

THE MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE RUSSIAN FEDERATION



Computing, Telecommunications and Control

**Vol. 14, No. 3
2021**

Peter the Great St. Petersburg
Polytechnic University
2021

EDITORIAL COUNCIL

Head of the editorial council

Prof. Dr. *Rafael M. Yusupov* (corresponding member of the Russian Academy of Sciences)

Members:

Prof. Dr. *Sergey M. Abramov* (corresponding member of the Russian Academy of Sciences),
Prof. Dr. *Dmitry G. Arseniev*,
Prof. Dr. *Vladimir V. Voevodin* (corresponding member of the Russian Academy of Sciences),
Prof. Dr. *Vladimir S. Zaborovsky*,
Prof. Dr. *Vladimir N. Kozlov*,
Prof. Dr. *Alexandr E. Fotiadi*,
Prof. Dr. *Igor G. Chernorutsky*.

EDITORIAL BOARD

Editor-in-chief

Prof. Dr. *Alexander S. Korotkov*, Peter the Great St. Petersburg Polytechnic University, Russia;

Members:

Assoc. Prof. Dr. *Vladimir M. Itsykson*, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Philippe Ferrari*, Head of the RF and Millimeter-Wave Lab IMEP-LAHC Microelectronics, Electromagnetism and Photonic Institute, Grenoble Alpes University, France;
Prof. Dr. *Yevgeni Koucheryavy*, Tampere University of Technology, Finland.
Prof. Dr. *Wolfgang Krautschneider*, Head of Nanoelectronics Institute, Hamburg University of Technology, Germany;
Prof. Dr. *Fa-Long Luo*, Affiliate Full Professor University of Washington, USA, Chief Scientist Micron Technology, Inc., Milpitas, USA, Chairman IEEE SPS Industry DSP Technology Standing Committee;
Prof. Dr. *Sergey B. Makarov*, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Emil Novakov*, IMEP-LAHC Microelectronics, Electromagnetism and Photonic Institute, Grenoble, France;
Prof. Dr. *Nikolay N. Prokopenko*, Don State Technical University, Rostov-on-Don, Russia;
Prof. Dr. *Mikhail G. Putrya*, National Research University of Electronic Technology, Moscow, Russia;
Sen. Assoc. Prof. Dr. *Evgeny Pyshkin*, School of Computer Science and Engineering, University of Aizu, Japan;
Prof. Dr. *Viacheslav P. Shkodyrev*, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Peter V. Trifonov*, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Igor A. Tsikin*, Professor, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Sergey M. Ustinov*, Peter the Great St. Petersburg Polytechnic University, Russia;
Prof. Dr. *Lev V. Utkin*, Peter the Great St. Petersburg Polytechnic University, Russia.

The journal is included in the List of Leading PeerReviewed Scientific Journals and other editions to publish major findings of PhD theses for the research degrees of Doctor of Sciences and Candidate of Sciences.

The journal is indexed by Ulrich's Periodicals Directory, Google Scholar, EBSCO, ProQuest, Index Copernicus, VINITI RAS Abstract Journal (Referativnyi Zhurnal), VINITI RAS Scientific and Technical Literature Collection, Russian Science Citation Index (RSCI) database © Scientific Electronic Library and Math-Net.ru databases.
ISSN 2687-0517

The journal is registered with the Federal Service for Supervision in the Sphere of Telecom, Information Technologies and Mass Communications (ROSKOMNADZOR). Certificate ЭЛ No. ФC77-77378 issued 25.12.2019.

No part of this publication may be reproduced without clear reference to the source.

The views of the authors can contradict the views of the Editorial Board.

The address: 195251 Polytekhnicheskaya Str. 29, St. Petersburg, Russia.



Информатика, телекоммуникации и управление

**Том 14, № 3
2021**

ИНФОРМАТИКА, ТЕЛЕКОММУНИКАЦИИ И УПРАВЛЕНИЕ

РЕДАКЦИОННЫЙ СОВЕТ ЖУРНАЛА

Председатель

Юсупов Р.М., чл.-кор. РАН;

Редакционный совет:

Абрамов С.М., чл.-кор. РАН;

Арсеньев Д.Г., д-р техн. наук, профессор;

Воеводин В.В., чл.-кор. РАН;

Заборовский В.С., д-р техн. наук, профессор;

Козлов В.Н., д-р техн. наук, профессор;

Фотиади А.Э., д-р физ.-мат. наук, профессор;

Черноруцкий И.Г., д-р техн. наук, профессор.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ ЖУРНАЛА

Главный редактор

Коротков А.С., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Редакционная коллегия:

Ицыксон В.М., канд. техн. наук, доцент, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Prof. Dr. *Philippe Ferrari*, Head of the RF and Millimeter-Wave Lab IMEP-LAHC Microelectronics, Electromagnetism and Photonic Institute, Grenoble Alpes University, France;

Prof. Dr. *Wolfgang Krautschneider*, Head of Nanoelectronics Institute, Hamburg University of Technology, Germany;

Кучерявый Е.А., канд. техн. наук, профессор, Tampere University of Technology, Finland.

Prof. Dr. *Fa-Long Luo*, Affiliate Full Professor University of Washington, USA, Chief Scientist Micron Technology, Inc., Milpitas, USA, Chairman IEEE SPS Industry DSP Technology Standing Committee;

Макаров С.Б., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Prof. Dr. *Emil Novakov*, IMEP-LAHC Microelectronics, Electromagnetism and Photonic Institute, Grenoble, France;

Прокопенко Н.Н., д-р техн. наук, профессор, Донской государственный технический университет, г. Ростов-на-Дону, Россия;

Путря М.Г., д-р техн. наук, профессор, Национальный исследовательский университет «Московский институт электронной техники», Москва, Россия;

Пышкин Е.В., канд. техн. наук, доцент, School of Computer Science and Engineering, University of Aizu, Japan;

Трифонов П.В., д-р техн. наук, доцент, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Устинов С.М., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Уткин Л.В., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Цикин И.А., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия;

Шкодырев В.П., д-р техн. наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, Россия.

Журнал с 2002 года входит в Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

Сведения о публикациях представлены в Реферативном журнале ВИНТИ РАН, в международной справочной системе «Ulrich`s Periodical Directory», в базах данных Российский индекс научного цитирования (РИНЦ), Google Scholar, EBSCO, Math-Net.Ru, ProQuest, Index Copernicus

ISSN 2687-0517

Журнал зарегистрирован Федеральной службой по надзору в сфере информационных технологий и массовых коммуникаций (Роскомнадзор). Свидетельство о регистрации Эл № ФС77-77378 от 25.12.2019.

При перепечатке материалов ссылка на журнал обязательна.

Точка зрения редакции может не совпадать с мнением авторов статей.

Адрес редакции: Россия, 195251, Санкт-Петербург, ул. Политехническая, д. 29.

Тел. редакции (812) 552-62-16.

Contents

Intellectual Systems and Technologies

| | |
|---|----|
| Konstantinov A.V. Deep gradient boosting for regression problems | 7 |
| Kudriashov N.S. Dynamic energy consumption rationing based on machine learning algorithms for oil refining tasks | 20 |
| Svistunova A.S., Khasanov D.S. Improving the efficiency of traffic management in a metropolis based on computer simulation | 33 |

Circuits and Systems for Receiving, Transmitting and Signal Processing

| | |
|--|----|
| Rachitskaya A.P. Semi-natural modeling for GNSS integrity monitoring algorithm | 43 |
| Aslanov G.K., Aslanov T.G., Kazibekov R.B., Musaibov R.R. Influence of transients in the information processing channel of the airport automatic radio direction finder on the direction finding accuracy | 56 |
| Assim A., Balashov E.V. Zero-drift operational amplifiers | 64 |



Содержание

Интеллектуальные системы и технологии

| | |
|---|----|
| Константинов А.В. Глубокий градиентный бустинг для решения задач регрессии | 7 |
| Кудряшов Н.С. Динамическое нормирование потребления энергоресурсов для задач нефтепереработки на основе алгоритмов машинного обучения | 20 |
| Свистунова А.С., Хасанов Д.С. Повышение эффективности управления транспортными потоками мегаполиса на основе имитационного моделирования | 33 |

Устройства и системы передачи, приёма и обработки сигналов

| | |
|--|----|
| Ращицкая А.П. Полунатурное моделирование алгоритма контроля целостности навигационного поля ГНСС | 43 |
| Асланов Г.К., Асланов Т.Г., Казобеков Р.Б., Мусаилов Р.Р. Влияние переходных процессов в канале обработки информации аэродромного автоматического радиопеленгатора на точность пеленгования | 56 |
| Ассим А., Балашов Е.В. Операционные усилители с нулевым дрейфом | 64 |

DOI: 10.18721/JCSTCS.14301
УДК 004

DEEP GRADIENT BOOSTING FOR REGRESSION PROBLEMS

A.V. Konstantinov

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

Deep Forest is a new machine-learning algorithm that combines the advantages of Deep Neural Networks and Decision Trees. It uses representation learning and allows building accurate compositions with a small amount of training data. A significant disadvantage of this approach is the inability to apply it directly to regression problems. First, feature generation method should be determined. Secondly, when replacing classification models with regression models, the set of distinct values of the Deep Forest model becomes limited by the set of values of its last layer. To eliminate the shortcomings, a new model, the Deep gradient boosting is proposed. The main idea is to iteratively improve the prediction using a new feature space. Features are generated based on the predictions of previously constructed cascade layers, by transforming predictions to a probability distribution. To reduce the time of model construction and overfitting, a mechanism for points screening is proposed. Experiments show the effectiveness of the proposed algorithm, in comparison with many existing methods for solving regression problems.

Keywords: regression, Deep Forest, gradient boosting, ensembles, decision trees.

Citation: Konstantinov A.V. Deep gradient boosting for regression problems. Computing, Telecommunications and Control, 2021, Vol. 14, No. 3, Pp. 7–19. DOI: 10.18721/JCSTCS.14301

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

ГЛУБОКИЙ ГРАДИЕНТНЫЙ БУСТИНГ ДЛЯ РЕШЕНИЯ ЗАДАЧ РЕГРЕССИИ

А.В. Константинов

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

Глубокий лес является новым подходом машинного обучения, сочетающим преимущества глубоких нейронных сетей и деревьев решений. Он использует обучение представлением и позволяет строить точные композиции при условии небольшого количества обучающих данных. Существенный недостаток данного подхода – невозможность напрямую применить его к задачам восстановления регрессии. Во-первых, требуется определить способ генерирования набора признаков. Во-вторых, при замене классификационных моделей на регрессионные, множество значений модели Глубокого леса становится ограниченным множеством значений последнего слоя. Для устранения недостатков предложена новая модель Глубокого градиентного бустинга. Основная идея состоит в итеративном улучшении предсказания, с использованием нового пространства признаков. Генерирование признаков производится на основании предсказаний ранее построенных моделей, путём трансформации их к распределению вероятностей. Для снижения времени построения модели и переобучения предложен механизм раннего отсева точек. Эксперименты показывают эффективность предложенного алгоритма по сравнению со многими существующими методами решения задачи регрессии.

Ключевые слова: регрессия, Глубокий лес, градиентный бустинг, ансамбли, деревья решений.

Ссылка при цитировании: Konstantinov A.V. Deep gradient boosting for regression problems // Computing, Telecommunications and Control. 2021. Vol. 14. No. 3. Pp. 7–19. DOI: 10.18721/JCSTCS.14301

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>).

Introduction

Over the past few years, one of the most important developments in the field of ensemble methods, that is, methods for constructing model compositions [1–4], has been a combination of ensemble-based models, including Random Forest [5] and stacking [6], proposed by Zhou and Feng and called Deep Forest, or gcForest [7]. The Deep Forest structure consists of layers similar to multilayer neural networks, but each gcForest layer contains several Random Forests, instead of perceptrons. As it was noted [7], Deep Forest is much easier to train, and it also has the ability to obtain high-quality results with extremely small data sizes available for training unlike Deep Neural Network, which requires a lot of effort to select hyperparameters and large data sizes for training. A lot of numerical experiments provided in [7] show that Deep Forest is superior to many widely used methods, or demonstrates results similar to existing methods when using default parameters.

One of the main advantages of Deep Forest is the ability to build deep models using representation learning without differentiable modules, that is without using reverse error propagation. Representation training allows you to identify complex dependencies in the data, and Deep Forest is built in such a way that a small amount of training sample is sufficient for training. For this reason, many modifications of Deep Forest have been developed, as well as adaptations to specific application tasks. In particular, Wang et al. [8] suggested using Deep Forest for predicting health status, or diagnostic health monitoring. Yang et al. [9] used Deep Forest to solve the problem of finding ships using thermal maps. Zheng et al. [10] considered the application of Deep Forest to the pedestrian detection problem.

