkin A.A. Context modeling in problems of compressing hyperspectral remote sensing data Pattern Recognition and Image Analysis. 2020. Т. 30. № 2. С. 217-223.
13. Hua S., Zhao W., Liu J. Background suppression algorithms based on improved filter and image multi-scale transformation. Xi Tong Gong Cheng Yu Dian Zi Ji Shu. 2020. Т. 42. № 8. С. 1679-1684.
14. Li G., Lin Y., Qu X. An infrared and visible image fusion method based on multi-scale transformation and norm optimization. Information Fusion. 2021. Т. 71. С. 109-129.
15. Zhou J., Yao J., Zhang W., Zhang D. Multi-scale retinex-based adaptive gray-scale transformation method for underwater image enhancement. Multimedia Tools and Applications. 2021.
16. Yu J., Zhang B., Chen W., Liu H. Multi-scale analysis on the tensile properties of UHPC considering fiber orientation. Composite Structures. 2022. Т. 280. С. 114835.
17. Дворников С.В., Дворников С.С., Спирин А.М. Синтез манипулированных сигналов на основе вейвлет-функций // Информационные технологии. 2013. № 12. С. 52-55.
18. S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, 1999. EDN: YGCTMD.
19. Bozhokin S., Suslova I., Tarakanov D. Special techniques in applying continuous wavelet transform to non-stationary signals of heart rate variability. Communications in Computer and Information Science (см. в книгах). 2020. Т. 1211 CCIS. С. 291-310.
20. Астафьева Н.М. Вейвлет-анализ: основы теории и примеры применения // Успехи физических наук. – 1998.– Т. 166 – № 11 – С. 1145–1170

21. Yu Q., Avestimehr A.S. Coded computing for resilient, secure, and privacy-preserving distributed matrix multiplication. IEEE Transactions on Communications. 2021. Т. 69. № 1. С. 59-72.
22. Ahmed I.T., Hammad B.T., Der C.S., Jamil N. Contrast-distorted image quality assessment based on curvelet domain features. International Journal of Electrical and Computer Engineering. 2021. Т. 11. № 3. С. 2595-2603.
23. Saccenti, E., Hendriks, M.H.W.B. & Smilde, A.K. Corruption of the Pearson correlation coefficient by measurement error and its estimation, bias, and correction under different error models. Sci Rep 10, 438 (2020). https://doi.org/10.1038/s41598-019-57247-4.

Дворников Сергей Викторович — д-р техн. наук, профессор, кафедра радиотехнических и оптоэлектронных комплексов, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения»; научный сотрудник, Военная академия связи им. С.М. Буденного. Область научных интересов: построение помехозащищенных систем радиосвязи, способов формирования и обработки сигналов сложных структур. Число научных публикаций — 283. practicdsv@yandex.ru; Тихорецкий проспект, 3, 194064, Санкт-Петербург, Россия; р.т.: +7(812)247-9811.

Дворников Сергей Сергеевич — канд. техн. наук, старший преподаватель, кафедра конструирования и технологий электронных и лазерных средств, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения»; научный сотрудник, научно-исследовательский отдел, Военная академия связи им. С.М. Буденного. Область научных интересов: теория передачи сигналов, спектральная эффективность сигналов, помехозащищенность каналов управления и связи радиотехнических систем. Число научных публикаций — 114. dvornik.92@mail.ru; Тихорецкий проспект, 3, 194064, Санкт-Петербург, Россия; р.т.: +7(812)247-9820.

Устинов Андрей Александрович — д-р техн. наук, профессор, ведущий научный сотрудник, ФГУП «ГосНИИПП». Область научных интересов: теория связи, помехозащищенность инфокоммуникационных каналов радиотехнических систем, сжатие видео и речевой информации. Число научных публикаций — 150. ust_m_a@mail.ru; набережная Обводного канала, 29, 191167, Санкт-Петербург, Россия; р.т.: +7(812)274-3156.

DOI 10.15622/ia.21.5.6

S. DVORNIKOV, S. DVORNIKOV, A. USTINOV

ANALYSIS OF THE CORRELATION PROPERTIES OF THE WAVELET TRANSFORM COEFFICIENTS OF TYPICAL IMAGES

Dvornikov S., Dvornikov S., Ustinov A. **Analysis of the Correlation Properties of the Wavelet Transform Coefficients of Typical Images.**

Abstract. The increasing flow of photo and video information transmitted through the channels of infocommunication systems and complexes stimulates the search for effective compression algorithms that can significantly reduce the volume of transmitted traffic, while maintaining its quality. In the general case, the compression algorithms are based on the operations of converting the correlated brightness values of the pixels of the image matrix into their uncorrelated parameters, followed by encoding the obtained conversion coefficients. Since the main known decorrelating transformations are quasi-optimal, the task of finding transformations that take into account changes in the statistical characteristics of compressed video data is still relevant. These circumstances determined the direction of the study, related to the analysis of the decorrelating properties of the generated wavelet coefficients obtained as a result of multi-scale image transformation. The main result of the study was to establish the fact that the wavelet coefficients of the multi-scale transformation have the structure of nested matrices defined as submatrices. Therefore, it is advisable to carry out the correlation analysis of the wavelet transformation coefficients separately for the elements of each submatrix at each level of decomposition (decomposition). The main theoretical result is the proof that the core of each subsequent level of the multi-scale transformation is a matrix consisting of the wavelet coefficients of the previous level of decomposition. It is this fact that makes it possible to draw a conclusion about the dependence of the corresponding elements of neighboring levels. In addition, it has been found that there is a linear relationship between the wavelet coefficients within the local area of the image with a size of 8×8 pixels. In this case, the maximum correlation of submatrix elements is directly determined by the form of their representation, and is observed between neighboring elements located, respectively, in a row, column or diagonally, which is confirmed by the nature of the scattering. The obtained results were confirmed by the analysis of samples from more than two hundred typical images. At the same time, it is substantiated that between the low-frequency wavelet coefficients of the multi-scale transformation of the upper level of the expansion, approximately the same dependences are preserved uniformly in all directions. The practical significance of the study is determined by the fact that all the results obtained in the course of its implementation confirm the presence of characteristic dependencies between the wavelet transform coefficients at different levels of image decomposition. This fact indicates the possibility of achieving higher compression ratios of video data in the course of their encoding. The authors associate further research with the development of a mathematical model for adaptive arithmetic coding of video data and images, which takes into account the correlation properties of wavelet coefficients of a multi-scale transformation.

Keywords: image compression, wavelet multiscale transform coefficients, correlated value transformation, decorrelation transformations.

References

1. Shevchuk B., Brayko Y., Geraimchuk M., Ivakhiv O. Highly information and energy-efficient monitoring data transmission in iot networks. Journal of Mobile Multimedia. 2021. T. 17. № 4. C. 465-498.

2. Mizdos T., Uhrina M., Pocta P., Barkowsky M. How to reuse existing annotated image quality datasets to enlarge available training data with new distortion types. Multimedia Tools and Applications. 2021. Т. 80. № 18. С. 28137-28159.
3. Умбиталиев А.А., Дворников С.В., Оков И.Н., Устинов А.А. Способ сжатия графических файлов методами вейвлет-преобразований // Вопросы радиоэлектроники. Серия: Техника телевидения. 2015. № 3. С. 100-106.
4. Дворников С.В., Устинов А.А., Оков И.Н., Царелунго А.Б., Дворовой М.О., Цветков В.В. Способ сжатия графических файлов // Вопросы радиоэлектроники. Серия: Техника телевидения. 2017. № 4. С. 77-86.
5. Kamatar V.S., Baligar V.P., Savanur S.S. Wo phase image compression algorithm using diagonal pixels of image blocks. В сборнике: 2021 2nd International Conference for Emerging Technology, INCET 2021. 2. 2021. С. 9456290.
6. Tellez D., Litjens G., Van Der Laak J., Ciompi F. Neural image compression for gigapixel histopathology image analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021. Т. 43. № 2. С. 567-578.
7. Sokolova E.A., Nyrkov A.P., Ivanovskii A.N. 3D image compression with variable fragments. В сборнике: Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2021. 2021. С. 699-702.
8. Дворников С.В., Устинов А.А., Цветков В.В. Компенсация движения в видеокодеках, использующих трёхмерные ортогональные преобразования, на основе оптимальных разбиений кодируемых блоков во временной области // Вопросы радиоэлектроники. Серия: Техника телевидения. 2013. № 2. С. 98-111.
9. Messaoudi A., Benchabane F., Srairi K. DCT-based color image compression algorithm using adaptive block scanning. Signal, Image and Video Processing. 2019. Т. 13. № 7. С. 1441-1449.
10. Mander K., Jindal H. An improved image compression-decompression technique using block truncation and wavelets. International Journal of Image, Graphics and Signal Processing. 2017. Т. 9. № 8. С. 17-29.
11. Kumar G., Brar Er.S.S., Kumar R., Kumar A. Review: dwt-dct technique and arithmetic-huffman coding based image compression. International Journal of Engineering and Manufacturing. 2015. Т. 5. № 3. С. 20-33.
12. Pertsau D.Yu., Doudkin A.A. Context modeling in problems of compressing hyperspectral remote sensing data Pattern Recognition and Image Analysis. 2020. Т. 30. № 2. С. 217-223.
13. Hua S., Zhao W., Liu J. Background suppression algorithms based on improved filter and image multi-scale transformation. Xi Tong Gong Cheng Yu Dian Zi Ji Shu. 2020. Т. 42. № 8. С. 1679-1684.
14. Li G., Lin Y., Qu X. An infrared and visible image fusion method based on multi-scale transformation and norm optimization. Information Fusion. 2021. Т. 71. С. 109-129.
15. Zhou J., Yao J., Zhang W., Zhang D. Multi-scale retinex-based adaptive gray-scale transformation method for underwater image enhancement. Multimedia Tools and Applications. 2021.
16. Yu J., Zhang B., Chen W., Liu H. Multi-scale analysis on the tensile properties of UHPC considering fiber orientation. Composite Structures. 2022. Т. 280. С. 114835.
17. Дворников С.В., Дворников С.С., Спирин А.М. Синтез манипулированных сигналов на основе вейвлет-функций // Информационные технологии. 2013. № 12. С. 52-55.
18. S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, 1999. EDN: YGCTMD.

19. Bozhokin S., Suslova I., Tarakanov D. Special techniques in applying continuous wavelet transform to non-stationary signals of heart rate variability. Communications in Computer and Information Science (см. в книгах). 2020. Т. 1211 CCIS. С. 291-310.
20. Астафьева Н.М. Вейвлет-анализ: основы теории и примеры применения // Успехи физических наук. – 1998.– Т. 166 – № 11 – С. 1145–1170
21. Yu Q., Avestimehr A.S. Coded computing for resilient, secure, and privacy-preserving distributed matrix multiplication. IEEE Transactions on Communications. 2021. Т. 69. № 1. С. 59-72.
22. Ahmed I.T., Hammad B.T., Der C.S., Jamil N. Contrast-distorted image quality assessment based on curvelet domain features. International Journal of Electrical and Computer Engineering. 2021. Т. 11. № 3. С. 2595-2603.
23. Saccenti, E., Hendriks, M.H.W.B. & Smilde, A.K. Corruption of the Pearson correlation coefficient by measurement error and its estimation, bias, and correction under different error models. Sci Rep 10, 438 (2020). https://doi.org/10.1038/s41598-019-57247-4.

Dvornikov Sergey — Ph.D., Dr.Sci., Professor, Department of radio engineering and optoelectronic complexes, Federal State Autonomous Educational Institution of Higher Education "St. Petersburg State University of Aerospace Instrumentation"; Researcher, Military Academy of Communication. Research interests: construction of noise-immune radio communication systems, methods of generating and processing signals of complex structures. The number of publications — 283. practicdsv@yandex.ru; 3, Tikhoretsky Av., 194064, St. Petersburg, Russia; office phone: +7(812)247-9811.

Dvornikov Sergey — Ph.D., Senior lecturer, Department of design and technologies of electronic and laser means, Federal State Autonomous Educational Institution of Higher Education "St. Petersburg State University of Aerospace Instrumentation"; Researcher, Research department, Military Academy of Communication. Research interests: theory of signal transmission, spectral efficiency of signals, noise immunity of control and communication channels of radio engineering systems. The number of publications — 114. dvornik.92@mail.ru; 3, Tikhoretskiy Av., 194064, St. Petersburg, Russia; office phone: +7(812)247-9820.

Ustinov Andrew — Ph.D., Dr.Sci., Professor, Leading researcher, FGUP "GosNIIPP". Research interests: theory of communication, noise immunity of infocommunication channels of radio engineering systems, compression of video and speech information. The number of publications — 150. ust_m_a@mail.ru; 29, Emb. of the Obvodny Canal, 191167, St. Petersburg, Russia; office phone: +7(812)274-3156.

УДК 004.056.3 **DOI** 10.15622/ia.21.5.7

В.Н. Якимов

ВОССТАНОВЛЕНИЕ ДИСКРЕТНОЙ ВРЕМЕННОЙ ПОСЛЕДОВАТЕЛЬНОСТИ СИГНАЛА НА ОСНОВЕ ЛОКАЛЬНОЙ АППРОКСИМАЦИИ С ИСПОЛЬЗОВАНИЕМ РЯДА ФУРЬЕ ПО ОРТОГОНАЛЬНОЙ СИСТЕМЕ ТРИГОНОМЕТРИЧЕСКИХ ФУНКЦИЙ

Якимов В.Н. **Восстановление дискретной временной последовательности сигнала на основе локальной аппроксимации с использованием ряда Фурье по ортогональной системе тригонометрических функций.**

Аннотация. В статье рассмотрена разработка математического и алгоритмического обеспечения для восстановления отсчетов на проблемных участках дискретной последовательности непрерывного сигнала. Цель работы состояла в том, чтобы обеспечить восстановление утраченных отсчетов или участков отсчетов с непостоянной искаженной временной сеткой при осуществлении дискретизации сигнала с равномерным шагом и одновременно обеспечить снижение вычислительной сложности цифровых алгоритмов восстановления. Решение поставленной задачи осуществлено на основе метода локальной аппроксимации. Спецификой применения этого метода стало использование двух подпоследовательностей отсчетов, расположенных симметрично по отношению к восстанавливаемому участку последовательности. В качестве аппроксимирующей модели используется ряд Фурье по ортогональной системе тригонометрических функций. Оптимальное решение задачи аппроксимации основано на критерии минимума квадратичной погрешности. Для данного вида погрешности получены математические соотношения. Они позволяют оценить ее значение в зависимости от порядка модели и числа отсчетов подпоследовательностей, по которым осуществляется процедура восстановления. Особенность полученных в настоящей работе математических соотношений для восстановления сигнала заключается в том, что они не требуют предварительного вычисления коэффициентов ряда Фурье. Они обеспечивают непосредственно вычисление значений восстанавливаемых отсчетов. При этом в случае выбора четного числа отсчетов в подпоследовательностях, используемых для восстановления, не требуется выполнять операции умножения. Всё это обеспечило снижение вычислительной сложности разработанного алгоритма для восстановления сигнала. Экспериментальные исследования алгоритма осуществлялись на основе имитационного моделирования с использованием модели сигнала, представляющей собой аддитивную сумму гармонических компонент со случайной начальной фазой. Численные эксперименты показали, что разработанный алгоритм обеспечивает результат восстановления отсчетов сигнала с достаточно низкой погрешностью. Алгоритм реализован в виде программного модуля. Работа модуля осуществляется на основе асинхронного управления процессом восстановления отсчетов. Он может быть применен в составе метрологически значимого программного обеспечения систем цифровой обработки сигналов.

Ключевые слова: сигналы дискретного времени, последовательность отсчетов, восстановление сигнала, локальная аппроксимация, тригонометрический ряд Фурье.

1. Введение.

Восстановление значений отсчетов на поврежденных участках дискретной последовательности, полученной в результате равномерной дискретизации непрерывного во времени

сигнала, является активной областью исследований различных научных и прикладных дисциплин [1-3]. В частности это касается радиолокации, вибродиагностики, беспроводной передачи данных, обработки акустических сигналов и т.д.

Потери отсчетов и непреднамеренно возникшая неравномерность временной сетки на отдельных участках дискретной последовательности в процессе ее формирования и передачи, усложняют цифровую обработку сигнала и в итоге снижают достоверность конечных результатов. Это приводит к необходимости корректировки цифрового представления сигналов. Правильный выбор способа устранения возникшей проблемы зависит от прикладной области и задач проводимых исследований, а также от инструментальных средств, используемых для преобразования непрерывного сигнала в цифровой код и формирования однородной дискретной последовательности отсчетов во временной области.

В обычных условиях, когда цифровая обработка анализируемого сигнала осуществляется без пространственно-временных ограничений выполнения процедуры дискретизации, решение задачи по восстановлению проблемных значений отсчетов может быть осуществлено путем простого повторного формирования его дискретной выборки. В отдельных случаях качество представления сигнала в дискретном виде может быть улучшено за счет использования более совершенных устройств аналого-цифрового преобразования, если такая возможность существует. В условиях невозможности повторного воспроизведения процедуры дискретизации или ограниченности предоставляемого времени для преобразования сигнала в цифровой код такой подход оказывается невыполнимым. В отдельных случаях проблемные участки с неверными числовыми данными просто исключают из обрабатываемых последовательностей и их представляют в цифровом коде с использованием нулевых или усредненных значений. Это снижает вероятность грубых ошибок результатов вычисления оценок характеристик сигнала. Однако при коротких выборках погрешность таких результатов может оказаться весьма существенной.

Для того чтобы получить приемлемую оценку значений проблемных отсчетов кроме самой дискретной последовательности сигнала следует принимать во внимание и имеющуюся информацию о причинах, оказавших негативное влияние на ее формирование. Возможность восстановления отсчетов сигнала определяется как с учетом специфики непреднамеренно возникших технических ограничений, вызвавших снижение способности устройств цифрового

кодирования выполнять требуемые преобразования на отдельных временных интервалах, так и с учетом особенностей потерь из-за сложных условий распространения самого сигнала в процессе его дискретизации. В соответствии с этим характерной чертой разработки существующих алгоритмов восстановления дискретных последовательностей сигналов является их функциональная направленность осуществлять оценку значений проблемных отсчетов с учетом неопределенностей и ограничений для каждого конкретного приложения.

Подлежащие цифровой обработке дискретные последовательности сигналов приходится восстанавливать по причине потери отсчетов, как из-за ухудшения физических свойств самого источника сигнала, так и из-за аппаратных ограничений аналого-цифрового преобразования принимающей стороны [4-6]. Недоступность отсчетов сигнала может возникнуть, когда некоторые участки выборки были пропущены из-за внешнего искажающего воздействия или высокого уровня фоновых шумов [7]. Кроме того дискретные последовательности сигналов могут иметь участки отсчетов с непостоянной временной сеткой. Причиной этого может стать непреднамеренная вариация периода следования тактовых импульсов синхронизирующих устройств, вызывающая эффект дрожания, или системные особенности представления данных в цифровом виде, приводящие к неравномерной во времени выборке [8-11]. В отдельных случаях для достижения определенных полезных свойств дискретная выборка сигнала может быть преднамеренно сформирована нерегулярной за счет рандомизации процедуры аналого-цифрового преобразования. Для того чтобы можно было бы использовать хорошо зарекомендовавшие себя классические цифровые методы обработки сигналов осуществляют ее преобразование в последовательность с равномерной временной сеткой. В частности такая процедура осуществляется на основе спектральной оценки с последующей реконструкцией формы сигнала и равномерной передискретизацией. Это позволяет эффективно восстанавливать и обрабатывать сигналы высокой сложности, но требует выполнения и относительно сложных вычислений [12, 13]. Неравномерность дискретного представления сигнала также может возникнуть в многоканальных высокоскоростных системах цифровой обработки сигнала, использующих для дискретизации сигнала несколько параллельно работающих аналого-цифровых преобразователей с более низкой скоростью. Это связано с тем, что на участках объединения несинхронизированных однородных выборок в одну общую

дискретную последовательность могут иметь место относительные смещения и потери отсчетов [14-16].

С целью получения эффективных алгоритмов для восстановления дискретной последовательности сигнала с проблемными участками отсчетов при их разработке используют специальные методы частотно-временного преобразования или разложения во временные ряды [1, 17-25]. Первый вид методов позволяет за счет преобразований с применением различного рода адаптивных ядер реконструировать форму сигнала. Применение адаптивного ядра позволяет восстановить равномерную временную сетку дискретного представления сигнала с исходным разрешением и соответствующие ее узлам значения отсчетов, но при этом может оказать сглаживающее влияние, что может негативно отразиться на форме частотного спектра сигнала. Второй тип методов на основе аппроксимативного подхода предоставляет возможность восстанавливать утраченные выборки отсчетов непосредственно во временной области. Аппроксимация позволяет воспроизводить аналитическое описание обрабатываемого фрагмента сложного сигнала. Она обеспечивает хорошую сходимость и производительность алгоритмов восстановления. Однако эти характеристики ухудшаются по мере увеличения числа проблемных отсчетов. Это объясняется тем, что практическая реализация таких алгоритмов, как правило, приводит к необходимости выполнения большого числа операций. В особенности это касается операций умножения. Проблема заключается в том, что выполнение операций умножения занимает относительно много времени и оказывает отрицательное влияние на мультипликативную сложность цифровых вычислительных процедур. Следствием этого является снижение эффективности восстановления отсчетов сигнала непосредственно в реальном режиме обработки дискретной последовательности.

Таким образом, в процессе восстановления проблемных участков равномерной дискретной последовательности актуальной задачей является уменьшение искажающего влияния неопределенности получаемых оценок значений утраченных отсчетов на динамику изменения сигнала во времени и на конечные результаты цифровой обработки сигнала в целом. При этом математическое решение этой задачи должно быть направленно на разработку практичных и эффективных в вычислительном отношении алгоритмов, программная реализация которых обеспечивала бы низкую мультипликативную сложность вычисления оценок значений восстанавливаемых отсчетов.