The Deep Forest structure is cascaded [7]. Each layer of the cascade is an ensemble of decision trees. This structure embodies the idea of learning representations: the feature space changes from layer to layer, creating an increasingly informative representation of the input data. Each level of the cascade, or layer, receives a set of features generated by the previous layer, and the features obtained at the output of the layer are transmitted to the next level. The cascade architecture is shown in Fig. 1 (RF in the Figure is a random forest, and ET is Extra Trees, extremely randomized trees [11]). The diagram shows that each level of the cascade consists of two different pairs of random forests that generate probability distributions of classes (dimension 3), concatenated together with the original input. It is important to note that the structure of the layers can be modified to improve Deep Forest when solving problems of a specific type. The predictions of the last layer contain a set of probability distributions of classes, from which you can get the final prediction by averaging. Deep Forest's ability to learn representations can be supplemented by the second part of Deep Forest, called multi-grained scanning, which allows processing data with a spatial structure. Multi-level scanning uses the sliding window principle to process the original features. The result of such a scan is new feature vectors corresponding to windows of various sizes. Moreover, the scanning itself is performed using a set of decision trees trained with the original set of classes. In this paper, the first part of Deep Forest is considered, since the tasks under consideration do not require the use of multi-level scanning, which significantly affects the results in such areas as image or sound processing. For the input data instance, each forest in the layer determines an estimate of the probability distribution of classes by counting the frequencies of different classes for data instances that fell into the same sheet as this instance during training. After that, the overall estimate of the probability distribution of classes is calculated as the average of the estimates of each decision tree included in the random forest. The probability distribution of classes is represented by a vector, which is then concatenated with the original feature vector in order to use it as an input feature vector at the next stage. The use of the probability distribution vector for constructing

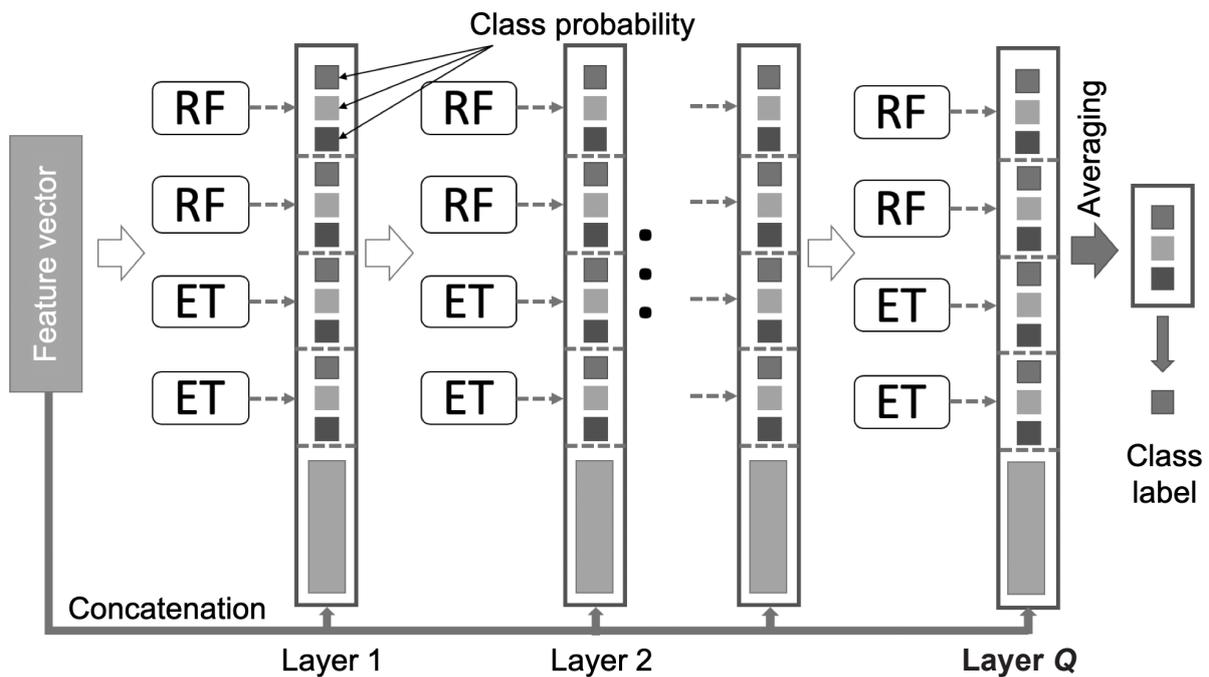


Fig. 1. Deep Forest cascade scheme

the next layer is similar to the idea of stacking [6]: in it, the basic models of the first layer are built using the original data set. After that, the stacking algorithm builds a new data set for training the second layer (meta-model), using the output vectors of the classes of the first layer as input features, and the same values are used as labels (values of the target variable) as for training the first layer. Also, the stacking algorithm often includes a cross-validation procedure to increase the generalizing ability when constructing the second layer. Unlike the standard stacking algorithm, Deep Forest uses both the original feature vector and the class vectors obtained from the previous layer on each layer. Zhou and Feng [7] also suggest using random forests of various types to provide a greater variety of predictions and create more informative features for the next layer.

Deep Forest has a significant shortcoming in comparison with neural networks: it is incapable of using arbitrary differentiable loss functions for any problems other than the standard classification one. In particular, traditional Deep Forest fails to solve such problems as transfer learning and distance metric learning. However, paper [12] eliminates this shortcoming by introducing weights allocated to decision trees, which allows the authors to solve a considerably wider range of problems with the help of the modified Deep Forest. Introduction of the weights for data instances similar to AdaBoost algorithms [13] together with the confidence screening mechanism allows both increasing the accuracy of classification and decreasing the time of building Deep Forest in a number of cases [15]. The AdaBoost algorithm [13] is based on the idea of boosting—iterative model construction as a combination of basic models, so that each subsequent basic model compensates for the shortcomings of the model in the previous iteration. According to [13], the training examples with the largest error in the previous iteration should correspond to large weights, but this approach is not the only possible one for applying the idea of boosting. Thus, one of the most popular approaches over the past decade is gradient boosting and its modifications [16]. Its key feature is the application of the principles of gradient descent in the functional space [17]. Each subsequent basic model is constructed in such a way as to approximate the value of the derivative of the loss function multiplied by minus one. Adding the predictions of the new base model to the already constructed combination corresponds to one step of gradient descent. Moreover, if the training example has an error close to

the minimum, the partial derivative will be close to zero, which is similar to the low weight of the training example in AdaBoost. Gradient boosting of decision trees is one of the most commonly used approaches for solving regression recovery problems with tabular data. The disadvantages of gradient boosting include the impossibility of parallel construction of ensemble elements, which leads to low performance with a large number of iterations, as well as a set of values determined by the training set in the leaves of each individual decision tree, which significantly limits the set of values of the entire ensemble. However, in [18], an algorithm for randomized gradient boosting (hereinafter referred to as RMGB) was proposed, based on partially random decision trees. This approach allows you to get a much larger set of values of the trained model, as well as improve performance by reducing the time of building each individual decision tree.

Despite Deep Forest’s great success in solving the classification problem, the cascade structure of models based on decision trees has not previously been used to solve the regression recovery problem, where the target variable has a continuous distribution. There are approaches that combine the principles of building neural networks and decision trees [19, 20]. Thus, in [19], an approach is described that allows the construction of regression models based on decision trees, in combination with convolutional neural networks, applied to the problem of determining the age from a photo. Its key disadvantage is the use of training the entire model using error back propagation, which greatly distinguishes this approach from Deep Forest, which has the advantages described above, in comparison with neural networks. Also, the approach proposed in [20], although it combines decision trees with neural networks, does not rely on a cascade structure, the construction of which can be performed iteratively without an error back propagation algorithm.

This paper discusses the problems of adapting Deep Forest to the regression problem, including the need to determine the method of combining predictions to obtain a new set of features, as well as the complexity of constructing functions, the set of values of which is quite large. Next, we propose a new approach based on the ideas of gradient boosting, as well as a method for generating a set of informative features. It allows you to build models using representation learning, without relying on an algorithm for back propagation of the error, obtaining arbitrarily accurate approximations of the original dependencies. To reduce the number of calculations and retraining, a mechanism for early elimination of points for which the predictions are already sufficiently accurate is also proposed.

Deep Forest

The main idea of a Deep Forest for the classification problem is to iteratively construct a set of layers, where each layer relies both on the original set of features and on an intermediate representation obtained from the previous layer. This approach allows you to increase the accuracy of the model simultaneously with the generalizing ability, while maintaining a low level of retraining. Let us imagine a deep forest as a composition of L layers S_i :

$$gcForest_L(x) = (S_L \circ S_{L-1} \circ \dots \circ S_1)(x).$$

Each layer can consist of arbitrary models that can be used to estimate the probabilities of classes. Namely, instead of predicting the class number, such models should be able to estimate the probability distribution of classes, such as, for example, compositions of decision trees: decision trees, random forests and extremely randomized trees. Indeed, the leaves of decision trees contain vectors consisting of the class frequencies of all elements of the training sample that fell into the corresponding leaf, and the compositions of decision trees allow us to estimate the probability distribution by averaging the class frequency vectors over all decision trees. The predicted frequency vectors of classes of different M_j^i compositions included in one layer of width W are concatenated into one vector, along with the original set of features:

$$S_i(\zeta) = \left(x, |M_1^i(\zeta)|, \dots, |M_w^i(\zeta)| \right)^T.$$

Each new layer either allows you to improve the representation of features from the previous layer, or has comparable performance. In the first case, the next layer is built in such a way that the target variable, that is, the class, is predicted from the set of features from the previous layer. In the second case, the composition is stopped. At the same time, a validation set is used to determine the quality of the representation of features.

Applicability of Deep Forest to the regression problem

This approach works well for classification, since the estimates of the probability distribution of classes contain significantly more information than the numbers of the predicted classes. And even the probability distribution, with the help of which it is impossible to make a correct prediction of the class, can be an informative set of features for constructing the next layer. Accordingly, this approach cannot be applied directly to the regression problem, for the following reasons. First, the complexity of generating a more suitable feature space for solving the problem, since it is impossible to directly calculate an analog of the class frequency vector for the regression problem, as in the case of the classification problem. Secondly, in addition to the representation of features suitable for solving the problem, in the case of regression, the ability to predict a large number of different values is also required, that is, the resolution, which for the entire Deep Forest will be limited by the resolution of the last layer. Thus, a Deep Forest in which classification trees are replaced with regression trees does not allow you to get better results than a random forest. Also, reducing the regression problem to classification will not allow achieving a higher quality of the model using Deep Forest, since such a reduction similarly reduces the resolution.

Gradient boosting based on Deep Forest

To solve the described problems, two significant changes are made to the Deep Forest model. First, each layer should consist of strong regression models, that is, models that allow you to approximate arbitrary functions with a given accuracy. For this purpose, random forests have been replaced with gradient boosting models of decision trees (hereinafter referred to as MGB). Secondly, the process of constructing a Deep Forest, as well as the process of deriving predictions, should be changed so that each next layer clarifies the general prediction of the model, and does not build it anew. The general model is built according to the principle of gradient boosting, that is, so that each next layer predicts additional terms, which add up to the estimate of the target variable. Let us imagine a new model of Deep gradient boosting from L layers, as:

$$DBF_L(x) = DBF_{L-1}(x) + \gamma M_L(x, F_0(x), \dots, F_{L-1}(x)),$$

where M_i – model of the i layer; γ – learning rate; F_i – algorithm for generating features in layer i . The scheme of the Deep gradient boosting model is shown in Fig. 2.