2. Разработка математического обеспечения для восстановления дискретной временной последовательности сигнала на основе метода локальной аппроксимации. Восстановление сигнала представляет собой процедуру, в процессе выполнения которой осуществляются операции с имеющимися отсчетами сигнала с целью оценки тех значений, которые имели бы отсчеты, если бы они не были утрачены. Математическая формулировка задачи восстановления во времени сигнала по его известным дискретным значениям сводится к нахождению приемлемой функциональной зависимости, позволяющей воспроизвести его исходную форму. Чтобы спроектировать эффективный алгоритм восстановления дискретной последовательности, необходимо учитывать характерные особенности потерь отсчетов, приводящие к ухудшению результатов цифровой обработки сигнала. Основная идея заключается в том, чтобы как можно точнее смоделировать проблемные участки с потерями отсчетов. Однако при этом следует принимать во внимание тот факт, что сложно найти общее представление о потерях отсчетов в дискретных последовательностях. Поэтому следует определить границы области использования разрабатываемого алгоритма, в пределах которых восстановление значений отсчетов должно быть численно стабильным и обеспечивать получение удовлетворительных результатов.

Поставленную задачу восстановления будем решать исходя из выполнения условия, что цифровой обработке подвергается непрерывный во времени и ограниченный по спектральному составу верхней граничной частотой $F_{\max}$ сигнал $x(t)$. Будем также считать, что сигнал удовлетворяет условиям стационарности или квазистационарности. Выполнение условия квазистационарности позволяет учесть также те сигналы, для которых частотно-временные характеристики в пределах заданного интервала времени наблюдения можно считать неизменными, и их распространение происходит по одному и тому же временному закону в медленно изменяющейся среде. Пусть в результате аналого-цифрового преобразования такого сигнала с равномерным шагом дискретизации будем иметь дискретную последовательность отсчетов x_i. При этом процедура дискретизации сигнала осуществлена с частотой превышающей минимально необходимую частоту дискретизации, определяемую согласно теореме Котельникова. Будем считать, что для $i \in [0; \nu]$ и $i \in [\nu + m + 1; N - 1]$ отсчеты x_i получены без искажения временной

сетки, и их значения известны точно. Значения m отсчетов x_i для $i \in [\nu+1;\nu+m]$ образуют проблемный участок последовательности. Они могут быть, как искажены или утрачены, так и не соответствовать равномерной временной сетке. В любом случае значения этих отсчетов необходимо восстановить.

Для восстановления проблемного участка отсчетов x_i воспользуемся методом локальной аппроксимации. Согласно этому методу восстанавливаемый участок последовательности аппроксимируется линейной комбинацией конечного набора известных и фиксированных базисных функций. При этом в качестве параметров локальности, определяющих временной интервал применения аппроксимирующей модели, для каждого проблемного участка в его окрестности выбирается область с известными значениями отсчетов. В пределах этой области определяется подпоследовательность отсчетов, на основе которой осуществляется построение аппроксимирующей модели, как решение локальной экстремальной задачи [26]. Применение локальной аппроксимации позволяет получать оценки значений проблемных отсчетов с учетом только динамики изменения известных значений подпоследовательностей в близких узлах. Одним из явных преимуществ такого метода является то, что он обеспечивает оперативность процедуры восстановления. Однако следует понимать, что локальная аппроксимация применима для восстановления значений отсчетов на временных интервалах, на которых прослеживается устойчивая зависимость значений проблемных участков и используемых для восстановления их отсчетов подпоследовательностей. Исходя из сказанного, применение данного метода на практике предполагает выбор аппроксимирующей модели, критерия оценки ее параметров, определение области и состава локально расположенных подпоследовательностей.

В соответствии методом локальной аппроксимации будем осуществлять восстановление значений отсчетов x_i для $i \in [\nu+1;\nu+m]$ в узлах временной сетки с равномерным шагом равным исходному интервалу дискретизации. При этом в окрестности восстанавливаемого участка выберем две локальные подпоследовательности отсчетов x_i с известными значениями, где $i \in [\nu-L+1;\nu]$ и $i \in [\nu+m+1;\nu+m+L]$. Схематично восстанавливаемый участок представлен на рисунке 1.

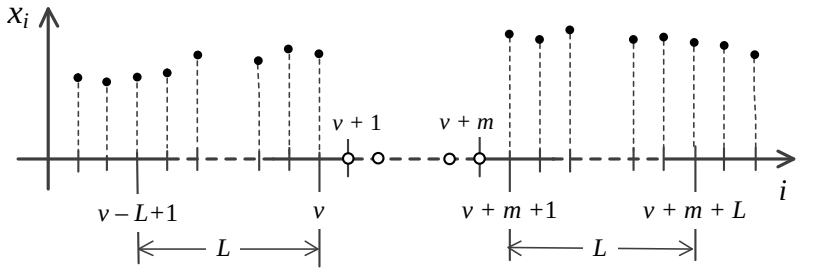


Рис. 1. Фрагмент последовательности с восстанавливаемым участком

Фактически будем использовать L предыдущих (слева) отсчетов и L последующих (справа) отсчетов от проблемного участка последовательности. Отметим, что L должно удовлетворять условиям:

$$\nu - L + 1 \geq 0 \text{ и } \nu + m + L \leq N - 1. \quad (1)$$

В качестве аппроксимирующей модели выберем тригонометрический ряд Фурье [27]:

$$\hat{x}_i = a_0 + \sum_{k=1}^{p-1} a_k \cos\frac{2\pi k}{L} i + \sum_{k=1}^{p-1} b_k \sin\frac{2\pi k}{L} i, \quad (2)$$

где p – порядок модели, a_k и b_k – постоянные коэффициенты.

Тригонометрический ряд Фурье представляет собой функциональный ряд, основу построения которого составляют периодические базисные функции синусов и косинусов с кратными частотами. В (2) функции синусов и косинусов попарно ортогональны и образуют ортогональные системы для $i \in [-L; L]$. Ортогональные системы характеризуются хорошими вычислительными свойствами и обеспечивают равномерную сходимость результата в пределах интервала аппроксимации во всей области определения сигнала [27-30].

Значения коэффициентов a_k и b_k для модели (2) порядка p будем находить из критерия минимума квадратичной погрешности:

$$\delta_p = \sum_i (\hat{x}_i - x_i)^2 \to \min, \quad (3)$$

где $i \in [\nu - L + 1; \nu]$ и $i \in [\nu + m + 1; \nu + m + L]$.

В соответствии с критерием (3) нахождение значений коэффициентов a_k и b_k аппроксимирующей модели осуществляется таким образом, чтобы сумма квадратов ошибок приближения в узлах временной сетки дискретизации была минимальной. При этом такой подход позволяет сглаживать возможные ошибки за счет использования избыточного числа отсчетов в подпоследовательностях.

Для обеспечения выполнения критерия (3) найдем частные производные первого порядка по a_0, a_n и b_n, где $n \in [1; p-1]$:

$$\frac{\partial \delta_p}{\partial a_0} = 0, \ \frac{\partial \delta_p}{\partial a_n} = 0 \ \text{и} \ \frac{\partial \delta_p}{\partial b_n} = 0. \tag{4}$$

С целью упрощения записей математических соотношений введем обозначения:

$$\alpha_{k,i} = \cos\frac{2\pi k}{L} i \ \text{и} \ \beta_{k,i} = \sin\frac{2\pi k}{L} i. \tag{5}$$

После вычисления частных производных (4) с учетом (5) будем иметь систему уравнений, решение которой позволяет получить соотношения для определения значений коэффициентов a_k и b_k:

$$2La_0 + \sum_{k=1}^{p-1} a_k \sum_i \alpha_{k,i} + \sum_{k=1}^{p-1} b_k \sum_i \beta_{k,i} = \sum_i x_i, \tag{6}$$

$$a_0 \sum_i \alpha_{n,i} + \sum_{k=1}^{p-1} a_k \sum_i \alpha_{k,i}\alpha_{n,i} + \sum_{k=1}^{p-1} b_k \sum_i \beta_{k,i}\alpha_{n,i} = \sum_i x_i \alpha_{n,i}, \tag{7}$$

$$a_0 \sum_i \beta_{n,i} + \sum_{k=1}^{p-1} a_k \sum_i \alpha_{k,i}\beta_{n,i} + \sum_{k=1}^{p-1} b_k \sum_i \beta_{k,i}\beta_{n,i} = \sum_i x_i \beta_{n,i}, \tag{8}$$

где $n \in [1; p-1]$; $i \in [\nu - L + 1; \nu]$ и $i \in [\nu + m + 1; \nu + m + L]$.

Ввиду того, что функции $\alpha_{k,i}$ и $\beta_{k,i}$ образуют ортогональные системы, справедливы соотношения [27, 30]:

$$\sum_i \alpha_{k,i}\alpha_{n,i} = \sum_i \cos\frac{2\pi k}{L} i \cos\frac{2\pi n}{L} i = \begin{cases} 0, & k \neq n; \\ L, & k = n. \end{cases} \quad (9)$$

$$\sum_i \beta_{k,i}\beta_{n,i} = \sum_i \sin\frac{2\pi k}{L} i \sin\frac{2\pi n}{L} i = \begin{cases} 0, & k \neq n; \\ L, & k = n. \end{cases} \quad (10)$$

$$\sum_i \cos\frac{2\pi k}{L} i \sin\frac{2\pi n}{L} i = 0\,,\ k \in [1; p-1] \text{ и } n \in [0; p-1]\,. \quad (11)$$

Принимая во внимание (9)-(11), для a_k и b_k получаем:

$$a_0 = \frac{1}{2L}\sum_i x_i\,,\ a_n = \frac{1}{L}\sum_i x_i\alpha_{n,i}\,,\ b_n = \frac{1}{L}\sum_i x_i\beta_{n,i}\,. \quad (12)$$

С учетом того, что $i \in [\nu - L + 1; \nu]$ и $i \in [\nu + m + 1; \nu + m + L]$, в явном виде для (12) будем иметь:

$$a_0 = \frac{1}{2L}\sum_{i=0}^{L-1}(x_{\nu-i} + x_{i+\nu+m+1})\,, \quad (13)$$

$$a_n = \frac{1}{L}\sum_{i=0}^{L-1}\left(x_{\nu-i}\cos\frac{2\pi n(\nu-i)}{L} + x_{i+\nu+m+1}\cos\frac{2\pi n(i+\nu+m+1)}{L}\right), \quad (14)$$

$$b_n = \frac{1}{L}\sum_{i=0}^{L-1}\left(x_{\nu-i}\sin\frac{2\pi n(\nu-i)}{L} + x_{i+\nu+m+1}\sin\frac{2\pi n(i+\nu+m+1)}{L}\right). \quad (15)$$

Соотношения (13)-(15) определяют вычислительный алгоритм нахождения коэффициентов аппроксимирующей модели (2) при заданном числе отсчетов подпоследовательностей L и порядке модели p в процессе восстановления m значений проблемных отсчетов последовательности x_i. При этом квадратическая погрешность равна:

$$\delta_p = \sum_{i=0}^{L-1}(x_{\nu-i}^2 + x_{i+\nu+m+1}^2) - 2La_0^2 - L\sum_{n=1}^{p-1}(a_n^2 + b_n^2)\,. \quad (16)$$

Согласно (16) при выбранном значении L погрешность δ_p полностью зависит от порядка модели p. Однако следует иметь ввиду, что увеличение порядка модели p ведет к усложнению вычислительного процесса восстановления. Рассмотрим вариант рационального выбора порядка модели p.

Из (14)-(15) следует:

$$a_n = a_{L-n} \text{ и } b_n = -b_{L-n}. \tag{17}$$

Отсюда приходим к простому выводу, что $0 \le n \le L/2$. При этом имеем, что $b_{L/2} = 0$. Как следствие из этого получаем, что число отсчетов L в каждой из подпоследовательностей, по которым осуществляется локальная аппроксимация, предпочтительно выбирать четным. В этом случае порядок модели целесообразно выбирать равным не более чем $p = (L/2+1)$. В свою очередь при предварительно выбранном порядке модели p число отсчетов по возможности следует задавать равным $L = 2(p-1)$ для каждой из подпоследовательностей. С учетом выполнения этих условий рациональная с точки зрения выбора порядка и его влияния на вычислительную сложность аппроксимирующая модель (2) будет иметь вид:

$$\hat{x}_i = a_0 + \sum_{k=1}^{L/2} a_k \cos\frac{2\pi k}{L} i + \sum_{k=1}^{L/2-1} b_k \sin\frac{2\pi k}{L} i. \tag{18}$$

Подставив $\hat{x}_i$ из (18) в (3), получаем, что с учетом (13)-(15) для $p = (L/2+1)$ квадратическая погрешность равна:

$$\delta_{\frac{L}{2}+1} = \frac{1}{2}\sum_{i=m+1}^{L}(x_{i+\nu} + x_{i+\nu-L})^2 + \frac{1}{2}\sum_{i=1}^{m}(x_{i+\nu-L} + x_{i+\nu+L})^2 + \\ + \frac{1}{2L}\left(\sum_{i=0}^{L-1}(-1)^i (x_{\nu-1} + (-1)^{m+1} x_{i+\nu+m+1})\right)^2. \tag{19}$$

Далее получим соотношения, которые позволяют упростить алгоритм восстановления значений отсчетов x_i. С этой целью примем, что $i = \nu + j$, где $j = 1,2,3,...m$. Тогда (2) примет вид:

$$\hat{x}_{\nu+j} = a_0 + \sum_{k=1}^{p-1} a_k \cos\frac{2\pi k(\nu+j)}{L} + \sum_{k=1}^{p-1} b_k \sin\frac{2\pi k(\nu+j)}{L}. \quad (20)$$

Подставим a_k и b_k из (13)-(15) в (20). После изменения порядка суммирования по индексам k и i, перегруппировки и приведения подобных членов будем иметь:

$$\hat{x}_{\nu+j} = \frac{1}{2L}\sum_{i=0}^{L-1}(x_{\nu-i} + x_{i+\nu+m+1}) + \frac{1}{L}\sum_{i=0}^{L-1} x_{\nu-i}\sum_{k=1}^{p-1}\left(C1_{k,i} + S1_{k,i}\right) + \\ + \frac{1}{L}\sum_{i=0}^{L-1} x_{i+\nu+m+1}\sum_{k=1}^{p-1}\left(C2_{k,i} + S2_{k,i}\right), \quad (21)$$

$$C1_{k,i} = \cos\frac{2\pi k(\nu+j)}{L}\cos\frac{2\pi k(\nu-i)}{L}, \quad (22)$$

$$C2_{k,i} = \cos\frac{2\pi k(\nu+j)}{L}\cos\frac{2\pi k(i+\nu+m+1)}{L}, \quad (23)$$

$$S1_{k,i} = \sin\frac{2\pi k(\nu+j)}{L}\sin\frac{2\pi k(\nu-i)}{L}, \quad (24)$$

$$S2_{k,i} = \sin\frac{2\pi k(\nu+j)}{L}\sin\frac{2\pi k(i+\nu+m+1)}{L}. \quad (25)$$

В (22) и (23) произведения косинусов могут быть представлены в виде суммы косинусов, а в (24) и (25) произведения синусов в виде разности косинусов. Осуществив соответствующие преобразования, после приведения подобных членов получаем:

$$\hat{x}_{\nu+j} = \frac{1}{2L}\sum_{i=0}^{L-1} x_{\nu-i}\left(1 + 2\sum_{k=1}^{p-1}\cos\frac{2\pi k(j+i)}{L}\right) + \\ + \frac{1}{2L}\sum_{i=0}^{L-1} x_{i+\nu+m+1}\left(1 + 2\sum_{k=1}^{p-1}\cos\frac{2\pi k(i+m+1-j)}{L}\right). \quad (26)$$

В (26) суммы по индексу k будут равны:

$$\sum_{k=1}^{p-1}\cos\frac{2\pi k(j+i)}{L}=\frac{1}{2}\left(\frac{\sin((2p-1)\pi(j+i)/L)}{\sin(\pi(j+i)/L)}-1\right), \quad (27)$$

$$\sum_{k=1}^{p-1}\cos\frac{2\pi k(i+m+1-j)}{L}=\frac{1}{2}\left(\frac{\sin((2p-1)\pi(i+m+1-j)/L)}{\sin(\pi(i+m+1-j)/L)}-1\right). \quad (28)$$

С учетом (27) и (28) будем иметь:

$$\hat{x}_{\nu+j}=\frac{1}{2L}\sum_{i=0}^{L-1}x_{\nu-i}\frac{\sin((2p-1)\pi(j+i)/L)}{\sin(\pi(j+i)/L)}+ \\ +\frac{1}{2L}\sum_{i=0}^{L-1}x_{i+\nu+m+1}\frac{\sin((2p-1)\pi(i+m+1-j)/L)}{\sin(\pi(i+m+1-j)/L)}\ . \quad (29)$$

В (29) в первой сумме заменим $j+i$ на i, а во второй сумме заменим $i+m+1-j$ на i. С учетом этого окончательно получаем:

$$\hat{x}_{\nu+j}=\frac{1}{2L}\left(\sum_{i=j}^{L-1+j}x_{\nu+j-i}D_i(2p-1,L)+\sum_{i=m+1-j}^{L+m-j}x_{\nu+j+i}D_i(2p-1,L)\right), \quad (30)$$

$$D_i(p,L)=\frac{\sin((2p-1)i\pi/L)}{\sin(\pi i/L)}. \quad (31)$$

Соотношения (30) и (31) определяют алгоритм восстановления отсчетов $\hat{x}_{\nu+j}$ для $j=1,2,3,...m$ без предварительного вычисления коэффициентов a_k и b_k. В (31) $\sin(\pi i/L)=0$ для $i=\lambda L$, где $\lambda=0,1,2,...$. В результате в (31) получаем деление на ноль. Эту ситуацию необходимо учитывать и не допускать в процессе выполнения процедуры восстановления. Для этого случая имеем:

$$D_i(p,L)=\frac{\sin(2p-1)\lambda\pi}{\sin\lambda\pi}=(2p-1). \quad (32)$$

Для того чтобы осуществлять вычисления $\hat{x}_{\nu+j}$ без проверки условия $i=\lambda L$ соотношение (30) запишем следующим образом:

$$\hat{x}_{\nu+j}=\frac{1}{2L}\Bigg((2p-1)(x_{\nu+j-L}+x_{\nu+j+L})+\sum_{i=j}^{L-1}x_{\nu+j-i}D_i(p,L)+$$
$$+\sum_{i=1}^{j-1}x_{\nu+j-i-L}D_i(p,L)+\sum_{i=m+1-j}^{L-1}x_{\nu+j+i}D_i(p,L)+\sum_{i=1}^{m-j}x_{\nu+j+i+L}D_i(p,L)\Bigg). \quad (33)$$

Полученное соотношение справедливо для любого порядка модели p. Однако как было показано выше с целью снижения вычислительной сложности в процессе восстановления отсчетов $\hat{x}_{\nu+j}$ порядок модели p при четном числе отсчетов подпоследовательностей L по возможности целесообразно выбирать равным $p=(L/2+1)$. С учетом такого выбора порядка (33) примет вид:

$$\hat{x}_{\nu+j}=\frac{x_{\nu+j-L}+x_{\nu+j+L}}{2}+\frac{(-1)^j}{2L}\sum_{i=0}^{L-1}(-1)^i(x_{\nu-i}-(-1)^m x_{\nu+i+m+1}). \quad (34)$$

С математической точки зрения соотношение (34) является достаточно простым. Надо только помнить, что число отсчетов L в подпоследовательностях, используемых для аппроксимации, обязательно должно быть четным. Согласно этому соотношению вычисление оценок значений отсчетов $\hat{x}_{\nu+j}$ в своей основе сводится к выполнению операций суммирования. При этом необходимость в выполнении операций умножения фактически отсутствуют. Это позволяет говорить о том, что практическая реализация соотношения (34) приводит к вычислительным процедурам с низкой мультипликативной сложностью. Такой результат повышает вычислительную эффективность восстановления последовательности отсчетов сигнала.

3. Разработка алгоритмического обеспечения для восстановление дискретной временной последовательности сигнала. Соотношения (33) и (34) положены в основу разработки вычислительных алгоритмов для восстановления дискретной временной последовательности сигнала. Алгоритмы разработаны с ориентированием на принципы модульного программирования. Ниже представлена краткая запись этих алгоритмов, отражающая выполнение основных действий при их реализации.

Алгоритм восстановления на основе соотношения (33).

1. Инициализация восстановления выборки отсчетов:
   1.1. идентифицируются параметры проблемного участка v и m;

1.2. задается порядок аппроксимирующей модели p;

1.3. задается число отсчетов L для подпоследовательностей с учетом условий $\nu - L + 1 \geq 0$ и $\nu + m + L \leq N - 1$.

2. Выполнение процедуры восстановления выборки отсчетов:

2.1. для $i = 1, 2, 3, ..., L - 1$ вычисляется массив отсчетов:

$$DpL[i] = \sin((2 \times p - 1) \times i \times \pi / L) / \sin(\pi \times i / L);$$

2.2. для $j = 1, 2, 3, ..., m$ выполняются действия с 2.2.1 по 2.2.6:

2.2.1. $\hat{x}[\nu + j] = (2 \times p - 1)(x[\nu + j - L] + x[\nu + j + L])$;

2.2.2. для $i = j, j + 1, j + 2, ..., L - 1$ вычисляется:

$$\hat{x}[\nu + j] = \hat{x}[\nu + j] + x[\nu + j - i] \times DpL[i];$$

2.2.3. для $i = 1, 2, 3, ..., j - 1$ вычисляется:

$$\hat{x}[\nu + j] = \hat{x}[\nu + j] + x[\nu + j - i - L] \times DpL[i];$$

2.2.4. для $i = m - j + 1, m - j + 2, m - j + 3, ..., L - 1$ вычисляется:

$$\hat{x}[\nu + j] = \hat{x}[\nu + j] + x[\nu + j - i - L] \times DpL[i];$$

2.2.5. для $i = 1, 2, 3, ...m - j$ вычисляется:

$$\hat{x}[\nu + j] = \hat{x}[\nu + j] + x[\nu + j + i + L] \times DpL[i];$$

2.2.6. для текущего значения j вычисляется оценка отсчета:

$$\hat{x}[\nu + j] = \hat{x}[\nu + j] / (2 \times L).$$

3. В исходной последовательности отсчеты $x[i]$ с индексами $i = \nu + 1, \nu + 2, \nu + 3, ..., \nu + m$ замещаются на оценки $\hat{x}[\nu + j]$.

Особенностью алгоритма на основе соотношения (33) является то, что значения отсчетов $DpL[i]$, которые вычисляются согласно (31), не зависят от значений параметров восстанавливаемого участка ν и m. Поэтому, если возникает ситуация, требующая восстановления нескольких проблемных участков в пределах одной и той же последовательности, то нет необходимости в повторном вычислении

массива этих отсчетов при неизменных значениях p и L. С учетом этого он формируется предварительно непосредственно после инициализации параметров восстановления перед выполнением основной процедуры вычисления оценок восстанавливаемых отсчетов. Повторное формирование этого массива осуществляется только при изменении параметров аппроксимации. Однако целесообразно сохранять все его предыдущие версии для случая возврата к ранее используемым параметрам аппроксимации. Отметим также то, что вычислительные процедуры 2.2.2 – 2.2.5 являются относительно независимыми при их выполнении. Поэтому их можно представить в виде отдельных потоков обработки отсчетов, которые могут выполняться параллельно. При этом многопоточность обработки будет учитываться в виде суммирования полученных результатов 2.2.2 – 2.2.5 при выполнении 2.2.6. Такой подход ускоряет вычислительный процесс и способствует повышению производительности процедуры восстановления отсчетов для случая независимого задания значений L и p.

Алгоритм восстановления на основе соотношения (34).

1. Инициализация восстановления выборки отсчетов:
   1.1. идентифицируются параметры проблемного участка ν и m;
   1.2. задается порядок аппроксимирующей модели p или четное число отсчетов L для подпоследовательностей с учетом выполнения условий $\nu - L + 1 \geq 0$ и $\nu + m + L \leq N - 1$;
      1.2.1. если задается порядок модели p, то $L = 2 \times (p - 1)$;
      1.2.2. если задается число отсчетов L для подпоследовательностей, то $p = (L / 2 + 1)$.
2. Выполнение процедуры восстановления выборки отсчетов:
   2.1. $Sum1 = 0$ и $Sum2 = 0$;
   2.2. для $i = 0, 2, 4, ..., L - 2$ вычисляются промежуточные суммы:

$$Sum1 = Sum1 + x[\nu - i] - x[\nu - i - 1];$$

$$Sum2 = Sum2 + x[\nu + i + m + 1] - x[\nu + i + m + 2];$$

   2.3. если m чётное, то $Sum = (Sum1 - Sum2) / 2L$,
   иначе $Sum = (Sum1 + Sum2) / 2L$;
   2.4. для $j = 1, 2, 3, ..., m$ выполняются действия:
      2.4.1. $\hat{x}[\nu + j] = (x[\nu + j - L] + x[\nu + j + L]) / 2$;
      2.4.2. для текущего значения j вычисляется оценка отсчета:

если j чётное, то $\hat{x}[\nu + j] = \hat{x}[\nu + j] + Sum$,

иначе $\hat{x}[\nu + j] = \hat{x}[\nu + j] - Sum$;

3. В исходной последовательности отсчеты $x[i]$ с индексами $i = \nu + 1, \nu + 2, \nu + 3, ..., \nu + m$ замещаются на оценки $\hat{x}[\nu + j]$.

Алгоритм на основе соотношения (34) на этапе инициализации его выполнения предлагает сделать альтернативный выбор между приоритетным заданием порядка аппроксимирующей модели p и числом отсчетов L для подпоследовательностей, которое обязательно должно быть только четным числом. При задании одного из этих параметров, второй рассчитывается согласно 1.2.1 или 1.2.2. Алгоритм не требует выполнения операций цифрового умножения. Основу его выполнения составляет вычисление промежуточных сумм $Sum1$ и $Sum2$ согласно 2.2. Эти суммы можно вычислять независимо друг от друга. Поэтому при программной реализации алгоритма процедуры вычисления сумм $Sum1$ и $Sum2$ можно реализовать в виде отдельных потоков обработки отсчетов подпоследовательностей, которые будут выполняться параллельно. Объединение результатов многопоточных вычислений будет учитываться при выполнении 2.3. Данный алгоритм совместно с многопоточной организацией определяемых им вычислительных процедур повышают производительность восстановления отсчетов. Однако следует учитывать тот факт, что применение алгоритма ограничено требованием выполнения обязательного условия взаимно однозначного соответствия порядка аппроксимирующей модели p и числа отсчетов L для подпоследовательностей.

В общем случае рассмотренные алгоритмы позволяют эффективно и рационально в вычислительном отношении организовать процесс восстановление отсчетов, поскольку большинство вычислительных процедур можно выполнять параллельно. Практическая реализация алгоритмов осуществлена в виде проблемно-ориентированного программного модуля. Функциональным назначением модуля является восстановление утраченных отсчетов или значений отсчетов на участках дискретной последовательности сигнала с непостоянной искаженной временной сеткой. Работа модуля осуществляется на основе асинхронного управления процессом восстановления отсчетов. Реализация асинхронной модели управления обеспечивает выполнение процедуры вычисления оценок значений отсчетов в режиме многозадачности. Модуль запускается по мере необходимостью восстановления конкретного участка дискретной последовательности сигнала.

Процедура восстановления проходит без блокировки выполнения вычислительного процесса, реализуемого основной программой по обработке сигнала. Это приводит к тому, что возможности вычислительных средств используются наиболее рационально. При этом предварительно устанавливается режим задания порядка аппроксимирующей модели и число отсчетов для подпоследовательностей, по которым осуществляется восстановление проблемных участков сигнала. Модуль может быть использован в качестве функционально самостоятельного компонента в составе интегрированного метрологически значимого программного обеспечения, выполняемого на прикладном уровне систем для комплексного цифрового анализа сложных сигналов [31, 32].

4. Численные эксперименты по восстановлению дискретных последовательностей сигнала. Исследования работоспособности и функциональных возможностей разработанных алгоритмов осуществлялось на основе имитационного моделирования дискретной последовательности сигнала с потерями отсчетов и последующего ее восстановления. В процессе моделирования для имитации работы модуля использовалась среда программирования TypeScript. Численные эксперименты проводились с применением набора тестовых моделей сигналов с ограниченным частотным спектром. В качестве математической основы представления таких моделей сигналов во временной области использовалась аддитивная смесь независимых в статистическом смысле гармонических компонент с нулевой постоянной составляющей. Аддитивный гармонический синтез модели сигнала предоставил возможность эффективно и просто регулировать его частотный состав и контролировать форму во временной области.

Значения частот гармонических компонент в процессе моделирования задавались в пределах от нуля до единицы. Они интерпретировались как нормированные частоты $f_k^н = f_k / F_{\max}$ по отношению к верхней граничной частоте $F_{\max}$ в спектре сигнала. Использование понятия нормированной частоты обеспечило моделирование различного частотного состава сигнала в единой полосе частот в пределах от нуля до единицы. Значения амплитуд гармонических компонент $A_k^н$ задавались в диапазоне от нуля до единицы и интерпретировались как нормированные по отношению к наибольшей амплитуде компоненты, присутствующей в составе сигнала. Начальные фазы гармонических компонент удовлетворяли условию $\phi_k \in [-\pi; +\pi]$. Они принимали случайные значения в

переделах указанного диапазона. Для этого использовался программный генератор чисел с равномерным законом распределения. Переход к нормированным значениям амплитуд и частот позволил формализовать проведение численных экспериментов и оценить потенциальные возможности разработанного математического и алгоритмического обеспечения для восстановления дискретных последовательностей сигналов с различными амплитудно-частотными характеристиками.

Для каждого эксперимента генерировалась реализация модели сигнала и формировалась его дискретная последовательность. Эксперименты повторялись для различных значений частот дискретизации с учетом присутствия в составе модели сигнала гармонической составляющей с наибольшей нормированной частотой. В частности, одна из таких реализаций модели содержала семь гармонических компонент, нормированные амплитуды и частоты которых имели значения: $A_1^н = 0,2$ и $f_1^н = 0,1$; $A_2^н = 0,35$ и $f_2^н = 0,25$; $A_3^н = 0,5$ и $f_3^н = 0,3$; $A_4^н = 0,7$ и $f_4^н = 0,35$; $A_5^н = 1$ и $f_5^н = 0,4$; $A_6^н = 0,7$ и $f_6^н = 0,45$; $A_7^н = 0,15$ и $f_7^н = 0,5$. В серии экспериментов дискретная последовательность данной реализации модели была получена в результате равномерной дискретизации с нормированной частотой $1000 f_7^н = 500$, которая является достаточно высокой. В этом случае соседние отсчеты дискретной последовательности могут иметь мало отличающиеся значения. Ограничение значений таких отсчетов до конечного числа разрядов позволило интерпретировать процесс восстановления в условиях аддитивного шума, обусловленного квантованием по уровню мало отличающихся значений сигнала.

В процессе моделирования часть экспериментов была направлена на оценку результатов восстановления в зависимости от выбора порядка p аппроксимирующей модели и числа отсчетов L в подпоследовательностях, используемых для восстановления. В таблицах 1 и 2 представлены результаты экспериментов по восстановлению участка с пятью отсчетами. В таблице 1 результаты получены для модели порядка p=5, а число отсчетов L каждой подпоследовательности было равно пяти, десяти и шестнадцати. В таблице 2 результаты получены для L=20, а порядок модели p был равен пяти, десяти и пятнадцати.

Эксперименты также проводились с учетом вариации числа восстанавливаемых отсчетов на проблемном участке при неизменном значении порядка модели и фиксированном числе отсчетов подпоследовательностей. Эти эксперименты имели особое значение

для оценки возможности восстановления утраченных отсчетов последовательностей сигналов, которые формируются в условиях ограниченного времени его наблюдения. В таблице 3 представлены результаты восстановления десяти, пятнадцати и двадцати отсчетов, где L=20 и p=5.

Для оценки отклонения результатов восстановления от истинных значений отсчетов использовалась относительная погрешность. Она рассчитывалась по формуле $\delta_i^x = (\hat{x}_i - x_i)/x_i$.

Результаты численных экспериментов, приведенные в таблицах 1, 2 и 3 не содержат грубых систематических ошибок, которые существенно бы искажали истинные значения восстанавливаемых отсчетов. Для каждого результата на всем интервале восстановления проблемного участка пределы изменения относительной погрешности отклонения вычисленных оценок отсчетов от их идеальных значений не превышают по абсолютной величине 0,01, т.е. одного процента.

Таблица 1. Результаты восстановления отсчетов для m=5, p=5 и L=5,10,16

j	$\nu + j$	$x_{\nu+j}$	L=5		L=10		L=16	
			$\hat{x}_{\nu+j}$	δ_i^x	$\hat{x}_{\nu+j}$	δ_i^x	$\hat{x}_{\nu+j}$	δ_i^x
1	201	1.9684	1.9628	-0.0028	1.9663	-0.0011	1.9678	-0.0003
2	202	1.9709	1.9679	-0.0015	1.9685	-0.0012	1.9618	-0.0046
3	203	1.9734	1.9729	-0.0002	1.9714	-0.0011	1.9674	-0.0031
4	204	1.9759	1.9779	-0.0010	1.9735	-0.0012	1.9731	-0.0015
5	205	1.9784	1.9828	0.0022	1.9763	-0.0010	1.9689	-0.0047

Таблица 2. Результаты восстановления отсчетов для m=5, L=20 и p=5,10,15

j	$\nu + j$	$x_{\nu+j}$	p=5		p=10		p=15	
			$\hat{x}_{\nu+j}$	δ_i^x	$\hat{x}_{\nu+j}$	δ_i^x	$\hat{x}_{\nu+j}$	δ_i^x
1	201	1.9684	1.9666	-0.0009	1.9596	-0.0044	1.9525	-0.0080
2	202	1.9709	1.9605	-0.0053	1.9616	-0.0047	1.9658	-0.0026
3	203	1.9734	1.9633	-0.0051	1.9647	-0.0044	1.9648	-0.0043
4	204	1.9759	1.9678	-0.0041	1.9665	-0.0047	1.9633	-0.0063
5	205	1.9784	1.9658	-0.0064	1.9695	-0.0044	1.9734	-0.0025

Таблица 3. Результаты восстановления отсчетов для p=5, L=20

j	$\nu + j$	$x_{\nu+j}$	m=10		m=15		m=20	
			$\hat{x}_{\nu+j}$	δ_i^x	$\hat{x}_{\nu+j}$	δ_i^x	$\hat{x}_{\nu+j}$	δ_i^x
1	201	1.9684	1.9673	-0.0005	1.9678	-0.0003	1.9711	0.0014
2	202	1.9709	1.9600	-0.0055	1.9599	-0.0056	1.9600	-0.0055
3	203	1.9734	1.9619	-0.0058	1.9613	-0.0061	1.9605	-0.0065
4	204	1.9759	1.9678	-0.0041	1.9676	-0.0042	1.9672	-0.0044
5	205	1.9784	1.9708	-0.0038	1.9714	-0.0035	1.9717	-0.0034
6	206	1.9807	1.9708	-0.0049	1.9715	-0.0047	1.9719	-0.0045
7	207	1.9831	1.9726	-0.0053	1.9722	-0.0055	1.9721	-0.0055
8	208	1.9854	1.9773	-0.0041	1.9760	-0.0047	1.9756	-0.0049
9	209	1.9876	1.9802	-0.0038	1.9801	-0.0037	1.9799	-0.0039
10	210	1.9898	1.9768	-0.0065	1.9813	-0.0042	1.9818	-0.0040
11	211	1.9920	–	–	1.9812	-0.0054	1.9818	-0.0051
12	212	1.9941	–	–	1.9839	-0.0051	1.9836	-0.0053
13	213	1.9962	–	–	1.9889	-0.0037	1.9878	-0.0042
14	214	1.9982	–	–	1.9909	-0.0036	1.9909	-0.0037
15	215	2.0002	–	–	1.9867	-0.0067	1.9907	-0.0047
16	216	2.0021	–	–	–	–	2.0022	-0.0058
17	217	2.0040	–	–	–	–	1.9945	-0.0047
18	218	2.0058	–	–	–	–	0.0027	-0.0027
19	219	2.0076	–	–	–	–	2.0001	-0.0037
20	220	2.0094	–	–	–	–	1.9905	-0.0093

Рассмотренный подход по восстановлению отсчетов был также апробирован с использованием разработанного программного модуля при анализе реальных сигналов, в частности при вибрационных исследованиях автобуса марки МАЗ-206067 Минского автомобильного завода. МАЗ-206067 – это автобус категории М3, класса I согласно классификации Правил ООН № 107 и Технического регламента Таможенного союза ТР ТС 018/2011 «О безопасности колесных транспортных средств». Автобус предназначен для перевозки пассажиров на городских и пригородных маршрутах средней загруженности. Он оснащен двигателем Mercedes-Benz OM904LA (дизель, 177 л.с.) и шестиступенчатой автоматической коробкой

передач ZF6HP504C. Для регистрации вибрационных сигналов использовались однокомпонентные акселерометры со встроенной электроникой ICP общего назначения 352C04 с разрешением 0,005 м/с². Они обеспечивают хорошее разрешение для измерения сигналов вибрации с малыми уровнями. Фирма изготовитель «PCB Piezotronics, Inc.». Установка датчиков и обработка вибрационных сигналов осуществлялись в соответствии с требованиями по измерению и представлению результатов измерений локальной вибрации, а также согласно требованиям по оценке воздействия локальной вибрации на пассажира и водителя [33-36]. Частотный диапазон исследования вибрационных сигналов находился в пределах до 200 Гц. Длительность формирования последовательности доходила до 110 секунд. Частота дискретизации была равна 51200 Гц. Один из результатов восстановления десяти отсчетов сигнала вибрации на пассажирском сиденье над двигателем представлен в таблице 4.

Таблица 4. Результаты восстановления десяти отсчетов сигнала вибрации на пассажирском сиденье над двигателем автобуса марки МАЗ-206067

j	$\nu + j$	$x_{\nu+j}$	p=5, L=20		p=10, L=20	
			$\hat{x}_{\nu+j}$	δ_i^x	$\hat{x}_{\nu+j}$	δ_i^x
1	30286	0.83	0.8285	-0.0028	0,8275	-0.0030
2	30287	0.81	0.8064	-0.0045	0.8069	-0.0038
3	30288	0.88	0.8759	-0.0047	0.8764	-0.0041
4	30290	0.79	0.7867	-0.0042	0.7867	-0.0042
5	30291	0.71	0.7065	-0.0050	0.7069	-0.0043
6	30292	0.72	0.7163	-0.0051	0.7170	-0.0041
7	30293	0.68	0.6767	-0.0049	0.6773	-0.0040
8	30294	0.69	0.6868	-0.0047	0.6874	-0.0038
9	30295	0.72	0.7172	-0.0039	0.7173	-0.0037
10	30296	0.62	0.6172	-0.0045	0.6178	-0.0035

Основываясь на полученных результатах численных экспериментов и примере восстановления реального вибрационного сигнала, можно говорить о том, что разработанный подход позволяет осуществлять восстановление отсчетов на проблемных участках дискретной последовательности сигнала с достаточно низкой погрешностью. В реальных условиях на полезный сигнал могут накладываться внешние аддитивные фоновые шумы и помехи, что приводит к искажению его исходной формы. Это оказывает влияние на ошибку вычисления оценок значений проблемных отсчетов и на выбор

количества отсчетов подпоследовательностей, необходимых для успешного выполнения процедуры восстановления. В общем случае можно считать, что в процессе восстановления происходит оценка и шумовой составляющей в отсчете сигнала. Поэтому по возможности следует принимать во внимание степень коррелированности полезного сигнала и внешних шумов, чтобы в последующем это можно было учесть при цифровой обработке восстановленной последовательности сигнала. При этом следует иметь в виду также то, что разработанный подход предполагает восстановление отсчетов последовательности непрерывного сигнала, который может быть представлен в виде стационарной или квазистационарной модели в пределах интервала времени, на котором осуществляется его дискретизация.

5. Заключение. В статье рассмотрена разработка математического и алгоритмического обеспечения для восстановления значений отсчетов на проблемных участках последовательности равномерно дискретезированного непрерывного во времени сигнала. Предполагается, что сигнал может рассматриваться как стационарный или квазистационарный в пределах интервала его обработки. Разработка осуществлена на основе метода локальной аппроксимации. Спецификой предложенного в настоящей работе подхода является то, что локальная аппроксимация осуществляется по двум подпоследовательностям отсчетов с известными значениями, которые находятся непосредственно перед и после восстанавливаемого участка дискретной последовательности сигнала. Такой подход позволяет получить оценки значений проблемных отсчетов непосредственно с учетом динамики изменения известных значений последовательности сигнала в близких по расположению к ним узлах временной сетки. В качестве аппроксимирующей модели используется ряд Фурье по ортогональной системе тригонометрических функций. Существенной особенностью полученных в настоящей работе математических соотношений для вычисления оценок проблемных отсчетов является то, что они не требуют предварительного вычисления коэффициентов ряда Фурье. Вычислительные процедуры, определяемые этими соотношениями, позволяют непосредственно получать оценки значений восстанавливаемых отсчетов во временной области. Кроме того, при четном числе отсчетов подпоследовательностей L, на основе обработки которых осуществляется восстановление проблемного участка сигнала, и выборе порядка аппроксимирующей модели равным $p = (L/2+1)$, вычислительные процедуры не требует выполнения операций умножения.

Полученные математические соотношения реализованы в виде вычислительных алгоритмов, которые являются неитеративными. Они основаны на выполнении простых арифметических и логических операциях, что снижает общую сложность вычислительных процедур. Показано, что отдельные вычислительные процедуры, входящие в состав алгоритмов, целесообразно выполнять параллельно. На практике эти процедуры могут быть реализованы с применением метода многопоточного программирования. Снижение сложности алгоритмов и использование методов параллельной обработки ведут к повышению вычислительной эффективности процесса восстановления отсчетов.

Исследование возможностей разработанного подхода по восстановлению отсчетов проблемных участков дискретной последовательности сигнала проводились с использованием имитационного моделировании. На примере численных экспериментов было показано, что обеспечивается устойчивое восстановление отсчетов сигнала с достаточно низкой погрешностью. Практическим результатом стала разработка специализированого функционально завершенного программного модуля. Работа данного модуля осуществляется в режиме асинхронного управления вычислительным процессом восстановления отсчетов без блокирования выполнения основной прикладной программы обработки сигнала. Модуль может быть применен в составе метрологически значимого программного обеспечения многофункциональных систем цифровой обработки сигналов.

Дальнейшее выполнение работ, связанных с восстановлением дискретных последовательностей сигналов предполагается проводить с целью обобщения полученных результатов на нестационарные участки сигнала. Это обосновывается тем, что локальная аппроксимация осуществляется на близких интервалах к проблемному участку.

Литература

1. Madisetti V.K. The Digital Signal Processing Handbook, Second edition: Digital Signal Processing Fundamentals // CRC Press, Taylor and Francis Group. 2010. 904 p.
2. Денисенко А.Н. Сигналы. Теоретическая радиотехника. Справочное пособие // М: Горячая линия-Телеком, 2005. 704 с.
3. Oppenheim A.V., Schafer R.W. Discrete-Time Signal Processing: Third edition // Pearson Higher Education. 2010. 1108 p.
4. Поршнев С.В., Кусайкин Д.В. Восстановление неравномерно дискретизированных сигналов с неизвестными значениями координат узлов временной сетки // Успехи современной радиоэлектроники. 2015. №6. С. 3–35.