Let the loss functional l be given and a Deep gradient boosting model with an $(L - 1)$ layer is already constructed. To minimize the loss functional, the next layer is constructed in such a way as to approximate the residual terms for each point of the training set:

$$r_i^{(L)} = \left| -\frac{\partial l(z, y_i)}{\partial z} \right|_{z = DBF_{L-1}(x_i)}.$$

As an input, similar to a deep forest, the next layer uses a set of features from the previous layers obtained using a feature generation algorithm:

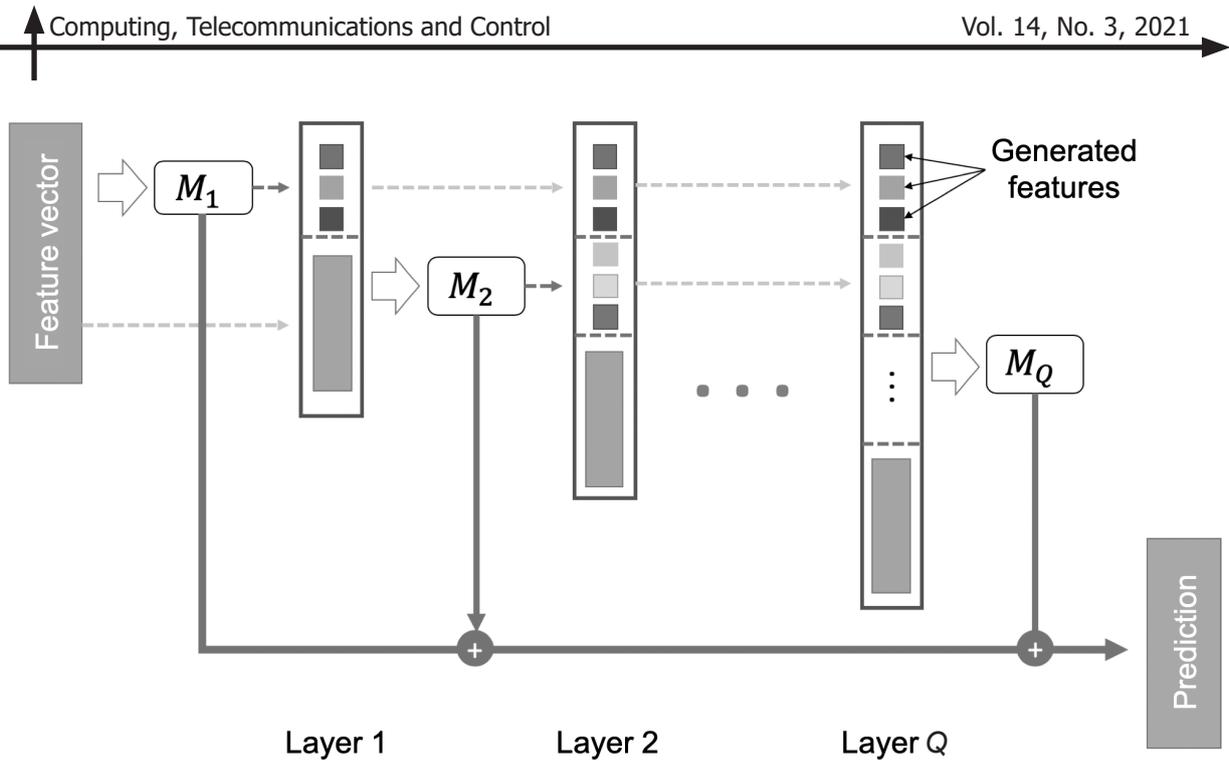


Fig. 2. Deep gradient boosting model scheme

$$I_M = (x_i, F_0(x_i), \dots, F_{M-1}(x_i)).$$

MGB M_L is trained to predict the residual terms from the generated features.

Thus, the general algorithm for constructing a Deep boosting model consists of the following steps:

1. Building a zero layer, like an MGB composition, namely training several different MGB.
2. Calculation of the residual terms of r_i – partial derivatives of the error function.
3. Generating a set of features based on predictions.
4. Formation of a new feature space: concatenation of the generated feature vectors with the feature vectors of the previous layers.

5. Construction of a new layer approximating the residual terms using the resulting feature space.

6. If the stop condition is not met, go to step 2.

This algorithm is general and leaves it possible to determine the algorithm for generating features and the stop condition. So, if you omit the feature generation procedure (that is, set a feature generation algorithm that matches zero-length vectors to predictions), and choose a decision tree construction algorithm as the algorithm for constructing a new layer, you can get a traditional gradient boosting algorithm for decision trees. The stop condition can be set as a condition for the number of layers M to exceed a certain constant. You can also set a threshold for the value of the empirical risk functional, after which the algorithm is stopped, or use a plateau detection criterion: during a certain number of iterations, the fluctuation of the empirical risk functional (the difference between the maximum and minimum values) should not exceed the constants. Both of the latter criteria can also be applied with a validation sample – a separate data set used for quality assessment that does not overlap with the training set.

The predictions of the MGB layer are used as the generated features, while the cumulative prediction of the entire composition is not taken into account. The features generated by layer i are defined as:

$$F_i(x) = [M_i^1(x, F_0(x), \dots, F_{i-1}(x)), \dots, M_i^K(x, F_0(x), \dots, F_{i-1}(x))]^T.$$

By analogy with a deep forest, each layer uses not one MGB, but a composition of independent models M_i^j , in order to obtain a more informative feature space. The layer prediction is calculated by averaging the predictions of the composition models. In this case, the method of generating features is based on the predictions of a set of composition models. The layer prediction is calculated as:

$$P_i(t) = \frac{1}{K} \sum_{j=1}^K M_i^j(t).$$

Potentially, the disadvantage of such formation of new features is the inability to build a decisive rule that splits a set of data according to the criterion of consistency of model predictions in a layer, as well as a high correlation of features. To eliminate these shortcomings, it is possible to transform the obtained features into an analog of the class frequency vector. For the convenience of the presentation, without detracting from the generality, we will assume that the value of each feature lies in the interval $[0, 1]$, and both boundaries of the interval are reached. If this is not the case, the values of all features should be pre-normalized, independently of each other at the model construction stage, and at the prediction stage, the same constants should be used for normalization as during training. The interval is divided into C disjoint half-intervals of equal length. Each prediction of the layer model is associated with a vector of C components:

$$u_i^j(x) = \left\langle 1 - \left| (F_i(x))_j - \frac{t}{C} \right| \right\rangle_{t=1}^C.$$

The vectors corresponding to each layer model are combined by averaging:

$$u_i(x) = \frac{1}{K} \sum_{j=1}^K u_i^j(x).$$

A normalized exponential function (SoftMax) with temperature θ is applied to the resulting vector, calculated as:

$$(\sigma(z, \theta))_i = \frac{\exp\left(\frac{z_j}{\theta}\right)}{\sum_j \exp\left(\frac{z_j}{\theta}\right)}.$$

Thus, the obtained analog of the probability vector of classes:

$$\hat{u}_i(x) = \sigma(u_i(x)).$$

The sum of all the elements of the vector $\hat{u}_i(x)$ will be equal to 1, and the value of each element will lie in the interval $[0, 1]$. If the predictions of the layer models are equal, the element of the vector $\hat{u}_i(x)$ corresponding to the predictions has the maximum value. By varying the temperature θ , it is possible to reduce or increase the influence of the predictions of individual models of the layer: the higher the temperature, the closer the elements of the vector are to $\frac{1}{K}$. The lower the temperature, the closer the value of one of the elements of the vector is to 1. The obtained representation of the features allows us to build decisive rules based on both the predicted values and the consistency of the layer models.

Let's pay attention to the physical meaning of the functions that are approximated by each layer: partial derivatives of the loss function. If the loss function is convex, then if its partial derivative at point x is zero, then the minimum of the loss function is reached at x . This means that no further iterations are required for the set of all such points Q . Each next layer of the model will need to predict values close to zero for Q points. If the predictions are different from zero, the difference will need to be compensated by the next layer after it. Thus, firstly, extra work will be performed for a certain set of points, and, secondly, the optimal prediction obtained at a certain iteration can be further worsened by adding small terms with an error. To eliminate these shortcomings, we will filter out data instances based on the absolute value of predictions. Let's consider two ways to filter out data instances: by a fixed threshold value d , and by the quantile of the empirical distribution of the absolute values of the predictions of the layer. For the first method, on each layer, we will determine the set of points for dropping out layer i , depending on the input features I_j , as:

$$Q_i = \{I_j : |P_i(I_j)| < d\}.$$

The difficulty of using this approach is that it is impossible to determine the number of points to be eliminated in advance, and also with a fixed value of d , the result strongly depends on the scale of the data. However, the threshold value can be selected automatically and individually for the layer. To do this, we introduce the parameter p – the proportion of points that need to be filtered out on each layer and find the quantile of the empirical distribution of the absolute values of the predictions:

$$d_i = \arg \max_d \left\{ \#\{(I, r) \in D^{(i)} : |P_i(I)| < d\} \leq p \cdot \#D^{(i)} \right\}.$$

The set of points for dropping out, that is those that will not be used for further construction of the model, is located similarly. This approach will allow you to leave a known amount of data for training at each step in advance. Too large, close to one, p values can lead to a situation in which the size of the training set will be insufficient to build an accurate model, and / or an increase in the value of the risk functional, due to too early sifted points.

In order to avoid situations in which points are eliminated too early, that is, inflated d_i values, it is possible to adjust the threshold values based on the known values of the residuals. To do this, we will introduce an additional threshold value d'_i , but we will calculate the quantile of the empirical distribution of the absolute values of the residuals, instead of predictions:

$$d'_i = \arg \max_d \left\{ \#\{(I, r) \in D^{(i)} : |r| < d\} \leq \eta \cdot \#D^{(i)} \right\}.$$

The adjusted threshold value will be calculated as:

$$\hat{d}_i = \min \{d_i, d'_i\}.$$

Thus, it remains possible to control the upper bound of the number of points that need to be eliminated on each layer, but if the model's prediction at a certain point is small due to an error, such a point will not affect the value of \hat{d}_i .

The choice of parameters of the basic models when constructing the composition has an extremely strong influence on the accuracy of the aggregate model. For example, if the approximate function depends on four features, and the maximum depth of the decision trees used in MGB is two, then it will be impossible to build an exact approximation using such an MGB. If the depth of the trees is too large, there may be a negative effect of retraining. This means that it is necessary to select the most suitable parameters of the model. Since each layer solves its own separate machine-learning task (both the approximate function and the set of features are different for each layer), the most suitable parameters should be selected for each

layer separately. To do this, we use a search of parameter values with cross-validation: a grid of parameters is set, including the maximum depth of the tree, the number of MGB iterations and the learning rate, and a combination of parameters is selected that gives the smallest error of cross-validation (cross-validation). All the basic models in the layer are trained using the locally obtained optimal parameters. Note that a similar improvement can also be applied to the classical gradient boosting algorithm for decision trees, but it is more justified in the case of the proposed deep gradient boosting, since it changes not only the objective function, but also the set of features at each iteration.

Experiments

For comparison with existing MGB and random forests, implementations from the “Scikit-Learn” package were used. To assess the quality of predictions in each experiment, the data were divided into two groups randomly: the training set (75 % of the original sample) and the test set (25 % of the original sample). Next, training was performed on the training set and quality assessment was performed on the test set. For each method, the experiment was repeated 100 times. Since MGB and random forest require the selection of parameters, such as the depth of trees, a procedure for automatically selecting the most suitable parameters with cross-checking was used. To do this, the training set was divided into 5 parts and such parameters were selected for which the average accuracy during cross-checking was maximum. After that, the model was built anew with the selected parameters using the entire training set. For a random forest, the maximum depth of the decision tree was chosen – an integer value from 2 to 7, or an unlimited depth of the tree, as well as the number of decision trees: 100 or 1000. For MGB, the parameters were selected in the same way, with the addition of the learning rate: 0.1, or 0.01. For deep boosting, the optimal values were not selected using cross-checking in advance.

To compare the methods on regression problems, the following data sets were used (Table 1):

- from the “StatLib” repositories and the “R” package: “ML housing data set”, “California housing data set”, “Diabetes”, “Longley’s Economic Regression Data”;
- from the Kaggle online data analysis competition platform: “House Prices: Advanced Regression Techniques”, based on the “Ames Housing data set”, with a large number of categorical features;
- simulated (artificially generated) data sets: “Friedman 1”, “Friedman 2”, “Friedman 3”, described in detail in [5]. And also “Regression” and “Sparse uncorrelated” from the “Scikit-Learn” package.

In Table 1, m denotes the number of features, and n is the number of observations.

On regression problems, the algorithms were compared: random forest, MGB, and the proposed Deep gradient boosting (with MGB on random decision trees). The standard error was used as the risk functional for all methods. Table 2 shows the values of standard errors (less – better) averaged over 100 experiments for data sets from repositories, and Table 3 – for artificial data sets, for:

- a random forest with the above-described method for selecting optimal hyperparameters (in the Table – the RF);
- extremely randomized trees with the above-described method of selecting optimal hyperparameters (in the Table – ERT);
- the classical model of gradient boosting with the above-described method of selecting optimal hyperparameters (in the Table – MGB);
- CatBoost gradient boosting models (in the Table – CatBoost);
- models of extreme gradient boosting XGBoost with the above-described method of selecting optimal hyperparameters (in the Table – XGBoost);
- the proposed Deep gradient boosting based on RMGB (in the Table – Deep RMGB);
- the proposed Deep gradient boosting based on extremely randomized trees (in the Table – Deep ERT).

As can be seen, the proposed algorithms allow us to build models with comparable or smaller root-mean-square errors than classical methods, as well as advanced widely used improvements to gradient boosting algorithms, such as CatBoost and XGBoost. The proposed deep gradient boosting RMGB in

Table 1

Description of regression data sets

| Title | Abbreviated designation | <i>m</i> | <i>n</i> |
|--|-------------------------|----------|----------|
| California housing data set | California | 8 | 20640 |
| House Prices: Advanced Regression Techniques | HouseART | 79 | 1460 |
| ML housing data set | Boston | 13 | 506 |
| Diabetes | Diabetes | 10 | 442 |
| Longley’s Economic Regression Data | Longley | 7 | 16 |
| Friedman 1 | Friedman 1 | 10 | 100 |
| Friedman 2 | Friedman 2 | 4 | 100 |
| Friedman 3 | Friedman 3 | 4 | 100 |
| Scikit-Learn Regression | Regression | 100 | 100 |
| Scikit-Learn Sparse uncorrelated | Sparse | 10 | 100 |

Table 2

Standard error on regression problems

| Method | California | HouseART | Boston | Diabetes | Longley |
|-----------|--|---------------------------------------|---------------------------------------|---------------------------------------|--|
| RF | 2.559×10^{-1} | 9.796×10^8 | 1.215×10^1 | 3.367×10^3 | 1.055×10^0 |
| ERT | 2.493×10^{-1} | 9.645×10^8 | 1.136×10^1 | 3.235×10^3 | 9.666×10^{-1} |
| MGB | 2.083×10^{-1} | 9.533×10^8 | 1.052×10^1 | 3.292×10^3 | 8.660×10^{-1} |
| CatBoost | 2.197×10^{-1} | 9.148×10^8 | 1.066×10^1 | 3.547×10^3 | 1.273×10^0 |
| XGBoost | 2.057×10^{-1} | 9.835×10^8 | 1.124×10^1 | 3.276×10^3 | 1.020×10^0 |
| Deep RMGB | 1.986×10^{-1} | 9.157×10^8 | 1.119×10^1 | 3.343×10^3 | 8.145×10^{-1} |
| Deep ERT | 2.472×10^{-1} | 9.220×10^8 | 1.082×10^1 | 3.315×10^3 | 9.045×10^{-1} |

Table 3

Standard error on artificial regression problems

| Method | Friedman 1 | Friedman 2 | Friedman 3 | Regression | Sparse |
|-----------|---------------------------------------|---------------------------------------|--|---------------------------------------|---------------------------------------|
| RF | 1.062×10^1 | 5.839×10^3 | 2.075×10^{-2} | 1.254×10^4 | 2.794×10^0 |
| ERT | 8.869×10^0 | 1.475×10^3 | 1.575×10^{-2} | 1.162×10^4 | 2.203×10^0 |
| MGB | 7.233×10^0 | 5.240×10^3 | 1.877×10^{-2} | 9.929×10^3 | 1.952×10^0 |
| CatBoost | 8.485×10^0 | 1.008×10^4 | 2.871×10^{-2} | 1.238×10^4 | 2.641×10^0 |
| XGBoost | 7.066×10^0 | 5.340×10^3 | 1.927×10^{-2} | 9.797×10^3 | 2.050×10^0 |
| Deep RMGB | 4.039×10^0 | 1.234×10^3 | 8.370×10^{-3} | 1.009×10^4 | 1.238×10^0 |
| Deep ERT | 7.916×10^0 | 1.490×10^3 | 1.326×10^{-2} | 1.141×10^4 | 1.779×10^0 |

more than half of the cases (on the considered data sets) allows us to get better results than other algorithms. The lower accuracy of Deep gradient boosting ERT, compared to Deep gradient boosting RMGB,

is consistent with the assumptions about the need to use MGB in deep boosting layers, instead of a random forest or ERT.

Conclusion

The algorithm of Deep gradient boosting is proposed, which allows to build deep models without using the algorithm of back propagation of the error for solving regression problems. It combines elements of representation learning characteristic of neural networks, due to which more accurate models are built in comparison with traditional gradient boosting and random forest, while maintaining the advantages of decision tree compositions, such as the ability to build models on small training sets and the ability to process categorical features. The described approach to generating features allows you to create more informative representations of input data improving the quality of the model. Experiments conducted on known regression problems and artificial data show the advantages of the proposed algorithm in comparison with existing algorithms.

This work is supported by the Russian Science Foundation under grant 21-11-00116, <https://rscf.ru/en/project/21-11-00116/>

REFERENCES

1. Ferreira A.J., Figueiredo M.A.T. Boosting algorithms: A review of methods, theory, and applications. *Ensemble Machine Learning*. Springer, Boston, MA, 2012, Pp. 35–85.
2. Rokach L. Ensemble-based classifiers. *Artificial Intelligence Review*, 2010, Vol. 33, No. 1-2, Pp. 1–39.
3. Woźniak M., Graña M., Corchado E. A survey of multiple classifier systems as hybrid systems. *Information Fusion*, 2014, Vol. 16, Pp. 3–17.
4. Zhou Z.-H. *Ensemble methods: Foundations and algorithms*. CRC Press, Boca Raton, 2012.
5. Breiman L. Bagging predictors. *Machine Learning*, 1996, Vol. 24, No. 2, Pp. 123–140.
6. Wolpert D.H. Stacked generalization. *Neural Networks*, 1992, Vol. 5, No. 2, Pp. 241–259.
7. Zhou Z.-H., Feng J. Deep Forest: Towards an alternative to Deep neural networks. *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI'17)*, AAAI Press, Melbourne, Australia, 2017, Pp. 3553–3559.
8. Wang C., Lu N., Cheng Y., Jiang, B. Deep Forest based multivariate classification for diagnostic health monitoring. *2018 Chinese Control and Decision Conference (CCDC)*, IEEE, 2018, Pp. 6233–6238.
9. Yang F., Xu Q., Li B., Ji Y. Ship detection from thermal remote sensing imagery through region-based Deep Forest. *IEEE Geoscience and Remote Sensing Letters*, 2018, Vol. 15, No. 3, Pp. 449–453.
10. Zheng W., Cao S., Jin X., Mo S., Gao H., Qu Y., Zhu Y. Deep Forest with local experts based on elm for pedestrian detection. *Pacific Rim Conference on Multimedia*, Springer, Cham, 2018, Pp. 803–814.
11. Geurts P., Ernst D., Wehenkel L. Extremely randomized trees. *Machine Learning*, 2006, Vol. 63, No. 1, Pp. 3–42.
12. Utkin L.V., Ryabinin M.A. A siamese Deep Forest. *Knowledge-Based Systems*, 2018, Vol. 139, Pp. 13–22.
13. Freund Y., Shapire R. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.*, 1997, Vol. 55, Pp. 119–139.
14. Pang M., Ting K.M., Zhao P., Zhou Z.H. Improving Deep Forest by confidence screening. *2018 IEEE International Conference on Data Mining (ICDM)*, IEEE, 2018, Pp. 1194–1199.
15. Utkin L., Konstantinov A., Meldo A., Ryabinin M., Chukanov V. A Deep Forest improvement by using weighted schemes. *2019 24th Conference of Open Innovations Association (FRUCT)*, IEEE, 2019, Pp. 451–456.
16. Natekin A., Knoll A. Gradient boosting machines, a tutorial. *Frontiers in Neurobotics*, 2013, Vol. 7, P. 21.
17. Friedman J.H. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 2001, Pp. 1189–1232.

18. **Konstantinov A., Utkin L., Muliukha V.** Gradient boosting machine with partially randomized decision trees. *2021 28th Conference of Open Innovations Association (FRUCT)*, IEEE, 2021, Pp. 167–173.
19. **Shen W., Guo Y., Wang Y., Zhao K., Wang B., Yuille A.L.** Deep regression forests for age estimation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, Pp. 2304–2313.
20. **Konstantinov A.V., Utkin L.V.** A generalized stacking for implementing ensembles of gradient boosting machines. *Cyber-Physical Systems: Digital Technologies and Applications*, 2020, P. 3.

Received 05.07.2021.

СПИСОК ЛИТЕРАТУРЫ

1. **Ferreira A.J., Figueiredo M.A.T.** Boosting algorithms: A review of methods, theory, and applications // *Ensemble Machine Learning*. Springer, Boston, MA, 2012. Pp. 35–85.
2. **Rokach L.** Ensemble-based classifiers // *Artificial Intelligence Review*. 2010. Vol. 33. No. 1-2. Pp. 1–39.
3. **Woźniak M., Graña M., Corchado E.** A survey of multiple classifier systems as hybrid systems // *Information Fusion*. 2014. Vol. 16. Pp. 3–17.
4. **Zhou Z.-H.** Ensemble methods: Foundations and algorithms. CRC Press, Boca Raton, 2012.
5. **Breiman L.** Bagging predictors // *Machine Learning*. 1996. Vol. 24. No. 2. Pp. 123–140.
6. **Wolpert D.H.** Stacked generalization // *Neural Networks*. 1992. Vol. 5. No. 2. Pp. 241–259.
7. **Zhou Z.-H., Feng J.** Deep Forest: Towards an alternative to Deep neural networks // *Proc. of the 26th Internat. Joint Conf. on Artificial Intelligence*. AAAI Press, Melbourne, Australia, 2017. Pp. 3553–3559.
8. **Wang C., Lu N., Cheng Y., Jiang, B.** Deep Forest based multivariate classification for diagnostic health monitoring // *2018 Chinese Control and Decision Conf. IEEE*, 2018. Pp. 6233–6238.
9. **Yang F., Xu Q., Li B., Ji Y.** Ship detection from thermal remote sensing imagery through region-based Deep Forest // *IEEE Geoscience and Remote Sensing Letters*. 2018. Vol. 15. No. 3. Pp. 449–453.
10. **Zheng W., Cao S., Jin X., Mo S., Gao H., Qu Y., Zhu Y.** Deep Forest with local experts based on elm for pedestrian detection // *Pacific Rim Conference on Multimedia*. Springer, Cham, 2018. Pp. 803–814.
11. **Geurts P., Ernst D., Wehenkel L.** Extremely randomized trees // *Machine Learning*. 2006. Vol. 63. No. 1. Pp. 3–42.
12. **Utkin L.V., Ryabinin M.A.** A siamese Deep Forest // *Knowledge-Based Systems*. 2018. Vol. 139. Pp. 13–22.
13. **Freung Y., Shapire R.** A decision-theoretic generalization of on-line learning and an application to boosting // *J. Comput. Syst. Sci.* 1997. Vol. 55. Pp. 119–139.
14. **Pang M., Ting K.M., Zhao P., Zhou Z.H.** Improving Deep Forest by confidence screening // *2018 IEEE Internat. Conf. on Data Mining. IEEE*, 2018. Pp. 1194–1199.
15. **Utkin L., Konstantinov A., Meldo A., Ryabinin M., Chukanov V.** A Deep Forest improvement by using weighted schemes // *2019 24th Conf. of Open Innovations Association (FRUCT)*. IEEE, 2019. Pp. 451–456.
16. **Natekin A., Knoll A.** Gradient boosting machines, a tutorial // *Frontiers in Neurorobotics*. 2013. Vol. 7. P. 21.
17. **Friedman J.H.** Greedy function approximation: A gradient boosting machine // *Annals of Statistics*. 2001. Pp. 1189–1232.
18. **Konstantinov A., Utkin L., Muliukha V.** Gradient boosting machine with partially randomized decision trees // *2021 28th Conf. of Open Innovations Association (FRUCT)*. IEEE, 2021. Pp. 167–173.
19. **Shen W., Guo Y., Wang Y., Zhao K., Wang B., Yuille A.L.** Deep regression forests for age estimation // *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*. 2018. Pp. 2304–2313.
20. **Konstantinov A.V., Utkin L.V.** A generalized stacking for implementing ensembles of gradient boosting machines // *Cyber-Physical Systems: Digital Technologies and Applications*. 2020. P. 3.

Статья поступила в редакцию 05.07.2021.

THE AUTHOR / СВЕДЕНИЯ ОБ АВТОРЕ

Konstantinov Andrei V.
Константинов Андрей Владимирович
E-mail: andrue.konst@gmail.com

© Санкт-Петербургский политехнический университет Петра Великого, 2021

DOI: 10.18721/JCSTCS.14302
УДК 004

DYNAMIC ENERGY CONSUMPTION RATIONING BASED ON MACHINE LEARNING ALGORITHMS FOR OIL REFINING TASKS

N.S. Kudriashov

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

Energy consumption rationing is necessary for high-quality production planning, and allows optimizing their use. This paper provides an analysis of various approaches to building a model of energy consumption, describes their limitations and new approaches to dynamic rationing. As the object of modeling the ELOU-AVT-6 (CDU/VDU-6) unit has been taken. Such units are intended for desalination and primary fractionation of oil. Functional requirements for the algorithms have been formed, based on real production needs. As the solution, models based on machine learning algorithms have been analyzed. These algorithms include CatBoost Regressor, Gradient tree boosting, Random Forest, ElasticNet and artificial neural networks. The analysis of the modeling results and comparison of the accuracy of the models is carried out. The paper also demonstrates a scenario of using a dynamic rationing model to analyze the causes of deviations of the actual consumption values from the planned ones.

Keywords: energy consumption rationing, machine learning, digital twin, oil refining, factor analysis.