5. Khan N.A., Ali S. Robust Sparse Reconstruction of Signals with Gapped Missing Samples from Multi-Sensor Recordings // Digital Signal Processing. 2022. vol. 123. 103392.
6. Aceska R., Bouchot J.-L., Li S. Local Sparsity and Recovery of Fusion Frame Structured Signals // Signal Processing. 2020. vol. 174. 107615.
7. Stankovic L., Stankovic S., Amin M. Missing samples analysis in signals for applications to L-estimation and compressive sensing // Signal Processing. 2014. vol. 94. pp. 401–408.
8. Aldroubi A., Leonetti C. Non-Uniform Sampling and Reconstruction from Sampling Sets with Unknown Jitter // Sampling Theory in Signal and Image Processing. 2008. vol. 7. no. 2. pp. 187–195.
9. Nordio A., Chiasserini C-F., Viterbo E. Signal Reconstruction Errors in Jittered Sampling // IEEE Transactions on signal Processing. 2009. vol. 57. no. 12. pp. 4711–4718.
10. Maymon S. Oppenheim A.V. Sinc Interpolation of Nonuniform Samples // IEEE Transactions on Signal Processing. 2011. vol. 59. no. 10. pp. 4745–4758.
11. Andras I., Dolinsky P., Michaeli L., Saliga J. A Time Domain Reconstruction Method of Randomly Sampled Frequency Sparse Signal // Measurement. 2018. vol. 127. pp. 68–77.
12. Якимов В.Н., Машков А.В. Знаковый алгоритм анализа спектра амплитуд и восстановления гармонических составляющих сигналов в условиях присутствия некоррелированных фоновых шумов // Научное приборостроение. 2017. Т. 27. № 2. С. 83–90.
13. Bilinskis I. Digital Alias-free Signal Processing // Wiley. 2007. 454 p.
14. Lu Y. M., Vetterli M. Multichannel Sampling with Unknown Gains and Offsets: A Fast Reconstruction Algorithm // Proc. Allerton Conference on Communication, Control and Computing. Monticello. 2010.
15. Sbaiz L., Vandewalle P., Vetterli M. Groebner Basis Methods for Multichannel Sampling with Unknown Offsets // Applied and Computational Harmonic Analysis. 2008. vol. 25. no. 3. pp. 277–294.
16. Liu N., Tao R., Wang R., Deng Y., Li N., Zhao S. Signal Reconstruction from Recurrent Samples in Fractional Fourier Domain and Its application in Multichannel SAR // Signal Processing. 2017. vol. 131. pp. 288–299.
17. Allen R.L., Mills D.W. Signal Analysis: Time, Frequency, Scale, and Structure // IEEE Press (Wiley-Interscience). 2004. 966 p.
18. Sejdic E., Orovic I., Stankovic S. Compressive sensing meets time-frequency: An overview of recent advances in time-frequency processing of sparse signals // Digital Signal Processing. 2018. vol. 77. pp. 22–35.
19. Teke O., Gurbuz A.C., Arikan O. A robust compressive sensing based technique for reconstruction of sparse radar scenes // Digital Signal Processing. 2014. vol. 27, pp. 23-32.
20. Жукова Н.А., Соколов И.С. Метод восстановления структуры группового телеметрического сигнала на основе графовой модели // Труды СПИИРАН. 2010. Вып. 13. С. 45–66.
21. Khan N.A., Ali S. Reconstruction of gapped missing samples based on instantaneous frequency and instantaneous amplitude estimation // Signal Processing. 2022. vol. 193. no. 108429.
22. Dokuchaev N. On Recovery of Discrete Time Signals from Their Periodic Subsequences // Signal Processing. 2019. vol. 162. pp. 180–188.
23. Annaby M.H., Al-Abdi I.A., Abou-Dina M.S., Ghaleb A.F. Regularized Sampling Reconstruction of Signals in the Linear Canonical Transform Domain // Signal Processing. 2022. vol. 198. 108569.

24. Yue C., Liang J., Qu B., Han Y., Zhu Y., Crisalle O.D. A Novel Multiobjective Optimization Algorithm for Sparse Signal Reconstruction // Signal Processing. 2020. vol. 167. 107292.
25. Wijenayake C., Scutts J., Ignjatovic A. Signal recovery algorithm for 2-level amplitude sampling using chromatic signal approximations // Signal Processing. 2018. vol. 153. pp. 143–152.
26. Катковник В.Я. Непараметрическая идентификация и сглаживание данных: метод локальной аппроксимации // М.: Главная редакция физико-математической литературы. 1985. 336 с.
27. Толстов Г.П. Ряды Фурье // М.: Наука. 1980. 381 с.
28. Brandt S. Data Analysis. Statistical and Computational Methods for Scientists and Engineers // Springer. 2014. 523 p.
29. Bernatz R. Fourier Series and Numerical Methods for Partial Differential Equations // Wiley. 2010. 318 p.
30. Edwards R.E. Fourier Series: A Modern Introduction. vol. 1 //Springer-Verlag. 1979. 224 p.
31. Yakimov V.N., Gorbachev O.V. Firmware of the Amplitude Spectrum Evaluating System for Multicomponent Processes // Instruments and Experimental Techniques. 2013. vol. 56. no. 5. pp. 540–545.
32. Yakimov V.N., Zaberzhinskij B.E., Mashkov A.V., Bukanova Yu.V. Multi-threaded Approach to Software High-speed Algorithms for Spectral Analysis of Multi-component Signals // XXI International Conference Complex Systems: Control and Modeling Problems (CSCMP). 2019. IEEE. pp. 698–701.
33. ГОСТ 31191.1-2004 (ИСО 2631-1:1997) Вибрация и удар. Измерение общей вибрации и оценка ее воздействия на человека. Часть I. Общие требования. Введ. 2008-07-01. М.: Стандартинформ, 2010. 25 с.
34. ГОСТ 55855-2013. Автомобильные транспортные средства. Методы измерения и оценки общей вибрации. Введ. 2014-09-01. М.: Стандартинформ, 2014. 21 с.
35. ГОСТ ИСО 10326-1-2002. Вибрация. Оценка вибрации сидений транспортных средств по результатам лабораторных испытаний. Часть 1. Общие требования. Введ. 2007-11-01. М.: Стандартинформ, 2017. 12 с.
36. ГОСТ ИСО 8002-99. Вибрация. Вибрация наземного транспорта. Представление результатов измерений. Введ. 2001-01-01. Минск: Межгос. совет по стандартизации, метрологии и сертификации, 2000. 16 с.

Якимов Владимир Николаевич — д-р техн. наук, профессор, кафедра автоматизации и управления технологическими процессами, Самарский государственный технический университет. Область научных интересов: методы и средства статистических измерений, диагностика технического состояния и испытания систем, цифровая обработка сигналов, корреляционный и спектральный анализ сигналов, математическое и имитационное моделирование. Число научных публикаций — 295. yvnr@hotmail.com; улица Молодогвардейская, 244, 443100, Самара, Россия; р.т.: +7(846)279-0354.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ (проект № 19-08-00228-а).

DOI 10.15622/ia.21.5.7

V. YAKIMOV

DISCRETE TIME SEQUENCE RECONSTRUCTION OF A SIGNAL BASED ON LOCAL APPROXIMATION USING A FOURIER SERIES BY AN ORTHOGONAL SYSTEM OF TRIGONOMETRIC FUNCTIONS

Yakimov V. **Discrete Time Sequence Reconstruction of a Signal Based on Local Approximation Using a Fourier Series by an Orthogonal System of Trigonometric Functions.**

Abstract. The article considers the development of mathematical and algorithmic support for the sample's reconstruction in problem sections of a discrete sequence of a continuous signal. The work aimed to ensure the reconstruction of lost samples or sections of samples with a non-constant distorted time grid when sampling a signal with a uniform step and at the same time to reduce the computational complexity of digital reconstruction algorithms. The solution to the stated problem is obtained based on the local approximation method. The specific of this method application was the use of two subsequences of samples located symmetrically concerning the reconstructed section of the sequence. The approximating model is a Fourier series on an orthogonal system of trigonometric functions. The optimal solution to the approximation problem is based on the minimum square error criterion. Mathematical equations are obtained for this type of error. They allow us to estimate its value depending on the model order and the samples number in the subsequences used in the reconstruction process. The peculiarity of the mathematical equations obtained in this paper for signal reconstruction is that they do not require the preliminary calculation of the Fourier series coefficients. They provide a direct calculation of the values of reconstructed samples. At the same time, when the number of samples in the subsequences used for reconstruction will be even, it is not necessary to perform multiplication operations. All this made it possible to reduce the computational complexity of the developed algorithm for signal reconstruction. Experimental studies of the algorithm were carried out based on simulation modeling using a signal model that is an additive sum of harmonic components with a random initial phase. Numerical experiments have shown that the developed algorithm provides the reconstruction result of signal samples with a sufficiently low error. The algorithm is implemented as a software module. The operation of the module is carried out on the basis of asynchronous control of the sampling reconstruction process. It can be used as part of metrologically significant software for digital signal processing systems.

Keywords: discrete time signals, sampled sequence, signal reconstruction, local approximation, trigonometric Fourier series.

References

1. Madisetti V.K. The Digital Signal Processing Handbook, Second edition: Digital Signal Processing Fundamentals. CRC Press, Taylor and Francis Group. 2010. 904 p.
2. Denisenko A.N. Signaly. Teoreticheskaya radiotekhnika. Spravochnoye posobiye [Signals. Theoretical radio engineering. Reference manual]. Moscow: Goryachaya Liniya-Telekom. 2005. 704 p. (In Russ.).
3. Oppenheim A.V., Schafer R.W. Discrete-Time Signal Processing: Third edition. Pearson Higher Education. 2010. 1108 p.
4. Porshnev S.V., Kusaykin D.V. [Reconstruction of non-uniform sampled discrete-time signals with unknown sampling locations]. Uspekhi sovremennoy radioelektroniki – Journal Achievements of Modern Radioelectronics. 2015. no. 6. pp. 3–35. (In Russ.).

5. Khan N.A., Ali S. Robust Sparse Reconstruction of Signals with Gapped Missing Samples from Multi-Sensor Recordings. Digital Signal Processing. 2022. vol. 123. 103392.
6. Aceska R., Bouchot J.-L., Li S. Local Sparsity and Recovery of Fusion Frame Structured Signals. Signal Processing. 2020. vol. 174. 107615.
7. Stankovic L., Stankovic S., Amin M. Missing samples analysis in signals for applications to L-estimation and compressive sensing. Signal Processing. 2014. vol. 94. pp. 401–408.
8. Aldroubi A., Leonetti C. Non-Uniform Sampling and Reconstruction from Sampling Sets with Unknown Jitter. Sampling Theory in Signal and Image Processing. 2008. vol. 7. no. 2. pp. 187–195.
9. Nordio A., Chiasserini C-F., Viterbo E. Signal Reconstruction Errors in Jittered Sampling. IEEE Transactions on signal Processing. 2009. vol. 57. no. 12. pp. 4711-4718.
10. Maymon S. Oppenheim A.V. Sinc Interpolation of Nonuniform Samples. IEEE Transactions on Signal Processing. 2011. vol. 59. no. 10. pp. 4745–4758.
11. Andras I., Dolinsky P., Michaeli L., Saliga J. A Time Domain Reconstruction Method of Randomly Sampled Frequency Sparse Signal. Measurement. 2018. vol. 127. pp. 68–77.
12. Yakimov V.N., Mashkov A.V. [The binary algorithm for the analysis of the spectrum amplitude and recover of harmonic components signals in the presence of uncorrelated background noise]. Nauchnoe priborostroenie – Scientific Instrument Making. 2017. vol. 27. no.2. pp. 83–90. (In Russ.).
13. Bilinskis I. Digital Alias-free Signal Processing. Wiley. 2007. 454 p.
14. Lu Y. M., Vetterli M. Multichannel Sampling with Unknown Gains and Offsets: A Fast Reconstruction Algorithm. Proc. Allerton Conference on Communication, Control and Computing. Monticello. 2010.
15. Sbaiz L., Vandewalle P., Vetterli M. Groebner Basis Methods for Multichannel Sampling with Unknown Offsets. Applied and Computational Harmonic Analysis. 2008. vol. 25. no. 3. pp. 277–294.
16. Liu N., Tao R., Wang R., Deng Y., Li N., Zhao S. Signal Reconstruction from Recurrent Samples in Fractional Fourier Domain and Its application in Multichannel SAR. Signal Processing. 2017. vol. 131. pp. 288-299.
17. Allen R.L., Mills D.W. Signal Analysis: Time, Frequency, Scale, and Structure. IEEE Press (Wiley-Interscience). 2004. 966 p.
18. Sejdic E., Orovic I., Stankovic S. Compressive sensing meets time-frequency: An overview of recent advances in time-frequency processing of sparse signals. Digital Signal Processing. 2018. vol. 77. pp. 22–35.
19. Teke O., Gurbuz A.C., Arikan O. A robust compressive sensing based technique for reconstruction of sparse radar scenes. Digital Signal Processing. 2014. vol. 27, pp. 23-32.
20. Zhukova N.A., Sokolov I.S. [Method of reconstructing the structure of the group telemetric signals based on the graph model]. Trudy SPIIRAN – SPIIRAS Proceedings. 2010. no. 13. pp. 45–66. (in Russ.)
21. Khan N.A., Ali S. Reconstruction of gapped missing samples based on instantaneous frequency and instantaneous amplitude estimation. Signal Processing. 2022. vol. 193. no. 108429.
22. Dokuchaev N. On Recovery of Discrete Time Signals from Their Periodic Subsequences. Signal Processing. 2019. vol. 162. pp. 180–188.
23. Annaby M.H., Al-Abdi I.A., Abou-Dina M.S., Ghaleb A.F. Regularized Sampling Reconstruction of Signals in the Linear Canonical Transform Domain. Signal Processing. 2022. vol. 198. 108569.

24. Yue C., Liang J., Qu B., Han Y., Zhu Y., Crisalle O.D. A Novel Multiobjective Optimization Algorithm for Sparse Signal Reconstruction. Signal Processing. 2020. vol. 167. 107292.
25. Wijenayake C., Scutts J., Ignjatovic A. Signal recovery algorithm for 2-level amplitude sampling using chromatic signal approximations. Signal Processing. 2018. vol. 153. pp. 143–152.
26. Katkovnik V.Ya. Neparametricheskaya identifikatsiya i sglazhivaniye dannykh: metod lokal'noy approksimatsii [Nonparametric identification and data smoothing: local approximation method]. Moscow: Glavnaya redaktsiya fiziko-matematicheskoy literatury. 1985. 336 p. (In Russ.).
27. Tolstov G.P. Ryady Fur'ye [Fourier series]. Moscow: Nauka. 1980. 381 p. (In Russ.).
28. Brandt S. Data Analysis. Statistical and Computational Methods for Scientists and Engineers. Springer. 2014. 523 p.
29. Bernatz R. Fourier Series and Numerical Methods for Partial Differential Equations. Wiley. 2010. 318 p.
30. Edwards R.E. Fourier Series: A Modern Introduction. vol. 1. Springer-Verlag. 1979. 224 p.
31. Yakimov V.N., Gorbachev O.V. Firmware of the Amplitude Spectrum Evaluating System for Multicomponent Processes. Instruments and Experimental Techniques. 2013. vol. 56. no. 5. pp. 540–545.
32. Yakimov V.N., Zaberzhinskij B.E., Mashkov A.V., Bukanova Yu.V. Multi-threaded Approach to Software High-speed Algorithms for Spectral Analysis of Multi-component Signals. XXI International Conference Complex Systems: Control and Modeling Problems (CSCMP). 2019. IEEE. pp. 698–701.
33. State Standard 31191.1-2004 (ISO 2631-1:1997). Vibratsiya i udar. Izmereniye obshchey vibratsii i otsenka yeye vozdeystviya na cheloveka. Chast' I. Obshchiye trebovaniya (Vibration and shock. Measurement and evaluation of human exposure to whole-body vibration. Part 1. General requirements), Moscow, Standartinform Publ., 2010. 25 p. (In Russ.).
34. State Standard 55855-2013. Avtomobil'nyye transportnyye sredstva. Metody izmereniya i otsenki obshchey vibratsii (Motor vehicles. Methods of measurement and evaluation of general vibrations), Moscow, Standartinform Publ., 2014, 21 p. (In Russ.).
35. State Standard ISO 10326-1-2002. Vibratsiya. Otsenka vibratsii sideniy transportnykh sredstv po rezul'tatam laboratornykh ispytaniy. Chast' 1. Obshchiye trebovaniya (Vibration. Laboratory method for evaluating vehicle seat vibration. Part 1. Basic requirements), Moscow, Standartinform Publ., 2017, 12 p. (In Russ.).
36. State Standard ISO 8002-99. Vibratsiya. Vibratsiya nazemnogo transporta. Predstavleniye rezul'tatov izmereniy (Mechanical vibration. Land vehicles. Method for reporting measured data), Minsk, Interstate council for standardization, metrology and certification Publ., 2000, 16 p. (In Russ.).

Yakimov Vladimir — Ph.D., Dr.Sci., Professor, Department of automation and control of technological processes, Samara State Technical University. Research interests: methods and means of statistical measurements, diagnostics of technical condition and testing of systems, digital signal processing, correlation and spectral analysis of signals, mathematical and simulation modeling. The number of publications — 295. yvnr@hotmail.com; 244, Molodogvardeyskaya St., 443100, Samara, Russia; office phone: +7(846)279-0354.

Acknowledgements. This research is supported by RFBR (grant 19-08-00228-a).

УДК 004.056:621.391 **DOI** 10.15622/ia.21.5.8

А.С. Гвоздарев, Т.К. Артёмова, П.Е. Патралов, Д.М. Мурин

ВЕРОЯТНОСТНЫЙ АНАЛИЗ БЕЗОПАСНОСТИ БЕСПРОВОДНОЙ СИСТЕМЫ СВЯЗИ ДЛЯ КАНАЛА ТИПА BEAULIEU-XIE С ЗАТЕНЕНИЯМИ

Гвоздарев А.С., Артёмова Т.К., Патралов П.Е., Мурин Д.М. **Вероятностный анализ безопасности беспроводной системы связи для канала типа Beaulieu-Xie с затенениями.**

Аннотация. В работе рассмотрена задача анализа безопасного сеанса на физическом уровне беспроводной системы связи в условиях многолучевого канала распространения сигнала и наличия канала утечки информации. Для обобщения эффектов распространения была выбрана модель канала Beaulieu-Xie с затенениями. Для описания безопасности процесса передачи информации использовалась такая метрика, как вероятность прерывания безопасного сеанса связи. В рамках исследования было получено аналитическое выражение вероятности прерывания связи. Проведён анализ её поведения в зависимости от характеристик канала и системы связи: среднего значения отношения сигнал-шум в основном канале и канале утечки, эффективного значение показателя потерь на пути распространения сигнала, относительного расстояния между законным приемником и прослушивающим приёмником и пороговой пропускной способности, нормированной на пропускную способность гладкого гауссова канала. Рассмотрены совокупности параметров, которые покрывают важные сценарии функционирования беспроводных систем связи. К ним относятся как глубокие замирания (отвечающие гиперрэлеевскому сценарию), так и малые замирания. Учитываются условия наличия существенной по величине компоненты прямой видимости и значительного количества многопутевых кластеров, затенения доминантной компоненты и многопутевость волн, а также всевозможные промежуточные варианты. Обнаружено, что величина энергетического потенциала, необходимого для гарантированной безопасной связи с заданной скоростью, определяется в первую очередь мощностью многопутевых компонент, а также наличие неснижаемой вероятности прерывания безопасного сеанса связи с ростом для каналов с сильным общим затенением компонент сигнала, что с практической точки зрения важно учитывать при предъявлении требований к величинам отношения сигнал/шум и скорости передачи данных в прямом канале, обеспечивающим желаемую степень безопасности беспроводного сеанса связи.

Ключевые слова: беспроводной канал, замирания, затенение, модель Beaulieu-Xie с затенениями, вероятность прерывания безопасного сеанса связи.

1. Введение. Развитие современных стандартов и технологий беспроводной связи неминуемо приводит к увеличению числа стационарных и мобильных устройств, обменивающихся информацией. Объекты сетей Интернета вещей (Internet-of-Things, IoT) или Интернета всего (Internet-of-Everything, IoE) взаимодействуют друг с другом. При этом используются технологии «устройство – устройству» (device to device, D2D), «машина – машине» (machine to machine, M2M), «человек – человеку» (human to human, H2H) [1], «транспорт – транспорту» (vehicle to vehicle, V2V) [2], а также

гибридные технологии, например, «человек – машине» (H2M, [3]) или «транспорт – любому устройству связи» (V2X) [4], позволяющие объединить различные датчики, управляющие элементы и аппаратуру передачи данных, размещённые в домах, на производствах, на транспорте, на элементах городской инфраструктуры, и мобильные, размещённые на теле человека или в элементах одежды, устройства в сети умного дома, умного города, умного жизнеобеспечения и телемедицины. Часто между устройствами при этом передаётся информация, подлежащая защите, в том числе персональные медицинские, финансовые и другие данные. От надёжности и своевременности передачи такой информации могут напрямую зависеть, например, эффективность работы городского транспорта или безопасность человеческой жизни. Поэтому начавшийся переход от индивидуальных решений в области умных домов к повсеместному внедрению умных технологий в последнее время придаёт особую остроту вопросам обеспечения безопасности сеансов связи [5-6].