Citation: Kudriashov N.S. Dynamic energy consumption rationing based on machine learning algorithms for oil refining tasks. Computing, Telecommunications and Control, 2021, Vol. 14, No. 3, Pp. 20–32. DOI: 10.18721/JCST-CS.14302

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

ДИНАМИЧЕСКОЕ НОРМИРОВАНИЕ ПОТРЕБЛЕНИЯ ЭНЕРГОРЕСУРСОВ ДЛЯ ЗАДАЧ НЕФТЕПЕРЕРАБОТКИ НА ОСНОВЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Н.С. Кудряшов

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

Нормирование потребления энергоресурсов необходимо для качественного планирования производства и позволяет рационализировать их использование. В статье приведен анализ различных подходов к построению модели потребления энергоресурсов, определены их недостатки и представлен новый подход к динамическому нормированию. В качестве объекта моделирования рассмотрен процесс суммарного потребления топлива для установки ЭЛОУ-АВТ-6, предназначенной для обессоливания и первичного фракционирования нефти. Сформированы функциональные требования к разрабатываемым алгоритмам, исходя из актуальных задач, диктуемых производством. В качестве решения рассмотрены модели на основе алгоритмов машинного обучения, такие как Catboost регрессор, Градиентный бустинг деревьев, Случайный лес, ElasticNet и искусственные нейронные сети. Проведен анализ результатов моделирования и сравнения точности моделей. Продемонстрирован сценарий использования модели динамического нормирования для анализа причин отклонения фактических значений потребления от плановых.

Ключевые слова: нормирование потребления энергоресурсов, машинное обучение, цифровой двойник, нефтепереработка, факторный анализ.

Ссылка при цитировании: Kudriashov N.S. Dynamic energy consumption rationing based on machine learning algorithms for oil refining tasks // Computing, Telecommunications and Control. 2021. Vol. 14. No. 3. Pp. 20–32. DOI: 10.18721/JCSTCS.14302

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>).

Introduction

High speed of technological growth and global trends towards the production digitalization and Industry 4.0 concepts dictate an increasing volume of requirements for the industry. As an illustrative example, the oil refining industry poses a task of active integration of new technologies to optimize technological processes, increase the quality of the final product and reduce its production costs. Consequently, in the last 10 years engineers have begun developing so-called digital twins of technological processes. A digital twin is a complex program entity based on accurate process models, statistical data, regulatory values and machine learning algorithms [1, 2].

Rationing of energy consumption is not an exception of such digitization processes. On the one hand, a digital twin of production allows us to implement various scenarios, such as: forecasting, retrospective analysis, simulation of various operating conditions of equipment, etc. On the other hand, the consumption rationing model should allow us to calculate the necessary and sufficient resource consumption for a particular scenario. The ability to analyze the actual deviations in energy consumption is a key feature, which can lead us to more rational cost management.

One of the key issues of building a model for the regulation of energy consumption, is a large number of influencing external and internal factors. Such factors impact often cannot be detected analytically. These factors include the parameters of the technological process, the parameters of raw input materials, meteorological conditions, time interval etc. All existing methods and algorithms for the regulation of energy consumption are often based on some sort of regulatory values and do not allow them to be effectively applied in such digitized scenarios, described previously [3].

The main goal of this work is to develop an approach and determine the methods of developing the models for dynamic rationing of energy consumption using the example of oil refining problems. For this case, the existing approaches to energy consumption rationing were analyzed and new methods, based on machine learning algorithms, were developed and approved. This new approach provides the ability to dynamically recalculate the rates of energy consumption within the process or environment changes.

Energy consumption rations

Energy consumption rations are the calculated values that characterize the maximum allowable expenditure of certain resources. At the same time, during the calculation, the operating conditions of the technological equipment and the environment at a particular moment of time are taken into account. The norms determine the calculation basis for planning the consumption of fuel and energy resources, and also allow you to control their expenditure and identify any potential saving reserves [4].

For industrial consumers, ration is an indicator of the planned consumption of some resources for the production of a unit of final product. There are two key groups of resources used as a subject of rationing: the main resources and the operation of production facilities.

In this paper, the rations of resources consumption, associated only with the provision of the main production process (fuel and energy resources) were considered. Therefore, we focus our attention on them. These values are evaluated in a form of a generalized indicator, expressed in Tones of Equivalent Fuel (TEF) per unit of production.

Let's consider the key existing methods for calculating consumption rations [5]:

- Experienced method;
- Computational and analytical method;
- Computational and statistical method.

Experienced method is, as the name suggests, an experiment, the results of which are used to form individual rations. A significant disadvantage of this approach is that the object of modeling must perform its work exclusively within the regimes provided at the stage of its testing [6]. Also, the process of such development norms entails large labor costs. It consists of methodology development, testing, analysis of results, etc. All this makes it difficult to replicate the solution. Another disadvantage is that applying this method makes it impossible to quickly revise the existing consumption rations due to changes in the process parameters.

Another common method, that allows us to ensure high accuracy of rationing, is the computational and analytical method. This method is based on a thorough study of technical regulations and design and engineering documentation [7]. During such process of rations calculation, a sequential division of the modeling object into separate aggregates is performed. After this, the engineer analyses such aggregates interaction. A significant disadvantage of this method is that the quality of the developed values is directly proportional to the quality and accuracy of the description of the object in the technical documentation. Thus, this approach may not take into account the real state of the object, which will entail a decrease in the quality of energy consumption rationing [8].

In turn, the computational and statistical method allows you to deal with the problem of inconsistency between the state of the object and its technical description. It provides the determination of rations, based on the reports data of the actual consumption of fuel and energy resources during the past periods. In other words, when applying this method, the values are interpolated by forming a function that characterizes the relationship between operating conditions and the amount of energy consumption [9].

For this work, during the building of the machine learning models, we employed the latter method first, and in particular the so-called analytical model. Mathematical formulation can be written as follows.

The process of energy consumption can be described as:

$$Y = f\left(X_1^{(in)}(T_1), \dots, X_N^{(in)}(T_N), X_1^{(out)}(T_{(N+1)}), \dots, X_K^{(out)}(T_{(N+K)}), \tau, \dot{X}\right) + \varepsilon(\tau),$$

where $X^{(in)}$ – interior industrial parameters values; $X^{(out)}$ – outer parameters values; T – different times-tamps, depending on the parameter; τ – current datetime value; \dot{X} – parameters, which are not kept in manufacture; ε – measurement error.

According to the amount of data we can use – our rationing model should have such mathematical description:

$$N(W) = f\left(\left(X_1^{(in)}(T_1), \dots, X_N^{(in)}(T_N), X_1^{(out)}(T_{(N+1)}), \dots, X_K^{(out)}(T_{(N+K)})\right), \tau\right).$$

As we see, the main goal of such model development is decreasing of the number of \dot{X} and to minimize the difference between such model values. In this case the mathematical description of the model development task will be as follow:

$$E = \arg \min_W \left(J(Y, N(W)) \right),$$

where J is an error function, more detailed description will be provided when it is referred to in the next paragraphs.

Another task for the development is the approach of choosing the $X^{(in)}$ and $X^{(out)}$ according to the analytical model.

One of the key advantages of this approach is its universality and applicability for various objects. Also, this approach is more flexible compared to its alternatives, described earlier.

This approach is based on the analytical construction of the energy consumption model. Therefore, its disadvantages include a high probability of a mistake in description of the nature of the function, the volume of which is directly proportional to the complexity of the modeled object.

The idea of building an analytical model is consonant with the task of machine learning. In the classical approach, at the stage of exploratory data analysis, we can involve specialists in energy consumption to form a list of factors, based on their knowledge about the modeling object [10]. However, this approach entails high labor costs. Thus, we have set the goal of reducing the use of knowledge about the subject area with no loss in quality: it will serve as a guarantee of successful replication of the developed approach.

Modeling object – CDU/VDU-6 unit

As it has been already mentioned, in this paper we took the oil refinery process as the object for energy consumption modeling. Modern oil separation involves piping crude oil through a sequence of hot furnaces. The resulting liquids and vapors are discharged into distillation units. Such products of the plant operation include diesel fuel, tar, fuel oil, kerosene, gasoline, etc.

In particular, we analyzed the data from the unit built according to the standard design called CDU/VDU-6. The feed-stock for the unit is crude oil coming from oil pumping stations.

To illustrate the refinery process, including CDU and VDU, Fig. 1 presents a simplified oil refinery schematic [11].

As it is shown in the figure, CDU/VDU unit separates crude oil into different products by boiling point differences and prepares feed for secondary processing units. Two main units for the CDU/VDU are CDU, an electrical desalting and oil dehydration unit, and the VDU, an atmospheric and vacuum distillation of oil unit.

During the process, the following products are obtained at this unit: fuel and liquefied gas, straight-run gasoline fractions, a fraction of straight-run diesel fuel, vacuum gas oil and tar. In addition, in the process of heat recovery, water vapor generates, which, looking ahead, is one of the targets predicted in this work.

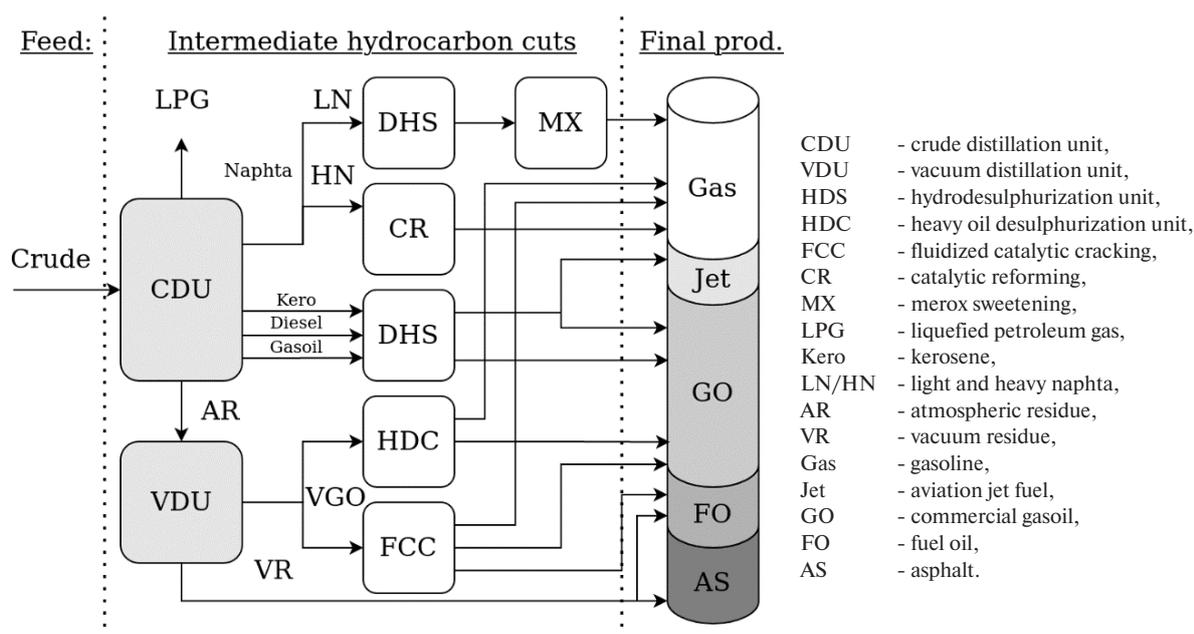


Fig. 1. Simplified schematic of an oil refinery

CDU/VDU-6 includes a gasoline stabilization unit and an intermediate tank farm for receiving, storing and dispensing raw materials of 6,000 cubic meters.

At the moment, for the modeled unit, an experimental method of rationing is used. The selection of norms is carried out from the corresponding reference book of norms with respect to the current mode of operation, month, and is multiplied by the actual load for raw materials. Such values are revised annually.

Such approach to rationing leads to ineffective plant management in terms of energy consumption. This is due to the low accuracy of the calculated consumption rates, since they do not take into account any real technological parameters, equipment degradation, environmental parameters, etc. Moreover, this approach does not allow us to analyze the factors and reasons of deviations in consumption from the calculated consumption rates. Such process is called factor analysis and it requires a more interpretative and transparent process of energy consumption.

Rationing model and digital twin

One of the main goals of using the dynamic rationing model is to ensure correct, well-grounded management of the technological process. This, in turn, can be achieved by the timely identification of deviations of the actual consumption from the planned one. Such approach would allow us to form operative corrective actions to improve the control quality.

To achieve this goal, it is necessary to establish a close interaction of the dynamic rationing model and a digital twin of the technological process [12]. Fig. 2 provides a more detailed description of such behavior.

In fact, there are two different scenarios of using the dynamic rationing models. The first one of them is the calculation of the planned y'' energy consumption rations, which are used as tasks for the plant operators. On the other hand, Fig. 2 shows a comparison between estimated consumption rations y' and actual consumption y'' . After reaching the point in time for which the consumption rate was planned, we can calculate its actual model value. This value should be close to the actual consumption of fuel and energy resources, otherwise, this algorithm will not be effective.