Альтернативой классическому использованию криптографических методов является подход, основанный на анализе безопасности на физическом уровне [7-9], учитывающий физические явления в канале связи [10]. Такой подход обладает рядом преимуществ перед криптографией [9]. Методы физического уровня позволяют обеспечить безопасный сеанс связи даже в случае, если у подслушивающего устройства имеются мощные вычислительные ресурсы. Они обеспечивают работу сетей различной иерархии, в том числе самоорганизующихся, с узлами, подключающимися и отключающимися от сети в произвольные моменты времени. Эти методы могут выступать в качестве дополнительного хорошо встраиваемого решения обеспечения безопасности связи вместе с традиционными.

Безопасность сеанса беспроводной системы связи на физическом уровне можно охарактеризовать различными показателями [11-13], среди которых важная роль отводится вероятности прерывания безопасного сеанса связи. Она оценивается с помощью предполагаемой статистической модели канала. Следовательно, чем ближе модель к реальным измерениям, тем выше эффективность методов физического уровня.

В настоящее время наблюдается тенденция отдавать предпочтение так называемым обобщенным моделям, учитывающим всё многообразие физических явлений, происходящих при распространении сигналов в системе связи [14-15]. Одной из самых новых в этой группе является модель Больё-Се (Beaulieu-Xie) с

затенениями для канала с замираниями и затенением [16] (далее будем обозначать её как ВХ-модель), которая, с одной стороны, достаточно сложна, чтобы учесть большее количество физических факторов, влияющих на распространение сигнала по сравнению с её аналогами, а с другой – позволяет получать аналитическое описание характеристик в замкнутой форме [17]. Хотя параметры канала в этой модели являются чисто стохастическими, некоторые из них, например, среднее отношение сигнал/шум, можно связать [13] с ослаблением сигнала в беспроводном канале через так называемый показатель потерь при распространении и таким образом учесть влияние физических факторов. Для этой модели на текущий момент существуют результаты, описывающие надёжность сеанса связи (вероятностью его прерывания) и качества связи (усреднённой битовой ошибкой) [16, 18], однако для анализа безопасности сеансов связи она ещё не применялась.

Таким образом, представляет интерес получение аналитического выражения для вероятности прерывания безопасного сеанса системы беспроводной связи при наличии прослушивания и замирания в соответствии ВХ-моделью с затенением. Оно должно связывать данную метрику с параметрами модели и величинами, описывающими влияние физических факторов. При этом в силу того, что ВХ-модель пригодна к описанию различных ситуаций, встречающихся в практике связи, её можно использовать как для основного канала связи, так и для канала утечки.

Целью данной работы является анализ характера поведения зависимости вероятности прерывания безопасного сеанса связи от величин, описывающих воздействие различных влияющих факторов, а также значений этой вероятности, достижимых в наиболее часто встречающихся условиях.

2. Модель системы связи с каналом прослушки. Будем использовать классическую модель системы связи с прослушиванием, предложенную Вайнером [19]. Законный отправитель «А» отправляет сообщение w законному получателю «Б» по дискретному во времени каналу с замираниями (т.е. основному каналу) (рисунок 1). При этом l-символьный блок сообщения w^l кодируется кодерами источника и канала (т.е. «А») в n-символьное кодовое слово $x^n = \{x(1), \ldots, x(i), \ldots, x(n)\}$. Законный получатель «Б» принимает на выходе основного канала n-символьное слово y^n, оценив его символы каким-либо образом, и с помощью декодеров канала и приёмника формирует принятое сообщение $\hat{w}^l$. В соответствии с классическим

подходом [20, 21] будем считать, что передаваемый сигнал также воспринимается пассивным [9] подслушивающим устройством («Е») по каналу прослушивания. В этих условиях предполагается, что канал прослушки не оказывает физического влияния на приём сигнала по основному каналу [13].

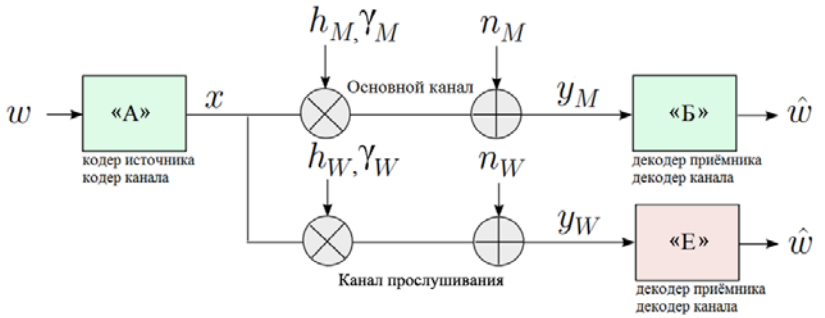


Рис. 1. Модель системы связи с каналом прослушивания

Распространение волн в обоих каналах считается многолучевым, что порождает замирание принимаемого сигнала. Обозначая переменные, соответствующие основному каналу, индексом «*M*», а каналу прослушивания – индексом «*W*», можно формализовать математическое описание системы следующим образом:

$$\begin{cases} y_M(i) = h_M(i)\,x(i) + n_M(i), \\ y_W(i) = h_W(i)\,x(i) + n_W(i), \end{cases} \tag{1}$$

где $x^n = \{x(1), \ldots, x(i), \ldots, x(n)\}$ – кодовое слово, $y_M(i)$, $y_W(i)$ – отсчёты слов, принятых законным получателем и подслушивающим устройством, i – номер отсчёта внутри слова, $n_M(i)$, $n_W(i)$ – комплексные гауссовы шумы с нулевым средним значением в обоих каналах, $h_M(i)$, $h_W(i)$ – стохастически изменяющиеся коэффициенты передачи основного канала и канала прослушивания при передаче i-го отсчёта, учитывающие замирания и затенения. Далее мы рассмотрим типичный случай медленных замираний или, что эквивалентно, каналов, квазистатических на интервале длительности кодового слова, т.е. имеющих $h_M(i) \approx h_M$ и $h_W(i) \approx h_W$ [13].

Для такой модели мгновенное и среднее отношение сигнал/шум (ОСШ) для основного канала и канала прослушивания могут быть определены следующим образом:

$$\begin{cases} \gamma_M(i) = \dfrac{P\left|h_M(i)\right|^2}{\sigma_M^2} = \dfrac{P\left|h_M\right|^2}{\sigma_M^2} = \gamma_M, \\ \gamma_W(i) = \dfrac{P\left|h_W(i)\right|^2}{\sigma_W^2} = \dfrac{P\left|h_W\right|^2}{\sigma_W^2} = \gamma_W, \end{cases} \tag{2}$$

$$\begin{cases} \overline{\gamma}_M(i) = \dfrac{P\mathbb{E}\left\{\left|h_M(i)\right|^2\right\}}{\sigma_M^2} = \dfrac{P\mathbb{E}\left\{\left|h_M\right|^2\right\}}{\sigma_M^2} = \overline{\gamma}_M, \\ \overline{\gamma}_W(i) = \dfrac{P\mathbb{E}\left\{\left|h_W(i)\right|^2\right\}}{\sigma_W^2} = \dfrac{P\mathbb{E}\left\{\left|h_W\right|^2\right\}}{\sigma_W^2} = \overline{\gamma}_W. \end{cases} \tag{3}$$

В выражениях (2)-(3) P обозначает среднюю мощность передаваемого сигнала, σ_M^2, σ_W^2 – мощности шума в каждом из каналов и $\mathbb{E}\{\bullet\}$ – оператор математического усреднения. Знаки равенства в (2)-(3) следуют из предположения о квазистатичности каналов.

Хорошо известно, что [13] наличие замираний позволяет дополнительно использовать информацию о характеристиках канала, оптимизируя стратегию функционирования передатчика и законного приёмника, максимизируя качество передачи (в терминах количества информации, скорости передачи данных, вероятности битовой ошибки [9]) для основного канала и минимизируя его для канала утечки в условиях потенциальной возможности наличия подслушивающего устройства. Это, в свою очередь, придаёт особую важность аналитическим исследованиям в части влияния параметров канала на систему связи.

Общее качество функционирования системы связи с каналом прослушивания будет зависеть от выбранных моделей, используемых для описания основного канала и канала утечки. С одной стороны, с практической точки зрения удобно выбрать наиболее общие модели, позволяющие учесть как можно большее количество эффектов, связанных с распространением сигналов. С другой стороны, важно

сохранить возможность аналитической трактовки выражений для выбранной метрики безопасности связи, для чего удобно использовать для обоих каналов одинаковые модели (с одинаковыми наборами, но с разными значениями параметров). Кроме того, рассмотрим ситуацию, когда законный и прослушивающий приёмники расположены рядом, что соответствует схожести условий распространения сигнала в основном канале и в канале прослушки. Таким образом, для обоих каналов будем использовать одну и ту же модель с одними и теми же параметрами. Основной канал и канал прослушки предполагаются статистически независимыми [13].

3. Модель канала. Одной из особенностей радиоканала является разнообразие условий излучения, распространения и приёма сигнала. Если антенная система передатчика формирует излучение в широком (относительно широком) секторе углов или имеет несколько лепестков диаграммы направленности, а антенная система приёмника способна принимать сигнал с широкого сектора направлений или с нескольких лучей, то создаются условия для попадания сигнала от передатчика к приёмнику несколькими путями. Многолучёвость проявляется как результат многократного переотражения и поглощения на пути распространения радиоволн. При этом в сигнале могут присутствовать, как прямая (доминантная), так и диффузная составляющие, каждая из которых образована множеством компонент с близкими, но всё же отличающимися характеристиками. Многолучевое распространение зависит от типа застройки, рельефа местности, вида растительности и скорости перемещения абонента. Сигналы разных лучей сдвинуты во времени относительно друг друга, что обусловлено различной длиной пути распространения. Компоненты с близкой мощностью, но отличающиеся фазой могут частично или полностью компенсировать друг друга. Этот эффект зависит от частоты. Таким образом, сигнал в радиоканале испытывает замирания. Кроме того, прямая компонента может испытывать частичное или полное затенение, например, вследствие дифракции на растительности или прохождении через элементы строительных конструкций или участки трассы, отличающиеся погодными условиями.

Коэффициент передачи канала на каждом из путей определяется длиной пути, а также условиями рассеяния и поглощения элементами застройки, рельефа и другими отражателями и препятствиями. Так как длины путей, а также коэффициенты отражения и прохождения, описывающие взаимодействие волн с объектами, расположенными на пути распространения сигнала, отличаются, то коэффициент передачи

канала является случайной величиной. Его статистические свойства определяются комбинацией физических условий распространения сигнала.

Будем использовать для основного канала и канала прослушки недавно предложенную обобщенную BX-модель с замираниями и затенениями [16].

В рамках рассматриваемой модели сигнал распространяется в виде кластеров электромагнитных волн, в которых выделяются доминантные и многопутевые компоненты (рисунок 2).

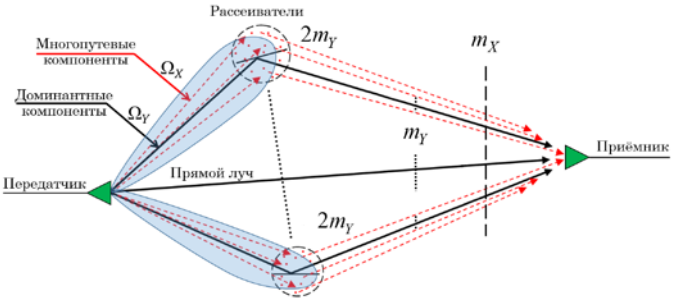


Рис. 2. Физическая ситуация, описываемая моделью канала

Модель канала параметризована в терминах средней мощности многопутевых (рассеянных) компонент Ω_X , средней мощности компонент прямой видимости Ω_Y , степени затенения компонент прямой видимости m_Y и общей степени затенения компонент m_X . Функция плотности вероятности (ФПВ) мгновенного отношения сигнал/шум для обоих каналов (основного и прослушиваемого) задаётся выражением (4) [16]:

$$f_{\gamma_i}(z_i) = \frac{e^{-\frac{m_X(\Omega_X+\Omega_Y)}{\overline{\gamma}_i\Omega_X}z_i}}{\Gamma(m_X)\sqrt{\frac{z_i\overline{\gamma}_i}{\Omega_X+\Omega_Y}}}\left(\frac{m_X}{\Omega_X}\right)^{m_X}\times$$
$$\times\left(\frac{z_i(\Omega_X+\Omega_Y)}{\overline{\gamma}_i}\right)^{m_X-\frac{1}{2}}\left(\frac{m_Y\Omega_X}{m_Y\Omega_X+m_X\Omega_Y}\right)^{m_Y}\times \quad (4)$$
$$\times {}_1F_1\left(m_Y; m_X; \frac{m_X^2(\Omega_X+\Omega_Y)\Omega_Y}{\overline{\gamma}_i\Omega_X(m_Y\Omega_X+m_X\Omega_Y)}z_i\right),$$

где $i=\{1;2\}$, $\overline{\gamma}_i=\begin{cases}\overline{\gamma}_M, & i=1,\\ \overline{\gamma}_W, & i=2\end{cases}$, $\overline{\gamma}_M$, $\overline{\gamma}_W$ обозначают средние значения отношений сигнал-шум для основного канала и канала прослушки соответственно, ${}_1F_1(\bullet)$ – вырожденная гипергеометрическая функция [22], а $\Gamma(\bullet)$ – гамма-функция Эйлера [22].

Следует отметить, что такая модель в нескольких конкретных случаях переходит в другие модели, в том числе широкораспространённые упрощённые: Райса, Рэлея, Накагами-*m*, BX-модель без затенения [23], в $\kappa-\mu$ с затенениями [17], а также в ряд других упрощённых моделей [16]. Более того, модель описывает, в том числе и те случаи, которые не охватываются классическими моделями, так как параметры интенсивности затенения m_X и m_Y могут принимать произвольные положительные значения (т.е. они не ограничены снизу значением 0,5, как это широко принято для упрощённых моделей), а значение $2m_Y$ равно числу доминантных компонентов [16].

Будем рассматривать ситуацию с одинаковыми параметрами затухания для основного канала и канала прослушивания, что физически соответствует сценарию, когда подслушивающий и законный приемник находятся на близком расстоянии в пределах зоны действия базовой станции (передатчика).

Можно заметить, что, хотя параметры модели Ω_X, Ω_Y, m_Y, m_X связаны с некоторыми физическими величинами, они являются статистическими. В отличие от них средние отношения сигнал-шум обычно связаны с расстояниями до «Б» (d_M) и «Е» (d_W) через так называемый показатель потерь на трассе α [24], т.е. $\overline{\gamma}_M \sim 1/d_M^{\alpha}$, $\overline{\gamma}_W \sim 1/d_W^{\alpha}$ [13]. Несмотря на то, что α является стохастическим, он характеризуется своим эффективным значением, которое описывает возможную среду распространения сигнала (например, $\alpha=2$ для распространения в свободном пространстве без замираний [24]). Величиной α можно управлять, например, с помощью интеллектуальных рассеивающих панелей [25-26].

4. Показатель качества безопасного функционирования беспроводной системы связи. Будем характеризовать качество функционирования описанной беспроводной системы связи вероятностью прерывания безопасного сеанса связи P_{out} (Secrecy

Outage Probability (англ.) – SOP). Она определяется как вероятность того, что пропускная способность C будет ниже порогового значения C_{th} [13], т.е.:

$$P_{out}\left(C_{th}\right)=\mathbb{P}\left(C<C_{th}\right)=\mathbb{P}\left(\gamma_M<\left(1+\gamma_W\right)2^{C_{th}}-1\right)= \\ =\int_0^{\infty}\int_0^{(1+\gamma_W)2^{C_{th}}-1} f_{\gamma_M,\gamma_W}\left(z_1,z_2\right)\mathrm{d}z_1\mathrm{d}z_2. \tag{5}$$

5. Полученное выражение для вероятности прерывания безопасного сеанса связи. В предположении о том, что основной канал и канал прослушивания статистически независимы, функция их совместной плотности распределения вероятностей мгновенных ОСШ может быть факторизована в виде произведения одномерных функций плотностей вероятности (4), т.е. $f_{\gamma_M,\gamma_W}\left(z_1,z_2\right)=f_{\gamma_M}\left(z_1\right)f_{\gamma_W}\left(z_2\right)$. В результате выражение (5) можно представить в виде:

$$P_{out}\left(C_{th}\right)=\int_0^{\infty}\int_0^{(1+z_2)2^{C_{th}}-1} f_{\gamma_M}\left(z_1\right)f_{\gamma_W}\left(z_2\right)\mathrm{d}z_1\mathrm{d}z_2= \\ =\frac{\left(\dfrac{m_Y\Omega_X}{m_Y\Omega_X+m_X\Omega_Y}\right)^{2m_Y}\left(\dfrac{m_X}{\Omega_X}\left(\Omega_X+\Omega_Y\right)\right)^{2m_X}}{\Gamma^2\left(m_X\right)\overline{\gamma}_M^{m_X}\overline{\gamma}_W^{m_X}}\times \\ \times\int_0^{\infty}\int_0^{(1+z_2)2^{C_{th}}-1} z_2^{m_X-1}z_1^{m_X-1}e^{-\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_W\Omega_X}z_2}e^{-\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X}z_1}\times \\ \times {}_1F_1\left(m_Y;m_X;\frac{m_X^2\left(\Omega_X+\Omega_Y\right)\Omega_Y}{\overline{\gamma}_W\Omega_X\left(m_Y\Omega_X+m_X\Omega_Y\right)}z_2\right)\times \\ \times {}_1F_1\left(m_Y;m_X;\frac{m_X^2\left(\Omega_X+\Omega_Y\right)\Omega_Y}{\overline{\gamma}_M\Omega_X\left(m_Y\Omega_X+m_X\Omega_Y\right)}z_1\right)\mathrm{d}z_1\mathrm{d}z_2. \tag{6}$$

Воспользуемся классическим разложением для вырожденной гипергеометрической функций ${}_1F_1\left(\bullet\right)$ в ряд ([22], выражение 13.2.2):

$$
{}_1F_1\left(m_Y; m_X; \frac{m_X^2\left(\Omega_X+\Omega_Y\right)\Omega_Y}{\overline{\gamma}_i\Omega_X\left(m_Y\Omega_X+m_X\Omega_Y\right)} z_i \right) =
$$

$$
= \sum_{p=0}^{\infty} \frac{\left(m_Y\right)_p}{\left(m_X\right)_p} \frac{\left(\frac{m_X^2\left(\Omega_X+\Omega_Y\right)\Omega_Y}{\overline{\gamma}_i\Omega_X\left(m_Y\Omega_X+m_X\Omega_Y\right)} z_i \right)^p}{p!}, \tag{7}
$$

где $\left(\bullet\right)_p$ – символ Похгаммера [22].

Подставляя разложение (7) в выражение (6) и перегруппировывая слагаемые, получим:

$$
P_{out}\left(C_{th}\right) = \frac{\left(\frac{m_Y\Omega_X}{m_Y\Omega_X+m_X\Omega_Y} \right)^{2m_Y} \left(\frac{m_X}{\Omega_X}\left(\Omega_X+\Omega_Y\right) \right)^{2m_X}}{\Gamma^2\left(m_X\right)\overline{\gamma}_M^{m_X}\overline{\gamma}_W^{m_X}} \times
$$

$$
\times \sum_{p=0}^{\infty}\sum_{q=0}^{\infty} \frac{\left(m_Y\right)_q}{\left(m_X\right)_q} \frac{\left(m_Y\right)_p}{\left(m_X\right)_p} \frac{\left(\frac{m_X^2\left(\Omega_X+\Omega_Y\right)\Omega_Y}{\Omega_X\left(m_Y\Omega_X+m_X\Omega_Y\right)} \right)^{p+q}}{q!\,p!\,\overline{\gamma}_W^p\overline{\gamma}_M^q} \times \tag{8}
$$

$$
\times \int_0^{\infty} z_2^{m_X+p-1} e^{-\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_W\Omega_X}z_2} \left\{ \int_0^{\left(1+z_2\right)2^{C_{th}}-1} z_1^{m_X+q-1} e^{-\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X}z_1} \mathrm{d}z_1 \right\} \mathrm{d}z_2 .
$$

Заметим, что внутренний определенный интеграл сводится к неполной нижней гамма-функции $\tilde{\gamma}\left(a,z\right)$ ([22], выражение 8.2.1):

$$
\int_0^{\left(1+z_2\right)2^{C_{th}}-1} z_1^{m_X+q-1} e^{-\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X}z_1} \mathrm{d}z_1 =
$$

$$
= \left(\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X} \right)^{-m_X-q} \tilde{\gamma}\left(m_X+q, \left(\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X} \right)\left(\left(1+z_2\right)2^{C_{th}}-1\right) \right). \tag{9}
$$

Учтем, что для $\tilde{\gamma}\left(a,z\right)$ существует разложение в виде конечной суммы ([22], выражение 8.4.7):

$$\tilde{\gamma}\left(m_X+q,\left(\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X}\right)\left(\left(1+z_2\right)2^{C_{th}}-1\right)\right)=$$

$$=\Gamma\left(m_X+q\right)\left(1-e^{-\left(\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X}\right)\left(\left(1+z_2\right)2^{C_{th}}-1\right)}\sum_{k=0}^{m_X+q-1}\frac{\left[\left(\left(1+z_2\right)2^{C_{th}}-1\right)\right]^k}{k!\left(\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X}\right)^{-k}}\right). \quad (10)$$

После подстановки (10) в (7) выражение для вероятности прерывания безопасного сеанса связи распадается на 2 слагаемых. Отметим, что в первом слагаемом двойной ряд распадается на произведение двух одинарных, для каждого из которых результат суммирования может быть получен в замкнутой форме путем сведения каждого из них к выражению для гипергеометрической функции Гаусса ${}_2F_1\left(m_Y,b;b;z\right)$ с совпадающими аргументами ([22] выражение 15.4.6):

$$\sum_{p=0}^{\infty}\frac{\left(m_Y\right)_p}{p!}\left(\frac{m_X\Omega_Y}{\left(m_Y\Omega_X+m_X\Omega_Y\right)}\right)^p=\sum_{p=0}^{\infty}\frac{\left(m_Y\right)_p\left(b\right)_p}{\left(b\right)_p}\frac{\left(\frac{m_X\Omega_Y}{\left(m_Y\Omega_X+m_X\Omega_Y\right)}\right)^p}{p!}$$

$$={}_2F_1\left(m_Y,b;b;\frac{m_X\Omega_Y}{\left(m_Y\Omega_X+m_X\Omega_Y\right)}\right)=\left(1-\frac{m_X\Omega_Y}{\left(m_Y\Omega_X+m_X\Omega_Y\right)}\right)^{-m_Y}. \quad (11)$$

Учитывая тождество:

$$\frac{1}{\left(m_X\right)_p}\Gamma\left(m_X+p\right)=\Gamma\left(m_X\right), \quad (12)$$

(вытекающее из определения символа Похгаммера), а также тот факт, что оставшийся несобственный интеграл в выражении (8) может быть вычислен с использованием вырожденной гипергеометрической функции Трикоми $U\left(a;b;z\right)$ ([22] выражение 13.4.4):

$$\int_0^\infty z_2^{m_X+p-1}\frac{\left(\left(1+z_2\right)2^{C_{th}}-1\right)^k}{\left(\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X}\right)^{-k}}e^{-\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\Omega_X}\left(\frac{1}{\overline{\gamma}_W}+2^{C_{th}}\frac{1}{\overline{\gamma}_M}\right)z_2}\mathrm{d}z_2=$$

$$=\left[\left(\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X}\right)\left(2^{C_{th}}-1\right)\right]^k\left(\frac{2^{C_{th}}-1}{2^{C_{th}}}\right)^{m_X+p}\Gamma\left(m_X+p\right)\times \quad (13)$$

$$\times U\left(m_X+p;m_X+p+1+k;\left(\frac{2^{C_{th}}-1}{2^{C_{th}}}\right)\left(\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\Omega_X}\left(\frac{1}{\overline{\gamma}_W}+2^{C_{th}}\frac{1}{\overline{\gamma}_M}\right)\right)\right),$$

итоговое выражение для вероятности прерывания безопасного сеанса связи представимо в виде:

$$P_{out}\left(C_{th}\right)=1-\frac{\left(\frac{2^{C_{th}}-1}{2^{C_{th}}}\right)^{m_X}\left(\frac{m_Y\Omega_X}{m_Y\Omega_X+m_X\Omega_Y}\right)^{2m_Y}}{\left(\frac{m_X}{\overline{\gamma}_W\Omega_X}\left(\Omega_X+\Omega_Y\right)\right)^{-m_X}}e^{-\frac{m_X}{\overline{\gamma}_W\Omega_X}\left(\Omega_X+\Omega_Y\right)\left(2^{C_{th}}-1\right)}\times$$

$$\times\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\frac{\left(m_Y\right)_q\left(m_Y\right)_p}{q!\,p!\,\overline{\gamma}_W^p\overline{\gamma}_M^q}\frac{\left(\frac{m_X^2\left(\Omega_X+\Omega_Y\right)\Omega_Y}{\overline{\gamma}_W\Omega_X\left(m_Y\Omega_X+m_X\Omega_Y\right)}\right)^p}{\left(\frac{2^{C_{th}}-1}{2^{C_{th}}}\right)^{-p}}\left(\frac{m_X\Omega_Y}{m_Y\Omega_X+m_X\Omega_Y}\right)^q\times \quad (14)$$

$$\times\sum_{k=0}^{m_X+q-1}\frac{\left[\left(\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\overline{\gamma}_M\Omega_X}\right)\left(2^{C_{th}}-1\right)\right]^k}{k!}\times$$

$$\times U\left(m_X+p;m_X+p+1+k;\left(\frac{2^{C_{th}}-1}{2^{C_{th}}}\right)\left(\frac{m_X\left(\Omega_X+\Omega_Y\right)}{\Omega_X}\left(\frac{1}{\overline{\gamma}_W}+2^{C_{th}}\frac{1}{\overline{\gamma}_M}\right)\right)\right).$$

Насколько известно авторам, полученное выражение (7) является новым.

Стоит отметить несколько фактов:

– Все используемые специальные функции являются классическими и программно реализованы во всех современных пакетах компьютерной алгебры (MatLAB, MathCAD, Mathematica, Maple и др.), а также в виде отдельных свободно распространяемых библиотек (например, пакет Special functions (scipy.special) для языка Python).

– Несмотря на то, что полученное итоговое выражение (14) представляется в виде двойного ряда, на практике только первых нескольких членов оказывается достаточно для достижения требуемой точности, что позволяет: 1) повысить скорость вычисления вероятности прерывания беспроводного сеанса связи по сравнению с исходным выражением (6) и 2) избежать сложностей, связанных с выбором метода численного интегрирования в выражении (6), в частности с необходимостью обеспечения компромисса между сходимостью интеграла, точностью проведения операции интегрирования (чувствительности выбранной схемы интегрирования к параметрам канала) и скоростью вычисления.

– Для дальнейшего повышения скорости вычисления для выражения (14) могут быть успешно использованы алгоритмы повышения скорости сходимости рядов, например, метод Шенкса [27].

Выражение (14) позволяют анализировать поведение вероятности прерывания безопасного сеанса связи в зависимости от параметров модели канала, а также от относительного расстояния между законным и подслушивающим приёмниками.

6. Результаты численной апробации. Чтобы продемонстрировать правильность полученных выражений и изучить эффекты, вызванные изменением канала распространения, было проведено численное моделирование. Исследовались как общее поведение вероятности прерывания безопасного сеанса связи с ростом среднего отношения сигнал/шум в основном канале для различных часто встречающихся на практике случаев, так и характер и величина влияния на эту вероятность различных факторов, которые позволяет учесть модель.

6.1. Параметры моделирования. Численное моделирование производилось в системе Wolfram Mathematica. Исследовалась вероятность прерывания безопасного сеанса связи (SOP), заданная выражениями (6)-(7).

Для моделирования была выбрана совокупность параметров, которая покрывает все практически важные сценарии функционирования: как глубокие замирания (малые значения m_X и m_Y, отвечающие гиперрэлеевскому сценарию), так и малые замирания

(большие значения m_X и m_Y), как в случае наличия существенной по величине компоненты прямой видимости и значительном количестве многопутевых кластеров (при больших значениях Ω_Y и Ω_X), так и при существенных затенениях доминантной компоненты и малого числа многопутевых волн, а также всевозможные промежуточные варианты.

Так как ВХ-модель является пригодной для описания разнообразных сценариев связи, то исследования проводились при параметрах, изменяющихся в широком диапазоне значений. Параметры модели приведены в таблице 1. При построении характеристик, в соответствии с теорией, описанной в работе [13], пропускная способность БСС нормировалась на пропускную способность Гауссовского канала связи без замираний и затенений, которая описывается классической формулой Шеннона [28]. В дальнейших упоминаниях в работе нормированная пропускная способность БСС будет называться просто пропускной способностью безопасного сеанса связи.

Таблица 1. Параметры модели

Название	Обозначение	Значения
Средняя мощность многопутевых компонент	Ω_X	-10..20 дБ
Средняя мощность компонент прямой видимости	Ω_Y	-10..20 дБ
Степень затенения компонент прямой видимости	m_Y	0,1..5
Общая степень затенения компонент	m_X	0,1..5
Показатель потерь на трассе	α	2 или 3
Отношение расстояний от передатчика до законного приёмника и до подслушивающего устройства	d_W / d_M	0,1..100
Среднее значение отношения сигнал-шум для основного канала связи	$\overline{\gamma}_M$	0..50 дБ
Среднее значение отношения сигнал-шум для канала прослушки	$\overline{\gamma}_W$	0..50 дБ
Пороговая пропускная способность системы связи, гарантирующая безопасную связь, нормированная на пропускную способность Гауссовского канала связи без замираний и затенений	C_{th}	0,1..0,6

Значение показателя потерь на трассе согласно [24] соответствовало ситуации распространения сигнала в условиях

городской местности с элементами инфраструктуры, создающими затенение.

6.2. Общие свойства. Влияние параметров канала. На рисунке 3 приведены зависимости вероятности прерывания безопасного сеанса связи от среднего ОСШ основного канала для ситуации канала с малой мощностью компонент и прямой видимости, и многопутевых, и сильным затенением, при показателе потерь на трассе $\alpha = 2$. Это практически гиперрэлеевский канал с глубокими замираниями [29]. В качестве параметра здесь использовано среднее значение ОСШ в канале прослушки (для значений, соответствующих -10, 0, 10, 20 дБ использованы линии с квадратными, круглыми, треугольными маркерами и маркерами в виде косых крестиков соответственно). Семейство линий, полученных при большой нормированной пороговой пропускной способности БСС (0,9), проходит существенно выше, чем семейство, полученное при малой C_{th} (0,1), и демонстрирует в отличие от них слабую зависимость от ОСШ в основном канале.

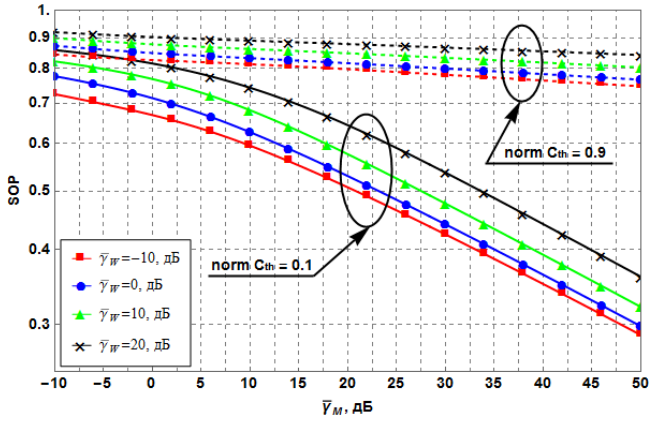


Рис. 3. Зависимость вероятности прерывания БСС от среднего ОСШ основного канала $\bar{\gamma}_M$ для канала с малой мощностью обеих компонент ($\Omega_X = \Omega_Y = -10$ дБ) и сильным затенением ($m_X = m_Y = 0{,}1$) при изменении среднего ОСШ в канале прослушки от -10 до 10 дБ и нормированной пороговой пропускной способности БСС C_{th} от малой (0,1) до большой (0,9)

Из рисунка 3 видно, что если допускать снижение скорости связи ниже заранее предопределённого значения C_{th}, равного 0,1, с вероятностью, например, 0,3, то для обеспечения БСС в таких

условиях необходимы очень большие значения среднего ОСШ законного приёмника (47-50 дБ), что с точки зрения практики очень тяжело реализовать. Что касается работы при высоких скоростях ($C_{th} = 0,9$), то при том же значительном ОСШ в основном канале 50 дБ значения вероятности прерывания лежат в диапазоне от 0,75 до 0,85, то есть, SOP = 0,3 недостижима при высоких скоростях и высокоскоростная передача информации по беспроводному каналу с большой долей вероятности будет небезопасной.

Как видно из рисунка 3 и последующих, рост ОСШ в канале утечки ухудшает ситуацию для всех значений остальных параметров модели, за исключением отношения расстояний до законного и подслушивающего приёмников (что будет рассмотрено отдельно), причём тем сильнее, чем выше $\overline{\gamma}_W$. Это объясняется тем, что при высоком ОСШ подслушивающему приёмнику легче и обнаружить передачу информации, и декодировать сигнал. В целом, для большинства систем передачи информации связь через канал, описываемый гиперрэлеевской моделью с глубокими замираниями, является небезопасной.

Серия рисунок 4-6 демонстрирует влияние условий распространения, т.е. параметров модели канала, на условия, при которых достигается заданная вероятность прерывания безопасного сеанса связи. В них приведены результаты для $C_{th} = 0,3$, а мощность либо компонент прямой видимости, либо многопутевых, либо и тех, и других, велика (что существенным образом отличает условия распространения от тех, для которых получены зависимости на рисунке 3). Основное отличие вероятности прерывания для ситуаций на рисунке 4-6 по сравнению со сценариями на рисунке 3 заключается в существенном уменьшении значений и в характерной форме зависимостей, обладающих «полочкой» на уровне высоких вероятностей прерывания при малых ОСШ в основном канале, завершающейся резким спадом, переходящим в плато неснижаемых малых значений вероятности при больших ОСШ. В целом зависимости смещаются в область меньших $\overline{\gamma}_M$ и достигают меньших значений на плато с уменьшением ОСШ в канале утечки, степени общего затенения m_X (рисунок 4), мощности компонент прямой видимости (рисунки 5 и 6).

На рисунке 4 зависимости вероятности прерывания БСС от ОСШ в основном и подслушивающем каналах приведены для ситуации, когда скорость связи ограничена нормированной пропускной способностью 0,3, а канал описывается моделью со

слабым затенением компонент прямой видимости ($m_Y = 5$) в отличие от многопутевых компонент и сильным прямым лучом ($\Omega_X = -10$ дБ, $\Omega_Y = 10$ дБ).

Наличие «полочки» при значениях SOP, близких к 1, свидетельствует о независимости вероятности прерывания безопасного сеанса связи от общей степени затенения компонент прямой видимости и многопутевых компонент. «Полочка» наблюдается в ситуациях, когда подслушивающее устройство имеет лучшее $\overline{\gamma}_W$, чем законный приёмник, т.е. $\overline{\gamma}_M < \overline{\gamma}_W$, например, при $\overline{\gamma}_M < 10$ дБ для линии, соответствующей $\overline{\gamma}_W = 10$ дБ.

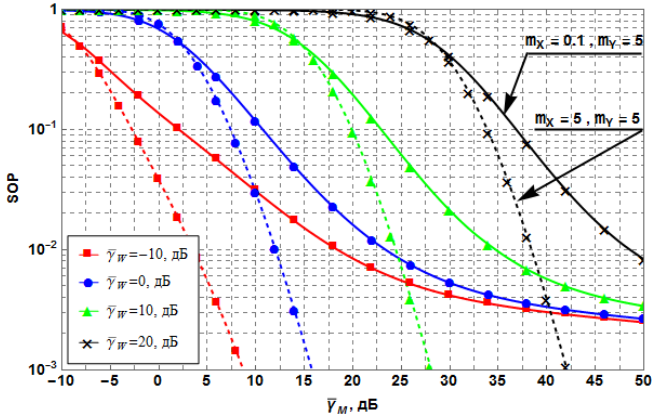


Рис. 4. Влияние общего затенения m_X на вероятность прерывания БСС для средней нормированной пропускной способности ($C_{th} = 0{,}3$) в канале со слабым затенением компонент прямой видимости ($m_Y = 5$) и сильным прямым лучом ($\Omega_X = -10$ дБ, $\Omega_Y = 10$ дБ)

Если же рассматривать случай, когда $\overline{\gamma}_M > \overline{\gamma}_W$, то при уменьшении общей степени затенения для обоих каналов сеанс связи станет более безопасным. С ростом $\overline{\gamma}_W$ улучшение вероятности прерывания БСС снижается. Например, при помеховой ситуации в основном канале, на 10 дБ лучшей, чем в канале утечки, при $\overline{\gamma}_W = -10$ дБ происходит снижение SOP с 0,15 до 0,04 (т.е. в 3,75 раз), а при $\overline{\gamma}_W = 10$ дБ – с 0,4 до 0,35 (т.е. в 1,14 раза), что практически нивелирует эффект. При большей разнице в ОСШ основного и

подслушивающего каналов эффект улучшения вероятности прерывания БСС при уменьшении общей степени затенения является более выраженным. Например, пусть подслушивающее устройство имеет среднее ОСШ $\overline{\gamma}_W = 10$ дБ, тогда при $\overline{\gamma}_M = 25$ дБ вероятность прерывания БСС при сильном затенении будет равна 0,07, а при слабой степени затенения – 0,007, то есть изменение составляет целый порядок.

Наличие неснижаемой вероятности прерывания БСС у зависимостей для каналов с сильным общим затенением компонент сигнала говорит о том, что при заданном количестве компонент прямой видимости ($2m_Y = 10$), нельзя будет снизить SOP никакими способами, связанными с повышением ОСШ в основном канале.

Стоит отметить, что чем большее ОСШ у подслушивающего устройства, тем большими значениями ОСШ в основном канале будут достигаться заданные значения вероятности прерывания БСС при заданных скоростях передачи. Например, зафиксируем вероятность прерывания безопасного сеанса связи на уровне 10^{-2} при среднем ОСШ подслушивающего устройства $\overline{\gamma}_W = -10$ дБ, тогда при сильных затенениях, то есть при $m_X = 0{,}1$, для обеспечения заданной SOP потребуется обеспечить $\overline{\gamma}_M = 18$ дБ, а в ситуации со слабыми затенениями достаточно будет 4 дБ. При улучшении помеховой обстановки в канале прослушки этот выигрыш в ОСШ законного приёмника уменьшается. Так, при вероятности прерывания безопасного сеанса связи 10^{-2} выигрыш равен 14, 11, 10 и 9 дБ для $\overline{\gamma}_W$, равного -10, 0, 10 и 20 дБ соответственно.

Из всего вышесказанного можно сделать вывод: для обеспечения достаточно малой вероятности прерывания БСС при большом количестве сильных доминантных компонент и малой мощности рассеянных компонент выгоднее будет канал связи со слабой степенью затенения всех компонент.

На рисунке 5 показана ситуация канала с невысокой степенью общего затенения ($m_X = 5$), в котором основной вклад в принятый сигнал вносят многопутевые компоненты, полученные за счёт множественных переотражений. При уменьшении параметра Ω_Y с 10 дБ до -10 дБ компонента прямой видимости для обоих каналов связи практически исчезает, а безопасность сеанса связи повышается. Наличие же прямой видимости в двух каналах одновременно приводит к тому, что безопасность сеанса связи снижается. Например, для обеспечения заданного порога скорости $C_{th} = 0{,}3$ с вероятностью

прерывания БСС 10^{-4} при $\overline{\gamma}_W = 0$ дБ и прямой видимости необходимо $\overline{\gamma}_M = 29$ дБ, а в ситуации, когда компонента прямой видимости практически исчезает, – всего 19 дБ.

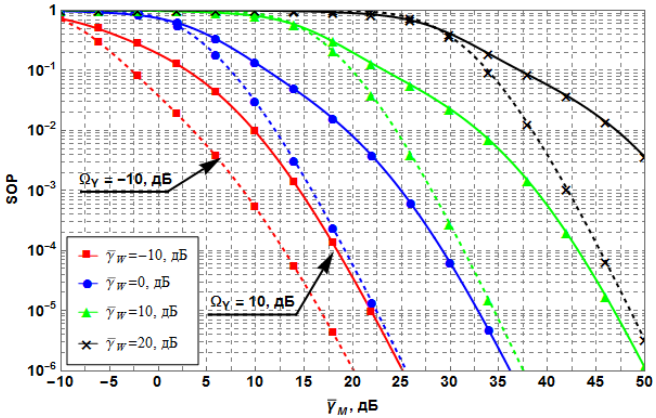


Рис. 5. Влияние средней мощности Ω_Y компонент прямой видимости на вероятность прерывания БСС для $C_{th} = 0{,}3$ в канале с сильным затенением компонент прямой видимости ($m_Y = 0{,}1$), слабым общим затенением ($m_X = 5$) и высокой мощностью многопутевых компонент ($\Omega_X = 10$ дБ)

Как показано на рисунке 6, в таком канале существенным фактором, влияющим на безопасность связи, является мощность многопутевых компонент.

Во-первых, она определяет характер поведения кривых вероятности БСС. Видно, что при $\Omega_X = -10$ дБ кривые для $\overline{\gamma}_W = -10$ дБ и $\overline{\gamma}_W = 0$ дБ практически сливаются друг с другом, хотя при $\Omega_X = 10$ дБ они сильно расходятся, на уровне SOP = 10^{-3} расхождение составляет примерно 10 дБ

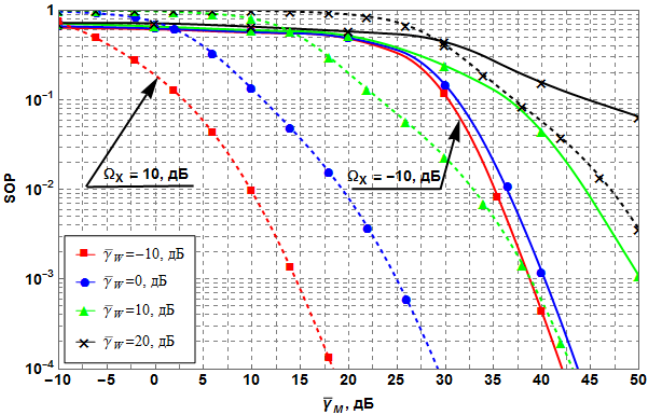


Рис. 6. Влияние средней мощности Ω_X многопутевых компонент на вероятность прерывания БСС для $C_{th} = 0{,}3$ в канале с сильным затенением компонент прямой видимости ($m_Y = 0{,}1$), слабым общим затенением ($m_X = 5$) и высокой мощностью компонент прямой видимости ($\Omega_Y = 10$ дБ)

Во-вторых, мощность многопутевых компонент определяет величину энергетического потенциала, необходимого для гарантированно безопасной связи с заданной скоростью. При этом рисунок 6 даёт представление о достижимых возможностях системы связи в ситуации, когда мощность многопутевых компонент мала, а компоненты прямой видимости испытывают сильное затенение.

Так, из рисунка видно, что для обеспечения достаточно низкой вероятности прерывания БСС, например, на уровне 10^{-3}, потребуется значительное ОСШ законного приёмника. Если ситуации с плохой помеховой обстановкой в канале утечки ($\bar{\gamma}_W = -10$ дБ) необходимо $\bar{\gamma}_M = 37$ дБ, что с токи зрения практики является сложно реализуемым, то при условии низких ($\bar{\gamma}_W = 10$ дБ) и очень низких ($\bar{\gamma}_W = 20$ дБ) помех на входе подслушивающего устройства необходимые ОСШ начинаются с 50 дБ, т.е. нереалистичны, поэтому обеспечить вероятность прерывания БСС на заданном уровне не представляется возможным.