The algorithm of the deviation causes analysis (factor analysis) should demonstrate the degree of influence of each parameter on the total difference between planned and actual consumption. This approach allows us to carry out the retrospective analysis. With the knowledge we get from such analysis, we can adjust the operating plan in order to minimize expenses and costs.

To implement the algorithm for analysis of the causes of deviations, the dynamic rationing model must be interpretable. We must be able to numerically assess the degree to which each parameter affects the result. Such coefficients should be normalized in relation to the difference in consumption. Resulting delta values will characterize the fraction of delta justified by this or that parameter [13].

Thus, the main functional requirements for the development of the dynamic consumption rationing model are:

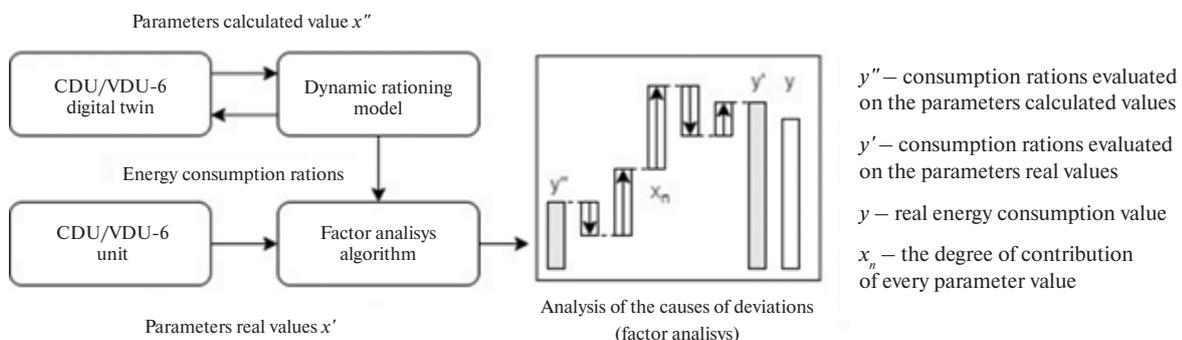


Fig. 2. Rationing model for the factor analysis issues

- interpretability of models or the possibility of using algorithms for analyzing the causes of deviations;
- the standardization error should not exceed 2 %, which is due to the absolute error of measuring devices for fuel and energy resources consumption;
- adaptability of the solution: the ability to revise consumption rates when critical changes in the process are detected;
- the model should take into account the degree of inertia of the process, since some parameters can contribute to consumption with different delta over time.

Initial modeling data

To start the dynamic rationing model development, it is necessary to assess the quality of the existing data and prepare them for transfer to the algorithm. For machine learning algorithms bad quality can be crucial, so we need to preprocess the data. Collecting, storage and primary generalization of data for the CDU/VDU-6 unit is carried out by PI Systems utilities. At the first step, ETL procedure was set up to download the data from the server.

For the analysis, a data interval of one and a half year was selected. Such choice is justified by the necessity of having a volume of data covering the entire annual cycle of the installation, as well as some reserve of data for validation and testing. Downloaded data contains technological information (unit temperatures, pressures, etc.), production data (plant load, quality characteristics), as well as environmental parameters (temperature, wind directions, etc.). The total fuel consumption was taken as an energy resource for rationing modeling.

The process of the control and monitoring of energy consumption is implemented in the context of hourly averaged values. Thus, the dynamic rationing model should also form the hourly average indicators of energy consumption. However, the data from the source were downloaded with a sampling rate of 1 minute, cleaned and averaged within an hour. This is due to the fact, that the data can contain a large number of anomalies. Such anomalies can be related to some functional failures of measuring instruments, breakdowns and production interruptions. In this case, it is necessary to filter the data somehow. For these purposes, the EllipticEnvelope algorithms [14], a high-pass filter and a moving average filter, were applied to achieve the best anomaly detection efficiency:

$$\hat{Y} = \theta(Y) * \frac{1}{60} \sum_{i=60} \vartheta(Y_{-i}),$$

where \hat{Y} – the resulting value of hourly consumption; θ – EllipticEnvelope operation giving [0, 1] values; ϑ – high-pass filter operation.

Fig. 3 shows an example of the preprocessed fuel consumption data.

Also, during the preparation, the data were preprocessed in order to reduce their volume with minimal loss in quality. Unfortunately, the use of dimensionality reduction methods (Principal component analysis, tSNE, etc.) is not further allowed for data interpretability, otherwise it would significantly complicate such an algorithm.

In such situation, to reduce the amount of data, a correlation analysis was carried out. The Spearman and Pearson coefficients were analyzed and the values were discarded according to the following criteria:

- parameters, the correlation coefficient of which with the target is more than 0.9, in order to avoid target leaks;
- to combat multicollinearity, all values with cross-correlations greater than 0.9 were also discarded.

Finally, the initial data volume contained 1083 parameters with 780 thousand values each. However, after their preprocessing, they were reduced to 562 parameters of 3 thousand averaged hourly values. For this publication, the data have been anonymized and replaced with tag_n for parameters and y for consumption not to violate the contribution rules. All the data were scaled from 0 to 1 for the same reason.

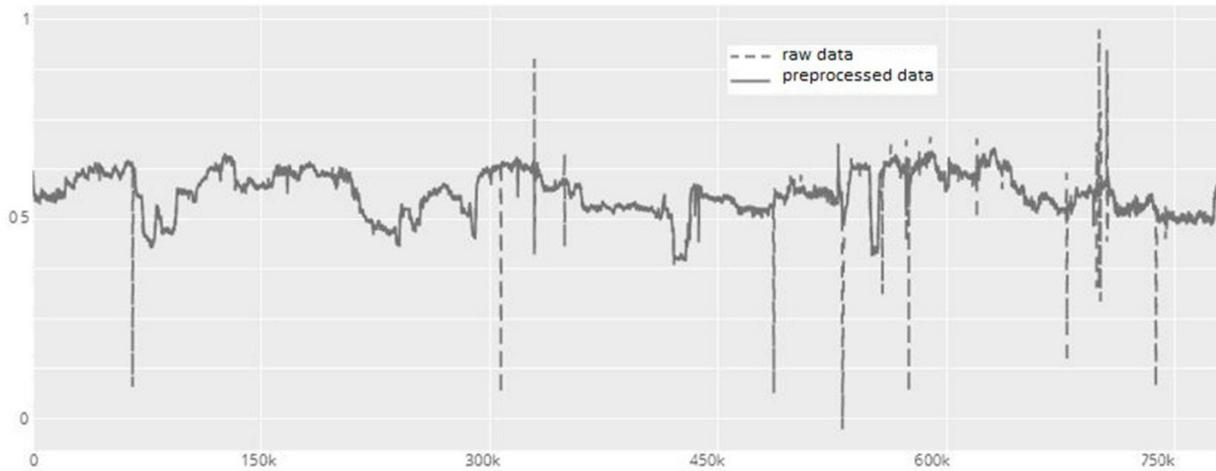


Fig. 3. The results of the consumption data preprocessing

Dynamic energy consumption rationing model

At the very beginning of the model development, the prepared dataset was divided into training, validation and test samples. The volume of the training dataset is 12 months of operation of the unit and the inner processes. Such dataset is used to train the model. The validation set serves to determine the generalizing ability of the model and is used to select its parameters. The validation process is based on the TimeSeries-Split cross-validation method with a sliding window of 1 month [15]. A test dataset with the length equal to 1 month is used to evaluate the simulation results.

Also, one of the important features of the modeled process is that some parameters can affect power consumption with some time delta. To handle such time deltas, we inserted a time lag parameters $\text{tag_n}^{(tk)}$ into our model. For this work, we have considered k value equal to 8, which implies that the maximum time lag can be 8 hours. Such delay is caused by the characteristics of the process, but have to be adjusted in future work.

In this research, the following machine learning approaches to the construction of a dynamic normalization model were considered [16]:

- linear models;
- models based on tree boosting algorithms;
- one-dimensional convolutional neural networks.

The Linear Regression algorithm is the most obvious and simplest solution to such problems. It is usually applied to create the baseline solution of the machine learning tasks. This algorithm is mainly suitable for linearly separable data. Although, it still often gives a satisfactory result for solving many real-world problems. While the generalizing ability of the model is usually weak, a significant advantage of the algorithm is its interpretability, which is crucial for the described task.

Simple Linear Regression was applied, but it showed low generalization ability. In this case, the ElasticNet algorithm was tested, which is a linear model with L1 and L2 regularizations. Fig. 4 shows a generalized diagram of the linear ElasticNet model and the result of forecasting consumption on a test sample.

The resulting vector of weights was pruned in case of finding the correct T (timestamp) for each parameter. After that, the model had been retrained. This operation can be described as follows.

From initial equation of Linear Regression:

$$N = W_{11}X_1(t_0) + \dots + W_{1n}X_1(t_{-n}) + \dots + W_{k1}X_k(t_0) + \dots + W_{kn}X_1(t_{-n}) + b,$$

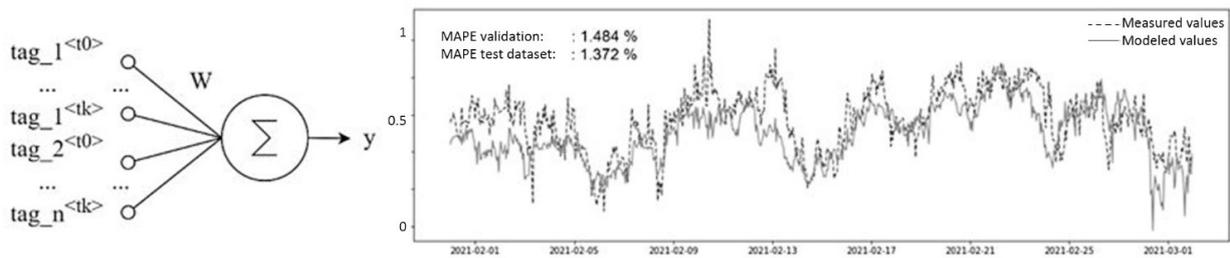


Fig. 4. Linear model of dynamic consumption rationing

where k is the number of parameters, n is the depth of time analysis, and b is a bias value. Model was transformed into:

$$N(W) = \sum_n W_i X_i(T_i) + b,$$

where T refers to different timestamps depending on the chosen parameter X . The same approach has been applied in all the models.

Another tested group of machine learning algorithms for consumption prediction is a family of tree-based boosting algorithms. Such algorithms are based on the concept of constructing a group of weak regressors to solve a more complex problem. In this work, a regressor model from the Catboost library was chosen. This algorithm has good generalization ability. Also, this algorithm has built-in functionality for analyzing the feature importance, based on the algorithm for analyzing and calculating trees (function `get_feature_importance`) [17].

Another realization of the boosting algorithms tested in this work is the classic Gradient boosting from scikit-learn library, which was applied to compare the results of modeling. However, the Random Forest algorithms were tested to compare the boosting to bagging approaches, but their further study was stopped due to the low accuracy. To solve the set task of analyzing lags in time, they are also added as parameters to the model. The resulting algorithm and mathematical description of the Catboost model development is described in Fig. 5.

An example of an element of the resulting tree and an assessment of the accuracy of the model on a test sample is shown in Fig. 6.

Another approach studied in this work is a family of artificial neural network (ANN) algorithms. Neural networks are accepted as more complex algorithms, due to huge number of parameters, that can be tuned for such models. In this way, ANNs have a high generalizing ability and are more flexible to use. Applying this group of algorithms provides us with a more elegant and appropriate approach of taking time lags into account due to the peculiarities of configuring the network architecture. In comparison with the previous approaches, the significant disadvantage of ANN is the considerable complexity and variability of its development process. At the same time, those models have a very complex structure, which complicates the process of interpreting the model. It entails the need of an analysis based on Shapley vectors, which can give us the appropriate information for the analysis of deviations cause [18].

In this work two architectures were tested as the solution of the research task. The first model is a simple sequential fully-connected neural network, for which all the time deltas were flattened into one input layer. The second tested model is the developed 1-dimensional convolutional neural network (CNN). The convolution is used to convolve timeseries data with several filters to define the time delta of the input-to-output dependency. It allows us to fold time intervals and transfer them to deeper layers of the neural network. Thus, the input to the model is not a vector of 562×8 values, but a matrix. Fig. 7 describes the generalized structure of the developed neural network.