6.3. Требования к скорости передачи. На рисунках 7-9 приведены зависимости вероятности прерывания безопасного сеанса связи от нормированной пороговой пропускной способности, ограничивающей скорость передачи информации, при фиксированном среднем ОСШ основного канала $\bar{\gamma}_M = 10$ дБ. В качестве параметра

здесь использовано среднее значение ОСШ в канале прослушки (для значений, соответствующих -10, 0, 10, 20 дБ использованы линии с квадратными, круглыми, треугольными маркерами и маркерами в виде косых крестиков соответственно). Анализируется влияние требований к скорости передачи для различных условий распространения, отображаемых наборами параметров ВХ модели канала, при показателе потерь на трассе $\alpha = 2$.

Для канала с одной компонентой прямой видимости, мощными многопутевыми компонентами и слабым общим затенением при фиксированном среднем ОСШ основного канала $\overline{\gamma}_M = 10$ дБ и различных средних мощностях компоненты прямой видимости Ω_Y зависимости вероятности прерывания БСС от пороговой пропускной способности БСС в канале приведены на рисунке 7.

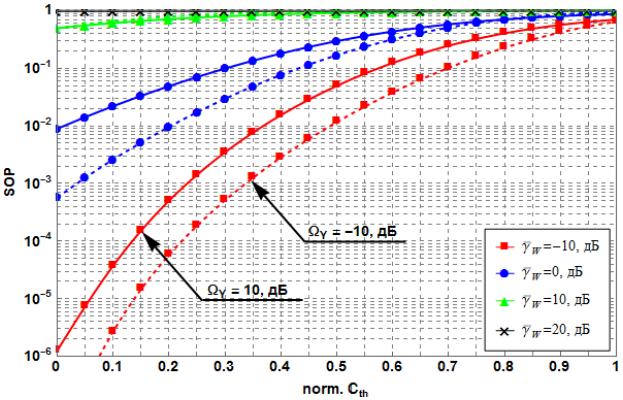


Рис. 7. Зависимость вероятности прерывания БСС от пороговой пропускной способности БСС в канале с одной компонентой прямой видимости, мощными многопутевыми компонентами и слабым общим затенением при фиксированном среднем ОСШ основного канала $\overline{\gamma}_M = 10$ дБ для высокой ($\Omega_Y = 10$ дБ) и низкой ($\Omega_Y = -10$ дБ) средней мощности компонент прямой видимости

Ситуацию, представленную на рисунке 7, можно трактовать как соответствующую каналу с замираниями Райса, в котором есть одна компонента прямой видимости ($m_Y = 0,5$) и наблюдается слабое общее затенение ($m_X = 5$). При этом отношение $K = \Omega_Y / \Omega_X$ определяет фактор Райса, принимающий в данном случае значения $K = 1$ (при $\Omega_Y = 10$ дБ) и $K = 0,01$ (при $\Omega_Y = -10$ дБ). Благодаря

слабому затенению информацию можно передавать быстро, так что вероятность прерывания БСС достигает низких значений.

Семейство сплошных линий, соответствующих высокой средней мощности компонент прямой видимости, проходит выше, чем семейство пунктирных линий, отображающих ситуацию $\Omega_Y = -10$ дБ. Это означает, что появление прямого луча с высокой мощностью повышает вероятность прерывания БСС, то есть, безопасность связи становится сложнее обеспечить. Ухудшение ситуации существенно больше при низкой нормированной пропускной способности C_{th} (в 15 раз при $C_{th} = 0{,}1$ и только в 1,5 раза – при $C_{th} = 0{,}9$ при неизменном $\overline{\gamma}_W = -10$ дБ, красные линии) и при высоком среднем ОСШ в канале утечки (в 15 раз при $\overline{\gamma}_W = -10$ дБ и в 10 раз при $\overline{\gamma}_W = 0$ дБ при неизменном $C_{th} = 0{,}1$). При дальнейшем увеличении $\overline{\gamma}_W$ влияние мощности прямой компоненты перестаёт сказываться, а сеанс связи перестаёт быть безопасным.

Для канала с множеством затенённых компонент прямой видимости (для примера 10 компонент $m_Y = 5$), слабыми многопутевыми компонентами ($\Omega_X = -10$ дБ) и сильным общим затенением ($m_X = 0{,}5$) при фиксированном среднем ОСШ основного канала $\overline{\gamma}_M = 10$ дБ и различных средних мощностях компоненты прямой видимости Ω_Y зависимости вероятности прерывания БСС от пороговой пропускной способности БСС в канале приведены на рисунке 8. Сеанс связи в таких условиях будет безопасным с меньшей вероятностью (SOP выше, чем в ситуации канала на рисунке 7). Благодаря представлению зависимостей только в области значений высоких вероятностей прерывания безопасного сеанса связи, на рисунке 7 является заметным эффект смены характера влияния мощности компонент прямой видимости при повышении ОСШ в канале прослушки до и более ОСШ в основном канале. При $\overline{\gamma}_W = -10$ дБ (что на 20 дБ ниже, чем $\overline{\gamma}_M$) добавление энергии компонентам прямой видимости понижает вероятность прерывания БСС, причём тем сильнее, чем ниже пороговое значение нормированной пропускной способности C_{th} (красная пунктирная линия проходит выше красной сплошной во всем диапазоне значений C_{th}). При $\overline{\gamma}_W = 10$ и 20 дБ (что равно или на 10 дБ больше, чем $\overline{\gamma}_M$, соответственно) смена ситуации с $\Omega_Y = -10$ дБ на $\Omega_Y = 10$ дБ уже повышает вероятность прерывания БСС, а величина изменения

зависит от C_{th}. В промежуточной ситуации, когда $\overline{\gamma}_W = \overline{\gamma}_M = 10$ дБ, существует такое пороговое значение C_{th}, соответствующее точке пересечения синей сплошной и пунктирной линий, ниже которого усиление компонент прямой видимости влияет положительно, а выше отрицательно.

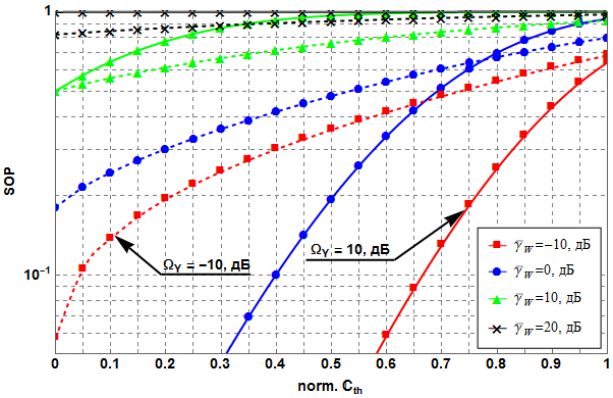


Рис. 8. Зависимость вероятности прерывания БСС от пороговой пропускной способности БСС в канале с 10 компонентами прямой видимости ($m_Y = 5$), слабыми многопутевыми компонентами ($\Omega_X = -10$ дБ) и слабым общим затенением ($m_X = 0{,}5$) при фиксированном среднем ОСШ основного канала $\overline{\gamma}_M = 10$ дБ для высокой ($\Omega_Y = 10$ дБ) и низкой ($\Omega_Y = -10$ дБ) средней мощности компонент прямой видимости

Ещё две показательные ситуации для сильного сигнала ($\Omega_X = 10$ дБ, $\Omega_Y = 10$ дБ) при $\overline{\gamma}_M = 10$ дБ приведены на рисунке 9: практическое отсутствие затенения и компонент прямой видимости, и многопутевых компонент ($m_X = m_Y = 5$, сплошные линии) и дружное сильное затенение всех компонент сигнала ($m_X = m_Y = 0{,}5$, пунктирные линии).

Влияние степени затенения компонент сигнала проявляется аналогично влиянию мощности компонент прямой видимости. При $\overline{\gamma}_W = -10$ дБ уменьшение дружного затенения компонент сигнала понижает вероятность прерывания БСС, причём тем сильнее, чем ниже пороговое значение нормированной пропускной способности C_{th} (красная пунктирная линия проходит выше красной сплошной во всем

диапазоне значений C_{th}). При $\overline{\gamma}_W = 10$ и 20 дБ смена степени затенения с большой на малую уже повышает вероятность прерывания БСС, а величина изменения зависит от C_{th} . В промежуточной ситуации, когда $\overline{\gamma}_W = \overline{\gamma}_M = 10$ дБ, наблюдается зависимость, аналогичная вышеописанной.

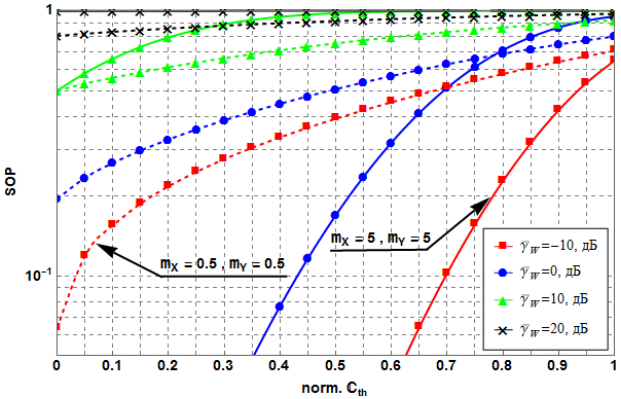


Рис. 9. Зависимость вероятности прерывания БСС от пороговой пропускной способности БСС при фиксированном среднем ОСШ основного канала $\overline{\gamma}_M = 10$ дБ для сильного сигнала ($\Omega_X = 10$ дБ, $\Omega_Y = 10$ дБ) и различной степени затенения

Таким образом, если в Гауссовском канале при $\overline{\gamma}_W \geq \overline{\gamma}_M$ нет такой пропускной способности, при которой сеанс связи был бы безопасным, то, как демонстрируют рисунок 8 и 9, при наличии затенений безопасный сеанс связи возможен до некоторого значения C_{th} , зависящего от соотношения между ОСШ в основном канале и канале прослушки.

Как показывает сравнение зависимостей на рисунке 9 с зависимостями, приведёнными на рисунке 7 сплошными линиями, в случае сильного сигнала и $\overline{\gamma}_M = 10$ дБ вероятность прерывания безопасного сеанса связи при общем слабом затенении, в основном приходящемся на компоненты прямой видимости, выше, чем в случае дружно слабо затенённых компонент. Это означает, что в канале с $m_X = m_Y = 5$ с той же вероятностью можно гарантировать более высокие скорости при безопасной передаче, например, при SOP 0,1

одинаковое затенение компонент обеспечит рост C_{th} с 0,51 до 0,70 при $\overline{\gamma}_W = -10$ дБ и с 0,30 до 0,43 при $\overline{\gamma}_W = 0$ дБ.

При выставлении высокой пороговой пропускной способности БСС, например, равной 0,9, эффект, связанный с затенениями, практически нивелируется и вероятность прерывания БСС будет всегда на высоком уровне вне зависимости от наличия или отсутствия затенений в канале связи.

6.4. Размещение законного и подслушивающего приёмников друг относительно друга. Рисунки 10 и 11 демонстрируют зависимость вероятности прерывания безопасного сеанса связи от отношения расстояний до подслушивающего и законного приёмников d_W / d_M на примерах ситуации с большим количеством ($2m_Y = 10$) мощных ($\Omega_Y = 5$ или 10 дБ) и слабо затенённых доминантных компонент при малой мощности рассеянных компонент ($\Omega_X = -10$ дБ) и фиксированном пороге пропускной способности БСС ($C_{th} = 0,6$) и различных степенях общего затенения m_X . Показатель ослабления сигнала в среде α зафиксирован на уровне 3, что соответствует городской местности [24].

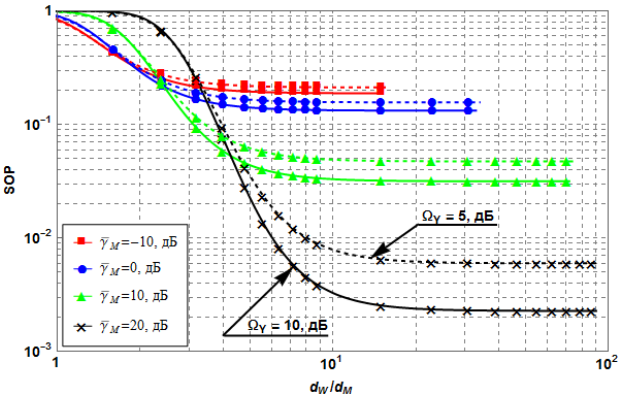


Рис. 10. Влияние отношения расстояний до подслушивающего и законного приёмников при $m_Y = 1$, $\Omega_X = -10$ дБ, $\Omega_Y = 5$ или 10 дБ, $C_{th} = 0,6$, $\alpha = 3$ для небольших затенений $m_X = 1$

Спад SOP с ростом d_W / d_M при всех наборах других параметров обладает общей особенностью: существует такое значение

$(d_W / d_M)_{порог}$, начиная с которого наблюдается неснижаемый уровень вероятности прерывания БСС $SOP_{\min}$. Пороговое значение $(d_W / d_M)_{порог}$ лежит в области около значения 10; его величина, а также значение $SOP_{\min}$ зависят от параметров канала и соотношения сигнал-шум в основном канале.

С ростом $\overline{\gamma}_M$ выход на эту «полочку» происходит при больших относительных расстояниях, а её уровень становится ниже. Например, при $\Omega_Y = 10$ дБ (сплошные линии на рисунке 10) рост $\overline{\gamma}_M$ с -10 дБ до 20 дБ обеспечивает появление неснижаемой вероятности $SOP_{\min} = 0,0025$ вместо 0,23 лишь при $(d_W / d_M)_{порог} = 15$ вместо 3,2.

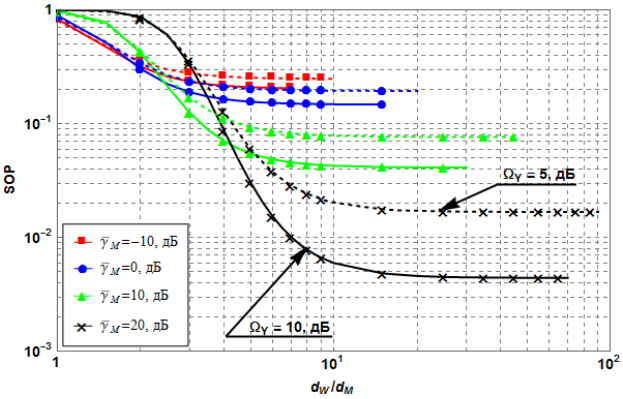


Рис. 11. Влияние отношения расстояний до подслушивающего и законного приёмников при $m_Y = 0,5$, $\Omega_X = -10$ дБ, $\Omega_Y = 5$ или 10 дБ, $C_{th} = 0,6$, $\alpha = 3$ для глубоких затенений $m_X = 0,5$

Величина мощности компонент прямой видимости, а также общая степень затенения практически не влияют на пороговое отношение дальностей приёмников, но определяют величину наилучшей достижимой вероятности прерывания БСС $SOP_{\min}$. Например, увеличение затенения (m_X уменьшается с 1 до 0,5) при $\overline{\gamma}_M = 20$ дБ порождает повышение $SOP_{\min}$ с 0,006 до 0,019 при $\Omega_Y = 5$ дБ (пунктирные чёрные линии на рисунках 10 и 11) и с 0,0023 до 0,0045 при $\Omega_Y = 10$ дБ (сплошные чёрные линии).

Общая степень затенения сигнала также влияет на величину изменения вероятности прерывания БСС при увеличении мощности компонент прямой видимости.

Сопоставление интервалов $\Delta SOP_{\min}$ между абсолютными значениями $SOP_{\min}$ для красных и чёрных линий на рисунках 10 и 11 показывает: по мере уменьшения m_X с 1 до 0,5 $\Delta SOP_{\min}$ растёт с 0,02 до 0,04 (в 2 раза) при $\overline{\gamma}_M = -10$ дБ и с 0,0037 до 0,0135 (в 2 раза) при $\overline{\gamma}_M = 20$ дБ.

7. Прогноз безопасности связи. Представляет интерес продемонстрировать оценки для вероятности прерывания безопасного сеанса связи и величины влияния на неё различных факторов на опубликованном примере реализованной передачи информации, для которого известны оценки качества передачи (вероятности ошибки, скорости передачи информации), как реально достигнутого в ходе эксперимента, так и прогнозы достижимых показателей качества при использовании наблюдаемого канала при других значениях доступных для изменения величин (в том числе ОСШ в канале – за счёт изменения мощности передатчика).

Будем опираться на эксперимент [30] с передачей данных в среде с мелкомасштабными замираниями в диапазоне 28 ГГц. В [16] авторы ВХ-модели оценили параметры этого канала путём минимизации тестовой статистики Колмогорова-Смирнова между эмпирическим распределением и предложенной моделью. Средние мощности многопутевых компонент и компонент прямой видимости сигнала и параметры затенения, извлечённые ими, составляют: $\Omega_X = -2$ дБ и $\Omega_Y = 0$ дБ, $m_X = 0,65$, $m_Y = 1,95$. Такие значения соответствуют каналу с сильным затенением практически четырёх многолучевых компонент, умеренным затенением компонент прямой видимости и почти равной средней мощностью многопутевых компонент и компонент прямой видимости. От системы передачи требовалась нормированная пропускная способность, равная 0,1.

Для таких условий была оценена с помощью полученных выражений вероятность прерывания безопасного сеанса связи. Результаты приведены на рисунках 12-14.

Как показывает рисунок 12, для данного канала пороговое значение d_W / d_M , при превышении которого увеличение ОСШ в основном канале даёт положительный результат (снижает SOP), равно 2. Вероятность прерывания безопасного сеанса связи перестаёт улучшаться при превышении (т.е. приближении законного приёмника

к передатчику и неподвижном прослушивающем приёмнике) d_W / d_M чуть ниже 10 для ОСШ в основном канале 10 дБ.

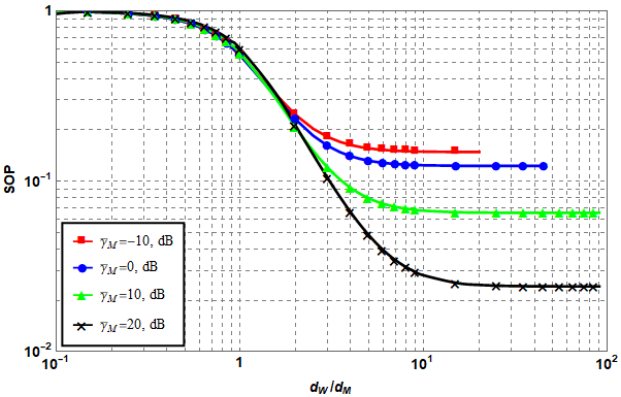


Рис. 12. Зависимость вероятности прерывания безопасного сеанса связи от отношения расстояний до подслушивающего и законного приёмников при $\alpha = 3$

Нормированная пропускная способность 0,1 при вероятности прерывания безопасного сеанса связи около 0,1 может быть для описываемого канала достигнута при ОСШ в основном канале 25 дБ, если в канале прослушки ОСШ равно 10 дБ (рисунок 13).

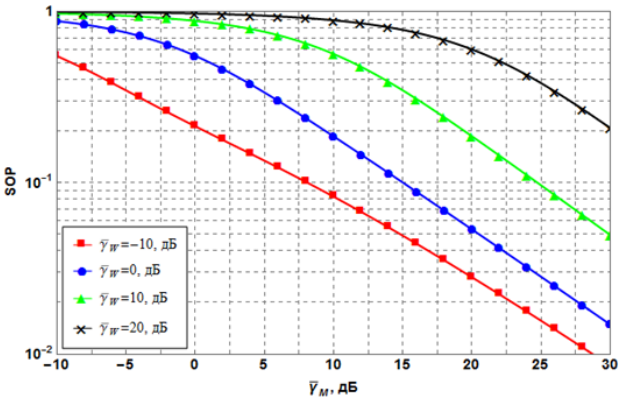


Рис. 13. Зависимость вероятности прерывания безопасного сеанса связи от ОСШ в основном канале

Рисунок 14 демонстрирует возможность управления вероятностью прерывания безопасного сеанса связи за счёт изменения

величины показателя потерь при распространении сигнала. В зависимости от среднего ОСШ в основном канале с таким набором параметров существуют значения α, вплоть до которых имеет смысл применять технические решения, способные повлиять в конечном итоге на величину SOP, в том числе использовать интеллектуальные рассеивающие панели [24-26]. При $\overline{\gamma}_M = 10$ дБ и $d_W / d_M = 4$ α не должно превышать 4, т.к. при больших α зависимости входят в режим насыщения.

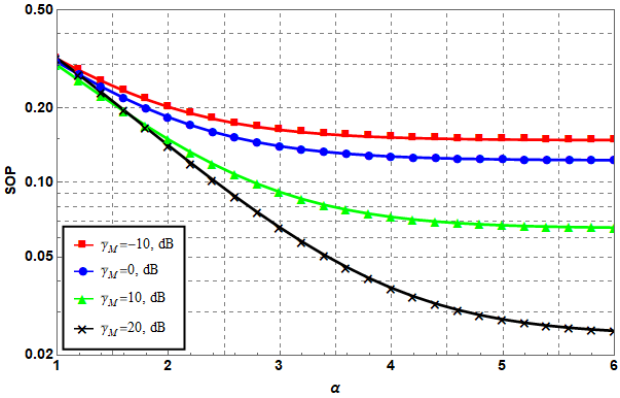


Рис. 14. Зависимость вероятности прерывания безопасного сеанса связи от показателя потерь на трассе при $d_W / d_M = 4$

8. Заключение. В работе получено аналитическое выражение в замкнутой форме для вероятности прерывания безопасного сеанса связи как показателя безопасности системы связи на физическом уровне для случая многолучевого ВХ канала с замираниями и наличием канала утечки. Был выполнен анализ для различных значений параметров: среднего значения ОСШ в основном канале и канале прослушки, эффективного значения показателя потерь на пути распространения сигнала, относительного расстояния между законным приемником и прослушивающим приёмником и пороговой пропускной способности, нормированной на пропускную способность гладкого гауссова канала.

Проведённый для наиболее типичных случаев свойств канала анализ вероятности прерывания безопасного сеанса связи показал:

– вероятность прерывания безопасного сеанса связи является высокой и не зависит от общей степени затенения компонент сигнала в

ситуации, когда подслушивающее устройство имеет лучшие шумовые характеристики, чем законный приёмник;

– для каналов с сильным общим затенением компонент сигнала существует неснижаемая с ростом среднего ОСШ в основном канале вероятность прерывания безопасного сеанса связи;

– величина энергетического потенциала, необходимого для гарантированной безопасной связи с заданной скоростью, определяется в первую очередь мощностью многопутевых компонент;

– существует такое пороговое значение минимальной требуемой скорости передачи, ниже которого усиление компонент прямой видимости или исчезновение общего затенения компонент сигнала уменьшает вероятность прерывания безопасного сеанса связи, а выше – повышает её, что наиболее ярко проявляется при одинаковых значениях ОСШ в основном канале и канале прослушки, равных 10 дБ;

– система связи при распространении сигнала в гиперрэлеевском канале с глубокими замираниями во многих практических случаях оказывается небезопасной с точки зрения защиты информации.

Полученные выражения, выводы и численные оценки могут быть использованы в рамках теоретических исследований для понимания влияния поведения сложных многопутевых каналов связи на характеристики безопасности систем связи. Они также применимы в рамках практических (инженерных) приложений, для обоснования рекомендаций по выбору или искусственному формированию структуры и условий распространения сигнала в зависимости от предъявляемых требований к вероятности прерывания безопасного сеанса связи и минимальной скорости передачи.

Литература

1. Kalyani V.L., Sharma D. IoT: machine to machine (M2M), device to device (D2D) internet of everything (IoE) and human to human (H2H): future of communication // Journal of Management Engineering and Information Technology (JMEIT). 2015. vol. 2. no. 6. pp. 17-23.
2. Jurgen R.K. (ed.). V2V/V2I communications for improved road safety and efficiency. // SAE International. 2012.
3. Lai K., Yanushkevich S.N., Shmerko V.P. Intelligent stress monitoring assistant for first responders // IEEE Access. 2021. vol. 9. pp. 25314-25329.
4. Shrestha R. et al. Evolution of V2X communication and integration of blockchain for security enhancements // Electronics. 2020. vol. 9. no. 9. p. 1338.
5. Qian Y., Ye F., Chen H.-H. Security in V2X communications // Security in Wireless Communication Networks, IEEE, 2022. pp. 311-331. doi: 10.1002/9781119244400.ch15.

6. Hasan M. et al. Securing vehicle-to-everything (V2X) communication platforms // IEEE Transactions on Intelligent Vehicles. 2020. vol. 5. no. 4. pp. 693-713. doi: 10.1109/TIV.2020.2987430.
7. Hamamreh J.M., Furqan H.M., Arslan H. Classifications and applications of physical layer security techniques for confidentiality: a comprehensive survey // IEEE Communications Surveys & Tutorials. 2018. vol. 21. no. 2. pp. 1773-1828. doi: 10.1109/COMST.2018.2878035.
8. Sánchez J.D.V. et al. Survey on physical layer security for 5G wireless networks // Annals of Telecommunications. 2021. vol. 76. no. 3. pp. 155-174.
9. Wu Y. et al. A survey of physical layer security techniques for 5G wireless networks and challenges ahead // IEEE Journal on Selected Areas in Communications. 2018. vol. 36. no. 4. pp. 679-695. doi: 10.1109/JSAC.2018.2825560.
10. Probability distributions relevant to radiowave propagation modelling // Recommendations ITU-R P.1057-6 (08/2019). URL: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.1057-6-201908-I!!PDF-E.pdf
11. Hyadi A., Rezki Z., Alouini M.S. An overview of physical layer security in wireless communication systems with CSIT uncertainty // IEEE Access. 2016. vol. 4. pp. 6121-6132. doi: 10.1109/ACCESS.2016.2612585.
12. Li S. et al. Amount of secrecy loss: a novel metric for physical layer security analysis // IEEE Communications Letters. 2020. vol. 24. no. 8. pp. 1626-1630. doi: 10.1109/LCOMM.2020.2995731.
13. Barros J., Rodrigues M.R.D. Secrecy capacity of wireless channels // 2006 IEEE international symposium on information theory. IEEE. 2006. pp. 356-360. doi: 10.1109/ISIT.2006.261613.
14. Fadnis C., Katiyar B. Review of higher order statistics for selection combining scheme in Weibull fading channel // 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC). IEEE. 2017. pp. 648-651. doi: 10.1109/CTCEEC.2017.8455182.
15. Peppas K.P., Nistazakis H.E., Tombras G.S. An overview of the physical insight and the various performance metrics of fading channels in wireless communication systems // Advanced trends in wireless communications. 2011. pp. 1-22. doi: 10.5772/15028.
16. Olutayo A., Cheng J., Holzman J.F. A new statistical channel model for emerging wireless communication systems // IEEE Open Journal of the Communications Society. 2020. vol. 1. pp. 916-926. doi: 10.1109/ojcoms.2020.3008161.
17. Gvozdarev A.S. A novel unified framework for energy-based spectrum sensing analysis in the presence of fading // Sensors. 2022. vol. 22. no. 5. pp. 1742. doi: 10.3390/s22051742.
18. Olutayo A., Cheng J., Holzman J.F. Performance bounds for diversity receptions over a new fading model with arbitrary branch correlation // EURASIP Journal on Wireless Communications and Networking. 2020. vol. 2020. no. 1. pp. 1-26.
19. Wyner A.D. The wire-tap channel // Bell system technical journal. 1975. vol. 54. no. 8. pp. 1355-1387. doi: 10.1002/j.1538-7305.1975.tb02040.x.
20. Liang Y. et al. Information theoretic security // Foundations and Trends in Communications and Information Theory. 2009. vol. 5. no. 4–5. pp. 355-580.
21. Liu R. Securing wireless communications at the physical layer. New York, NY, USA: Springer, 2010. vol. 7.
22. Olver F.W.J. et al. NIST digital library of mathematical functions, release 1.0.22. 2019. URL: http://dlmf.nist.gov/ (дата обращения: 1.07.2022).
23. Beaulieu N.C., Xie J. A novel fading model for channels with multiple dominant specular components // IEEE Wireless Communications Letters. 2014. vol. 4. no. 1. pp. 54-57. doi: 10.1109/LWC.2014.2367501.

24. Cho Y. S. et al. MIMO-OFDM wireless communications with MATLAB. John Wiley & Sons, 2010.
25. Li Z. et al. Enhancing indoor mmWave wireless coverage: small-cell densification or reconfigurable intelligent surfaces deployment? // IEEE Wireless Communications Letters. 2021. vol. 10. no. 11. pp. 2547-2551. doi: 10.1109/LWC.2021.3106821.
26. Gvozdarev A.S., Patralov P.E., Artemova T.K., Murin D.M. Reconfigurable intelligent surfaces' impact on the physical layer security of the Beaulieu-Xie shadowed fading channel // 2022 International Symposium on Networks, Computers and Communications (ISNCC). 2022. pp. 1-5.
27. Bender C.M., Orszag S., Orszag S.A. Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory. // Springer Science & Business Media. 1999. vol. 1.
28. Shannon C.E. Communication theory of secrecy systems // The Bell system technical journal. 1949. vol. 28. no. 4. pp. 656-715.
29. Frolik J. A case for considering hyper-Rayleigh fading channels // IEEE transactions on wireless communications. 2007. vol. 6. no. 4. pp. 1235-1239. doi: 10.1109/TWC.2007.348319.
30. Samimi M.K. et al. 28 GHz millimeter-wave ultrawideband small-scale fading models in wireless channels // 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring). IEEE. 2016. pp. 1-6. doi: 10.1109/VTCSpring.2016.7503970.

Гвоздарев Алексей Сергеевич — канд. физ.-мат. наук, доцент, кафедрf интеллектуальных информационных радиофизических систем физического факультета, Ярославский государственный университет им. П.Г. Демидова. Область научных интересов: статистическая обработка сигналов в беспроводных системах связи, применение методов математической статистики в задачах обработки и передачи информации. Число научных публикаций — 80. a.gvozdarev@uniyar.ac.ru; улица Советская, 14, 150003, Ярославль, Россия; р.т.: +7(4852)797-770.

Артёмова Татьяна Константиновна — канд. физ.-мат. наук, доцент, Ярославский государственный университет им. П.Г. Демидова. Область научных интересов: обработка сигналов в беспроводных системах связи, антенны, микроволновая техника и технологии. Число научных публикаций — 66. artemova@uniyar.ac.ru; улица Советская, 14, 150003, Ярославль, Россия; р.т.: +7(4852)797-770.

Патралов Павел Евгеньевич — магистр, кафедра интеллектуальных информационных радиофизических систем физического факультета, Ярославский государственный университет им. П.Г. Демидова. Область научных интересов: математическая статистика, теория вероятности, применение методов математического моделирования теории связи. Число научных публикаций — 5. p.patralov1@stud.uniyar.ac.ru; улица Советская, 14, 150003, Ярославль, Россия; р.т.: +7(4852)797-770.

Мурин Дмитрий Михайлович — канд. физ.-мат. наук, директор, институт информационной безопасности, Ярославский государственный университет им. П.Г. Демидова. Область научных интересов: информационная безопасность различных систем, блокчейн-системы. Число научных публикаций — 17. d.murin@uniyar.ac.ru; улица Советская, 14, 150003, Ярославль, Россия; р.т.: +7(4852)788-591.

Поддержка исследований. Работа выполнена при финансовой поддержке РНФ (проект № 22-29-01458).

DOI 10.15622/ia.21.5.8

A. GVOZDAREV, T. ARTEMOVA, P. PATRALOV, D. MURIN

THE STATISTICAL ANALYSIS OF THE SECURITY FOR A WIRELESS COMMUNICATION SYSTEM WITH A BEAULIEU-XIE SHADOWED FADING MODEL CHANNEL

Gvozdarev A., Artemova T., Patralov P., Murin D. **The Statistical Analysis of the Security for a Wireless Communication System with a Beaulieu-Xie Shadowed Fading Model Channel.**

Abstract. The paper considers the problem of a wireless communication system's physical level security for a multipath signal propagation channel and the presence of a wiretap channel. To generalize the propagation effects, the Beaulieu-Xie shadowed channel model was assumed. To describe the security of the information transfer process, such a metric as the secure outage probability of was considered. An analytical expression of the secure outage probability was obtained and analyzed depending on the characteristics of the channel and the communication system: the average value of the signal-to-noise ratio in the main channel and the wiretap channel, the effective path-loss exponent, the relative distance between the legitimate receiver and the wiretap and the threshold bandwidth, normalized to the bandwidth of a smooth Gaussian channel. The analysis considers the sets of parameters that cover all practically important scenarios for the functioning of a wireless communication systems: both deep fading (corresponding to the hyper-Rayleigh scenario) and small fading, both in the case of a significant line-of-sight component and a significant number of multipath clusters, and with significant shadowing of the dominant component and a small number of multipath waves, as well as all intermediate options. It is found out that the value of the energy requirements for guaranteed secure communication at a given speed is determined primarily by the power of multipath components, as well as the existence of an irreducible secure outage probability of a communication session with an increase for channels with strong overall shadowing of the signal components, which from a practical point of view is important to take into account when imposing requirements for the values of the signal-to-noise ratio and the data transfer rate in the direct channel, providing the desired degree of security of the wireless communication session.

Keywords: wireless channel, fading, shadowing, Beaulieu-Xie shadowed fading model, secrecy outage probability.

References

1. Kalyani V.L., Sharma D. IoT: machine to machine (M2M), device to device (D2D) internet of everything (IoE) and human to human (H2H): future of communication // Journal of Management Engineering and Information Technology (JMEIT). 2015. vol. 2. no. 6. pp. 17-23.
2. Jurgen R.K. (ed.). V2V/V2I communications for improved road safety and efficiency. // SAE International. 2012.
3. Lai K., Yanushkevich S.N., Shmerko V.P. Intelligent stress monitoring assistant for first responders // IEEE Access. 2021. vol. 9. pp. 25314-25329.
4. Shrestha R. et al. Evolution of V2X communication and integration of blockchain for security enhancements // Electronics. 2020. vol. 9. no. 9. p. 1338.
5. Qian Y., Ye F., Chen H.-H. Security in V2X communications // Security in Wireless Communication Networks, IEEE, 2022. pp. 311-331. doi: 10.1002/9781119244400.ch15.

6. Hasan M. et al. Securing vehicle-to-everything (V2X) communication platforms // IEEE Transactions on Intelligent Vehicles. 2020. vol. 5. no. 4. pp. 693-713. doi: 10.1109/TIV.2020.2987430.
7. Hamamreh J.M., Furqan H.M., Arslan H. Classifications and applications of physical layer security techniques for confidentiality: a comprehensive survey // IEEE Communications Surveys & Tutorials. 2018. vol. 21. no. 2. pp. 1773-1828. doi: 10.1109/COMST.2018.2878035.
8. Sánchez J.D.V. et al. Survey on physical layer security for 5G wireless networks // Annals of Telecommunications. 2021. vol. 76. no. 3. pp. 155-174.
9. Wu Y. et al. A survey of physical layer security techniques for 5G wireless networks and challenges ahead // IEEE Journal on Selected Areas in Communications. 2018. vol. 36. no. 4. pp. 679-695. doi: 10.1109/JSAC.2018.2825560.
10. Probability distributions relevant to radiowave propagation modelling // Recommendations ITU-R P.1057-6 (08/2019). URL: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.1057-6-201908-I!!PDF-E.pdf
11. Hyadi A., Rezki Z., Alouini M.S. An overview of physical layer security in wireless communication systems with CSIT uncertainty // IEEE Access. 2016. vol. 4. pp. 6121-6132. doi: 10.1109/ACCESS.2016.2612585.
12. Li S. et al. Amount of secrecy loss: a novel metric for physical layer security analysis // IEEE Communications Letters. 2020. vol. 24. no. 8. pp. 1626-1630. doi: 10.1109/LCOMM.2020.2995731.
13. Barros J., Rodrigues M.R.D. Secrecy capacity of wireless channels // 2006 IEEE international symposium on information theory. IEEE. 2006. pp. 356-360. doi: 10.1109/ISIT.2006.261613.
14. Fadnis C., Katiyar B. Review of higher order statistics for selection combining scheme in Weibull fading channel // 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC). IEEE. 2017. pp. 648-651. doi: 10.1109/CTCEEC.2017.8455182.
15. Peppas K.P., Nistazakis H.E., Tombras G.S. An overview of the physical insight and the various performance metrics of fading channels in wireless communication systems // Advanced trends in wireless communications. 2011. pp. 1-22. doi: 10.5772/15028.
16. Olutayo A., Cheng J., Holzman J.F. A new statistical channel model for emerging wireless communication systems // IEEE Open Journal of the Communications Society. 2020. vol. 1. pp. 916-926. doi: 10.1109/ojcoms.2020.3008161.
17. Gvozdarev A.S. A novel unified framework for energy-based spectrum sensing analysis in the presence of fading // Sensors. 2022. vol. 22. no. 5. pp. 1742. doi: 10.3390/s22051742.
18. Olutayo A., Cheng J., Holzman J.F. Performance bounds for diversity receptions over a new fading model with arbitrary branch correlation // EURASIP Journal on Wireless Communications and Networking. 2020. vol. 2020. no. 1. pp. 1-26.
19. Wyner A.D. The wire-tap channel // Bell system technical journal. 1975. vol. 54. no. 8. pp. 1355-1387. doi: 10.1002/j.1538-7305.1975.tb02040.x.
20. Liang Y. et al. Information theoretic security // Foundations and Trends in Communications and Information Theory. 2009. vol. 5. no. 4–5. pp. 355-580.
21. Liu R. Securing wireless communications at the physical layer. New York, NY, USA: Springer, 2010. vol. 7.
22. Olver F.W.J. et al. NIST digital library of mathematical functions, release 1.0.22. 2019. URL: http://dlmf.nist.gov/ (дата обращения: 1.07.2022).
23. Beaulieu N.C., Xie J. A novel fading model for channels with multiple dominant specular components // IEEE Wireless Communications Letters. 2014. vol. 4. no. 1. pp. 54-57. doi: 10.1109/LWC.2014.2367501.

24. Cho Y. S. et al. MIMO-OFDM wireless communications with MATLAB. John Wiley & Sons, 2010.
25. Li Z. et al. Enhancing indoor mmWave wireless coverage: small-cell densification or reconfigurable intelligent surfaces deployment? // IEEE Wireless Communications Letters. 2021. vol. 10. no. 11. pp. 2547-2551. doi: 10.1109/LWC.2021.3106821.
26. Gvozdarev A.S., Patralov P.E., Artemova T.K., Murin D.M. Reconfigurable intelligent surfaces' impact on the physical layer security of the Beaulieu-Xie shadowed fading channel // 2022 International Symposium on Networks, Computers and Communications (ISNCC). 2022. pp. 1-5.
27. Bender C.M., Orszag S., Orszag S.A. Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory. // Springer Science & Business Media. 1999. vol. 1.
28. Shannon C.E. Communication theory of secrecy systems // The Bell system technical journal. 1949. vol. 28. no. 4. pp. 656-715.
29. Frolik J. A case for considering hyper-Rayleigh fading channels // IEEE transactions on wireless communications. 2007. vol. 6. no. 4. pp. 1235-1239. doi: 10.1109/TWC.2007.348319.
30. Samimi M.K. et al. 28 GHz millimeter-wave ultrawideband small-scale fading models in wireless channels // 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring). IEEE. 2016. pp. 1-6. doi: 10.1109/VTCSpring.2016.7503970.

Gvozdarev Aleksey — Ph.D., Associate professor, Department of intelligent radiophysical information systems (physics faculty), P.G. Demidov Yaroslavl State University. Research interests: statistical signal processing in wireless communication systems, application of mathematical statistics methods for information transmission and processing. The number of publications — 80. a.gvozdarev@uniyar.ac.ru; 14, Sovetskaya St., 150003, Yaroslavl, Russia; office phone: +7(4852)797-770.

Artemova Tatiana — Ph.D., Associate professor, P.G. Demidov Yaroslavl State University. Research interests: signal processing in wireless communication systems, antennas, microwave devices and technologies. The number of publications — 66. artemova@uniyar.ac.ru; 14, Sovetskaya St., 150003, Yaroslavl, Russia; office phone: +7(4852)797-770.

Patralov Pavel — Master, Department of intelligent radiophysical information systems (physics faculty), P.G. Demidov Yaroslavl State University. Research interests: mathematical statistics, probability theory, application of mathematical modeling in wireless communications. The number of publications — 5. p.patralov1@stud.uniyar.ac.ru; 14, Sovetskaya St., 150003, Yaroslavl, Russia; office phone: +7(4852)797-770.

Murin Dmitriy — Ph.D., Director, Information security institute, P.G. Demidov Yaroslavl State University. Research interests: information safety of various systems, blockchain systems. The number of publications — 17. d.murin@uniyar.ac.ru; 14, Sovetskaya St., 150003, Yaroslavl, Russia; office phone: +7(4852)788-591.

Acknowledgements. This work was supported by Russian Science Foundation under Grant 22-29-01458.

Руководство для авторов

Взаимодействие автора с редакцией осуществляется через личный кабинет на сайте журнала «Информатика и автоматизация» http://ia.spcras.ru/. При регистрации авторам рекомендуется заполнить все предложенные поля данных. Подготовка статьи ведется с помощью текстовых редакторов MS Word 2007 и выше или LaTex. Объем основного текста (до раздела Литература) - от 20 до 30 страниц включительно. Переносы разрешены. Номера страниц не проставляются. Основная часть текста статьи разбивается на разделы, среди которых являются обязательными: введение, хотя бы один «содержательный» раздел и заключение. Допускается также мотивированное содержанием и структурой материал а выделение подразделов. В основную часть опускается помещать рисунки, таблицы, листинги и формулы. Правила их оформления подробно рассмотрены на нашем сайте в разделе «Руководство для авторов».

Author guidelines

Interaction between each potential author and the Editorial board is realized through the pesoal account on the website of the journal “Informatics and Automation” http://ia.spcras.ru/. At the registration the authors are requested to fill out all data fields in the proposed form. The submissions should be prepared using MS Word 2007, LaTex. The text of the paper in the main part should not exceed 30 pages. Pages are not numbered; hyphenations are allowed. Certain figures, tables, listings and formulas are allowed in the main section, and their typography is considered in more detail at the journal web.

Signed to print 01.10.2022
Printed in Publishing center GUAP, 67 litera A, B. Morskaya, St. Petersburg, Russia, 190000

The journal is registered in the Russian Federal Agency for Communications and Mass-Media Supervision, certificate ПИ № ФС77-79228 dated September 25, 2020
Subscription Index П5513, Russian Post Catalog

Подписано к печати 01.10.2022. Формат 60×90 1/16. Усл. печ. л. 12,26. Заказ № 293.
Тираж 300 экз., цена свободная.
Отпечатано в Редакционно-издательском центре ГУАП,
190000, Санкт-Петербург, Б. Морская, д. 67, лит. А

Журнал зарегистрирован Федеральной службой по надзору в сфере связи и массовых коммуникаций, свидетельство ПИ № ФС77-79228 от 25 сентября 2020 г.

Подписной индекс П5513 по каталогу «Почта России»