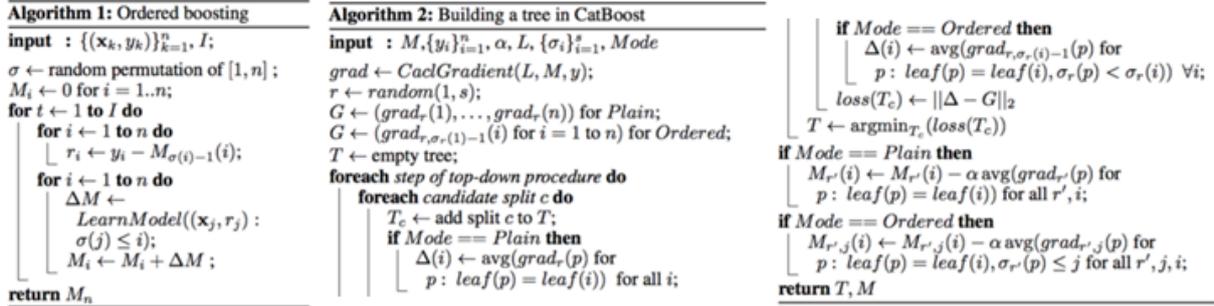


Fig. 5. CatBoost ordered boosting and tree building

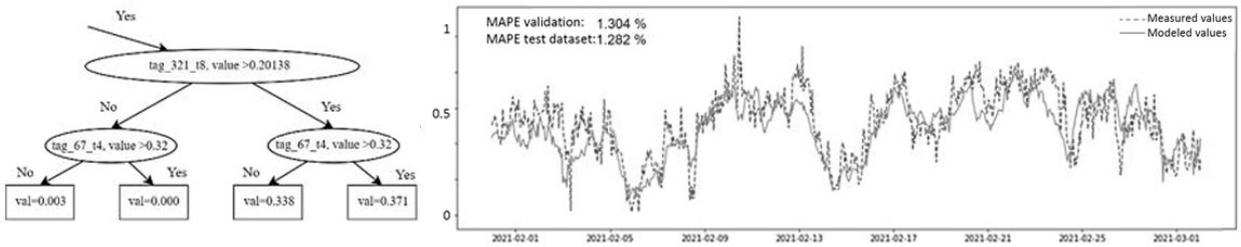


Fig. 6. CatBoost model of dynamic consumption rationing

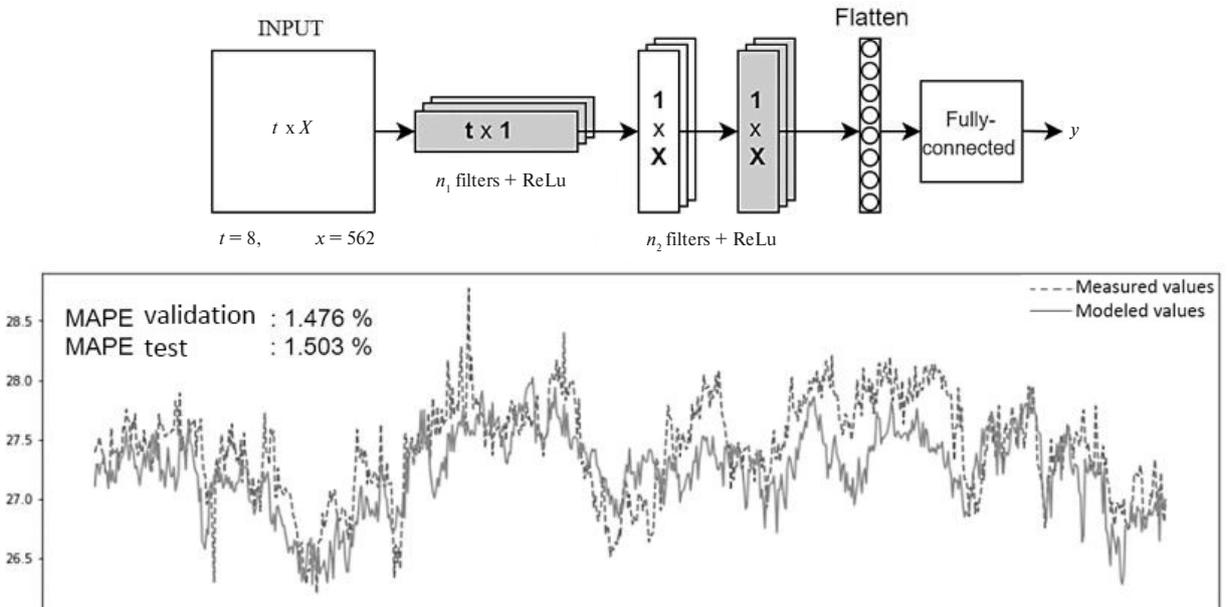


Fig. 7. CNN-based model of dynamic consumption rationing

Only three of the developed models produced satisfactory results on the described datasets. For this reason, only results of this three models consumption prediction was plotted in this publication. The analysis of the modeling results for all models is described in the next part of this work.

Modeling results analysis

During this research, a number of models of different nature have been developed. To test and validate those models, the mean absolute percentage error (MAPE) has been chosen.

The MAPE value calculation is similar to the absolute measurement error calculation [19]. Thus, we following formula was chosen to ensure the 2 % accuracy of developed algorithm:

$$M = \frac{1}{n} \cdot \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right|,$$

where n – number of measurements (data points) in the dataset; A_i – measured value; F_i – predicted value.

In this case, the task for the model development and evaluation can be rewritten as follows:

$$E = \arg \min_w \left(\frac{1}{n} * \sum_{i=1}^n \left| \frac{Y_i - N(W)_i}{Y_i} \right| \right).$$

Table 1 shows the results of evaluating the tested models. The models that performed the best were highlighted in the table.

Table 1

Model performance evaluation

| Model type | Model name | MAPE validation, % | MAPE test, % |
|----------------------------|-------------------|--------------------|--------------|
| Linear | ElasticNet | 1.484 | 1.372 |
| | Linear Regression | 1.013 | 2.422 (–) |
| Ensemble learning (Trees) | CatBoost | 1.304 | 1.282 |
| | Gradient Boosting | 2.021 (–) | 1.989 |
| | Random Forest | 3.199 (–) | 2.551 (–) |
| Artificial Neural Networks | MLPerceptron | 2.621 (–) | 2.305 (–) |
| | Convolutional NN | 1.476 | 1.503 |

Linear models performed well for this task. Classical linear regression shows the best generalizing ability on the validation set, however, the error increases dramatically on the test dataset. To solve this problem, L1 and L2 regularization was applied to the model, in the form of an ElasticNet model. This model performed well on both samples. A significant advantage of this model is also the simplicity of its interpretation in the form of regression coefficients.

On another hand, most of the ensemble learning models based on decision trees (Random Forest and Gradient Boosting) showed a slightly worse result for this task. Thus, Gradient Boosting and Random Forest showed a MAPE smaller than 2 %, which does not allow us to use them for efficient dynamic rationing.

As an exception, the model based on CatBoost Regressor showed the best result for all of the models. Also, this model can be easily interpreted by the tools of the CatBoost library itself. Further selection of hyperparameters and the use of other Gradient Boosting libraries can serve as possible ways of developing the model.

Models based on deep neural networks, generally, showed the worst prediction result among all models. The model developed on the basis of convolutional networks showed an acceptable result below 2 %. A feature of such models is a significantly larger number of parameters that can be configured. In this regard,

a further refinement of the architecture can serve as a possible way to improve the accuracy of the model. Also, the algorithm of deviation analysis is the most complex for this kind of models.

Thus, all three described models can be used to solve the problem. The priority at this stage is given to the model based on the Catboost regressor, since it showed a good generalization ability and low error rate. Moreover, this model allows us to calculate the significance of each parameter and use them to analyze the cause of deviations.

Conclusions

In this work, an approach to the dynamic rationing of energy consumption was developed. This approach is based on machine learning methods and algorithms of their interpretation.

Three main groups of algorithms have been tested, allowing us to implement dynamic normalization according to the described requirements: interpretability, adaptability, error lower than 2 %, taking into account the inertia of the process. Within these three groups, three models were built to achieve the modeling goals:

- Linear model based on EllipticEnvelope with regression coefficients as indicators of the significance of a parameter.
- CatBoost Regressor based model with parameter values (function `get_feature_importance`).
- Neural networks based model (one-dimensional convolutional neural network) with Shapley values.

Among those described models, the CatBoost Regressor based model produced the best result. This model performs the rationing of fuel and energy resources consumption with an accuracy of about 1.3 % MAPE. The built-in algorithm for calculating the significance of parameters allows us to characterize the reasons for the appearance of deviations in the data.

Despite the fact that the objectives of this study have been achieved, there is still a need of improvement of the accuracy for dynamic rationing models. Such improvement will cause the increasing of the economic benefits of the customer's manufacture.

There are two main ways to improve the obtained results. The first of them is further tuning of the parameters and hyperparameters of the developed models and expanding the training dataset. On the other hand, there are a number of algorithms that have not been tested in this work: Recurrent neural network based on GRU or LSTM; XGBoost library; LightGBM library.

REFERENCES

1. **Dozortsev V.M.** Tsifrovyye dvoyniki v promyshlennosti: genesis, sostav, terminologiya, tekhnologii, platformy, perspektivy. Chast 1. *Avtomatizatsiya v promyshlennosti*, 2020, No. 9, Pp. 3–11, (rus). DOI: 10.25728/avtprom.2020.09.01
2. **Kostenko D., Kudryashov N., Maystrishin M., Onufriev V., Potekhin V., Vasiliev A.** Digital twin applications: Diagnostics, optimisation and prediction. *Proceedings of the 29th DAAAM International Symposium*, 2018, Pp. 0574–0581, DOI: 10.2507/29th.daaam.proceedings.083
3. **Borky J.M.** Historical perspective: Energy monitoring and control systems. *The Military Engineer*, 2015, Vol. 107, No. 694, Pp. 91–92.
4. **Danilov O.L., Garyayev A.B., Yakovlev I.V.** *Energoberezheniye v teploenergetike i teplotekhnologiyakh*. Moscow: MEI Publ., 2017. (rus)
5. **Myo A.K., Portnov E.M., Kokin V.V.** Development of theoretical aspects of training systems construction for the basics of management and control over distributed power facilities. *2018 International Russian Automation Conference*, Sochi, IEEE, 2018. DOI: 10.1109/RUSAUTOCON.2018.8501758
6. **Gnatyuk V.I., Sheynin A.A.** ARS-Normirovaniye elektropotrebleniya infrastrukturykh ob'yektov. *Fedorovskiy Chteniya-2010. XL Vserossiyskaya Nauchno-Prakticheskaya Konferentsiya, Moscow, 16–19 nov. 2010*. Moscow: MEI Publ., 2010, Pp. 26–32. (rus)

7. **Gruntovich N.V., Kapanskiy A.A.** Raschetno-analiticheskiy metod normirovaniya raskhodov elektricheskoy energii v tekhnologicheskikh sistemakh vodosnabzheniya i vodootvedeniya. *Vestnik GGTU im. P.O. Sukhogo*, 2015, No. 2 (61). (rus)
8. **Kosharnaya Yu.V.** Razrabotka sistemy normirovaniya pokazateley elektropotrebleniya i otsenki ob'yemov energosberezheniya na primere metallurgicheskogo predpriyatiya. Chast 1. *Promyshlennaya Energetika*, 2015, No. 8, Pp. 13–17. (rus)
9. **Gofman I.V.** *Normirovaniye potrebleniya energii i energeticheskiye balansy promyshlennykh predpriyatiy*. Moscow: Energiya Publ., 1966. (rus)
10. **Tyralis H., Karakatsanis G., Tzouka K., Mamassis N.** Exploratory data analysis of the electrical energy demand in the time domain in Greece. *Energy*, 2017, No. 134, Pp. 902–918. DOI: 10.1016/j.energy.2017.06.074
11. **Galan A., De Prada C., Gutierrez G., Sarabia D.** *Implementation of RTO in a large hydrogen network considering uncertainty*. Springer, 2019. DOI: 10.1007/s11081-019-09444-3
12. **Kudriashov N., Markov S., Potekhin V.** Adaptive control system synthesis methods for complex manufacturing objects. *Proceedings of the 30th DAAAM International Symposium*. Publ. by DAAAM International, Vienna, Austria, 2019, Pp. 0493–0499. DOI: 10.2507/30th.daaam.proceedings.066
13. **Brown T.A.** *Confirmatory factor analysis for applied research*. 2nd ed. Guilford Press, 2015.
14. **Himeur Y., Ghanem Kh., Alsalemi A., Bensaali F., Amira A.** Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. *Applied Energy*, 2021, Vol. 287(3). DOI: 10.1016/j.apenergy.2021.116601
15. **Ojala M., Garriga G.C.** Permutation tests for studying classifier performance. *Proceedings of the 2009 9th IEEE International Conference on Data Mining*, 2009, Pp. 908–913. DOI: 10.1109/ICDM.2009.108
16. **Prince Waqas Khan, Yung-Cheol Byun, Sang-Joon Lee, Dong-Ho Kang, Jin-Young Kang, Hae-Su Park.** Machine learning-based approach to predict energy consumption of renewable and nonrenewable power sources. *Energies*, 2020, No. 13(18). DOI: 10.3390/en13184870
17. **Prokhorenkova L., Gusev G., Vorobev, Dorogush A.V., Gulin A.** *CatBoost: Unbiased boosting with categorical features*, V5 2019.
18. **Molnar Ch.** *Interpretable machine learning. A guide for making black box models explainable*, 2021.
19. **de Myttenaere A., Golden B., Le Grand B., Rossi F.** Mean absolute percentage error for regression models. *Neurocomputing*, 2016 arXiv:1605.02541.

Received 23.05.2021.

СПИСОК ЛИТЕРАТУРЫ

1. **Дозорцев В.М.** Цифровые двойники в промышленности: генезис, состав, терминология, технологии, платформы, перспективы. Ч. 1 // Автоматизация в промышленности. 2020. № 9. С. 3–11. DOI: 10.25728/avtprom.2020.09.01
2. **Kostenko D., Kudryashov N., Mastrishin M., Onufriev V., Potekhin V., Vasiliev A.** Digital twin applications: Diagnostics, optimisation and prediction // Proc. of the 29th DAAAM Internat. Symp. 2018. Pp. 0574–0581. DOI: 10.2507/29th.daaam.proceedings.083
3. **Borky J.M.** Historical perspective: Energy monitoring and control systems // *The Military Engineer*. 2015. Vol. 107. No. 694. Pp. 91–92.
4. **Данилов О.Л., Горяев А.Б., Яковлев И.В.** Энергосбережение в теплоэнергетике и теплотехнологиях. М.: ИД МЭИ, 2017.
5. **Myo A.K., Portnov E.M., Kokin V.V.** Development of theoretical aspects of training systems construction for the basics of management and control over distributed power facilities // 2018 Internat. Russian Automation Conf. Sochi, IEEE, 2018. DOI: 10.1109/RUSAUTOCON.2018.8501758

6. **Гнатюк В.И., Шейнин А.А.** ARS-нормирование электропотребления инфраструктурных объектов // Федоровские чтения-2010. XL Всерос. науч.-практ. конф. с междунар. уч. с элементами научной школы для молодежи. М.: ИД МЭИ, 2010. С. 26–32.
7. **Грунтович Н.В., Капанский А.А.** Расчетно-аналитический метод нормирования расходов электрической энергии в технологических системах водоснабжения и водоотведения // Вестник ГГТУ им. П.О. Сухого. 2015. № 2 (61).
8. **Кошарная Ю.В.** Разработка системы нормирования показателей электропотребления и оценки объемов энергосбережения на примере металлургического предприятия. Ч. 1 // Промышленная энергетика. 2015. № 8. С. 13–17.
9. **Гофман И.В.** Нормирование потребления энергии и энергетические балансы промышленных предприятий. М.: Энергия, 1966.
10. **Tyralis H., Karakatsanis G., Tzouka K., Mamassis N.** Exploratory data analysis of the electrical energy demand in the time domain in Greece // Energy. 2017. No. 134. Pp. 902–918. DOI: 10.1016/j.energy.2017.06.074
11. **Galan A., De Prada C., Gutierrez G., Sarabia D.** Implementation of RTO in a large hydrogen network considering uncertainty. Springer, 2019. DOI: 10.1007/s11081-019-09444-3
12. **Kudriashov N., Markov S., Potekhin V.** Adaptive control system synthesis methods for complex manufacturing objects // Proc. of the 30th DAAAM Internat. Symp. Publ. by DAAAM International, Vienna, Austria, 2019, Pp. 0493–0499. DOI: 10.2507/30th.daaam.proceedings.066
13. **Brown T.A.** Confirmatory factor analysis for applied research. 2nd ed. Guilford Press, 2015.
14. **Himeur Y., Ghanem Kh., Alsalemi A., Bensaali F., Amira A.** Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives // Applied Energy. 2021. Vol. 287(3). DOI: 10.1016/j.apenergy.2021.116601
15. **Ojala M., Garriga G.C.** Permutation tests for studying classifier performance // Proc. of the 2009 9th IEEE Internat. Conf. on Data Mining. 2009. Pp. 908–913. DOI: 10.1109/ICDM.2009.108
16. **Prince Waqas Khan, Yung-Cheol Byun, Sang-Joon Lee, Dong-Ho Kang, Jin-Young Kang, Hae-Su Park.** Machine learning-based approach to predict energy consumption of renewable and nonrenewable power sources // Energies. 2020. No. 13(18). DOI: 10.3390/en13184870
17. **Prokhorenkova L., Gusev G., Vorobev, Dorogush A.V., Gulina A.** CatBoost: Unbiased boosting with categorical features. V5 2019.
18. **Molnar Ch.** Interpretable machine learning. A guide for making black box models explainable. 2021.
19. **de Myttenaere A., Golden B., Le Grand B., Rossi F.** Mean absolute percentage error for regression models // Neurocomputing. 2016 arXiv:1605.02541.

Статья поступила в редакцию 23.05.2021.

THE AUTHOR / СВЕДЕНИЯ ОБ АВТОРЕ

Kudriashov Nikita S.
Кудряшов Никита Сергеевич
E-mail: niki94@yandex.ru

DOI: 10.18721/JCSTCS.14303

УДК 656.1, 004.94

IMPROVING THE EFFICIENCY OF TRAFFIC MANAGEMENT IN A METROPOLIS BASED ON COMPUTER SIMULATION

*A.S. Svistunova, D.S. Khasanov*St. Petersburg Federal Research Center of the Russian Academy of Sciences,
St. Petersburg, Russian Federation

This article considers an algorithm for traffic management. The process of simulation modeling of traffic flow at a road intersection in a metropolis is described. The optimization of the obtained simulation model is performed in the software system of simulation maintenance Anylogic. The application of the developed model is shown on the real example of a megalopolis road intersection. We considered the classification of highways, and analyzed the built model on the basis of the existing megalopolis crossroads, which allowed us to obtain the data comparable with the existing system. We considered and adopted methods to solve the high traffic problem by re-organizing the intersection using simulation modeling. The final system showed a 27 % throughput growth in the whole system and a 40 % growth in the main direction without traffic jams. The analysis shows the prospects of using the proposed simulation model to study real-world traffic flow management processes in order to study their behavior.

Keywords: megalopolis, control, simulation modeling, algorithm, model, transport stream, car roads.

Citation: Svistunova A.S., Khasanov D.S. Improving the efficiency of traffic management in a metropolis based on computer simulation. Computing, Telecommunications and Control, 2021, Vol. 14, No. 3, Pp. 33–42. DOI: 10.18721/JCST-CS.14303

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ УПРАВЛЕНИЯ ТРАНСПОРТНЫМИ ПОТОКАМИ МЕГАПОЛИСА НА ОСНОВЕ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ

*А.С. Свистунова, Д.С. Хасанов*Санкт-Петербургский Федеральный исследовательский центр РАН,
Санкт-Петербург, Российская Федерация

Рассмотрен алгоритм управления транспортными потоками. Описан процесс имитационного моделирования транспортного движения на перекрестке дорог в мегаполисе. Оптимизация полученной имитационной модели выполнена в программной системе имитационного обслуживания Anylogic. Применение разработанной модели показано на реальном примере автодорожного перекрестка мегаполиса. Рассмотрена классификация автомобильных дорог, на основании которой проведен анализ модели, построенной по существующему перекрестку мегаполиса, из которого были получены данные, сопоставимые с данными существующей системы. Изучены и приняты методы решения проблемы высокого трафика путём реорганизации перекрестка при помощи имитационного моделирования. Итоговая система показала прирост пропускной способности на 27 % во всей системе и 40 % прироста по основному направлению без образования заторов. Выполненный анализ обосновал перспективность использования предложенной имитационной модели для исследования реальных процессов управления транспортными потоками с целью изучения их поведения.

Ключевые слова: мегаполис, управление, имитационное моделирование, алгоритм, модель, транспортный поток, автомобильные дороги.

Ссылка при цитировании: Svistunova A.S., Khasanov D.S. Improving the efficiency of traffic management in a metropolis based on computer simulation // Computing, Telecommunications and Control. 2021. Vol. 14. No. 3. Pp. 33–42. DOI: 10.18721/JCSTCS.14303

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>).

Introduction

At present, the problem of organizing transport networks in megacities is the most urgent, since in all large cities traffic is developing at a much faster pace than the public road network. The capacity of most existing roads can only adequately support average traffic levels, but during “peak” hours, such as in the morning and evening, when the number of vehicles increases significantly, the problem of “traffic jams” arises. To solve this problem, effective management methods are needed to restructure traffic flows at the busiest locations, especially at complex interchanges [1, 2].

Obviously, it is difficult or impossible to conduct a full-scale experiment on traffic flow management in a megacity, so simulation modelling becomes the only tool for effective decision-making in this area. The advantage of this method is that, unlike the analytical method, simulation modelling of traffic flows allows an unlimited number of times to reproduce the system under study and determine its optimal state [3–5].

Study materials and methods

Initially an analysis of public roads should be carried out to identify the most problematic sections based on predetermined criteria, using appropriate road classification attributes and accumulated statistical data.

In general, these classification characteristics may include the presence and number of traffic lights, the number of pedestrians and vehicles, the presence of road markings, the number of traffic lanes, and their narrowing and widening [6, 7].

The main classification attributes of roads are shown in Table 1.

Obviously, the megacity traffic flow has features such as stochasticity, non-stationarity, incomplete controllability, and multiplicity of criteria for assessing its quality. These features affect the modeling process and make it impossible to build an adequate analytical model that allows studying this system and its characteristics under various operating conditions. The level of detail of the modelled process can be chosen depending on the respective approach, either macro- or micro-level models are used [8–11].

Macro-level models consider vehicles in groups without considering their condition individually. This makes it possible to estimate traffic characteristics at a particular point in time, which makes this approach preferable for simulations of complex roadway elements such as intersections.

At the micro-level models, it is possible to consider the individual characteristics of each vehicle and the interactions between them. This approach is preferable for modelling individual events at direct traffic points.

Consider an algorithm for controlling vehicle flows at an intersection of a metropolitan motorway, taking into account such factors as the functioning of traffic lights, the possibility of widening the road, and the possibility of using a new branch line [12–14].

To investigate the algorithm and conduct simulations, a transport scheme linking Ring Road 1 with High Speed Tollway 2, and the urban road network around Highways 3 and 4 was selected (Fig. 1).

The main problem with this scheme is the merging of three roads into one, with a final narrowing of 4 lanes. The traffic flows further to the junction with a pedestrian crossing, traffic lights and the ability to continue straight ahead, turn right, left and make a U-turn. Due to this large number of directions and the

huge flow of vehicles, traffic congestion is a constant problem. In order to find a solution to this problem, a simulation model of the intersection was implemented, with the possibility of upgrading it and reorganizing the corresponding traffic. As a new solution to the traffic flow management problem, we proposed a reorganization of the intersection based on a roundabout.

The development of the new solution included:

1. Construction of an initial simulation model of an intersection.
2. Computer experiment with the original model.
3. Development of recommendations for the reorganization of traffic flows.
4. Formation of a simulation model of an intersection taking into account the recommendations.

To implement this approach, we chose AnyLogic software system, which allows the use of road library data to build the required models. The main objective was to find possible solutions to the problem of the road layout in question using different simulation model variants.

Table 1

Technical classification of public roads

| Class of road | Category of road | Total number of lanes | Width of lane, m | Central dividing strip | Intersections with roads, bicycle and pedestrian paths | Crossings with railways and tram tracks | Access to the road from a single level abutment |
|-------------------------------|------------------|-----------------------|------------------|------------------------|---|---|---|
| Motorway | IA | 4 or more | 3.75 | Obligatory | At different levels | | Not allowed |
| High-speed road | IB | 4 or more | 3.75 | | | | Allowed without crossing straight ahead |
| The road of the ordinary like | IB | 4 or more | 3.75 | Obligatory | Allow single level crossings with traffic light control | At different levels | Allowed |
| | II | 4 or 3 | 3.75 / 3,5 | The absence of | | | |
| | III | 2 | 3.5 | Not required | Crossings at the same level are allowed | Allow for intersections on the same level | |
| | IV | 2 | 3.0 | | | | |
| | V | 1 | 4.5 or more | | | | |

To develop the simulation model of the transport scheme under study, we used the blocks of AnyLogic road library, such as: Source – responsible for generating cars on the selected section of the transport network; carMoveTo – allowing to determine the next road and direction for cars; carDispose – block; selectOutput – allows you to select the direction for cars, based on probabilistic estimates; roadNetworkDescriptor – makes it possible to show the degree of congestion of sections using a color scheme; trafficLight – sets the time interval of traffic lights operation. The considered traffic network has 29 routes and is shown in Fig. 2 and 3.

The computer experiment performed on the original simulation model showed the same results as the real traffic process, namely, a large congestion formed, which is shown in Fig. 3.

After analyzing the existing situation, the model was adjusted based on the decision made to reorganize the traffic flow by creating a roundabout, i.e. to exclude the intersection and the corresponding traffic



Fig. 1. Interchange with junction

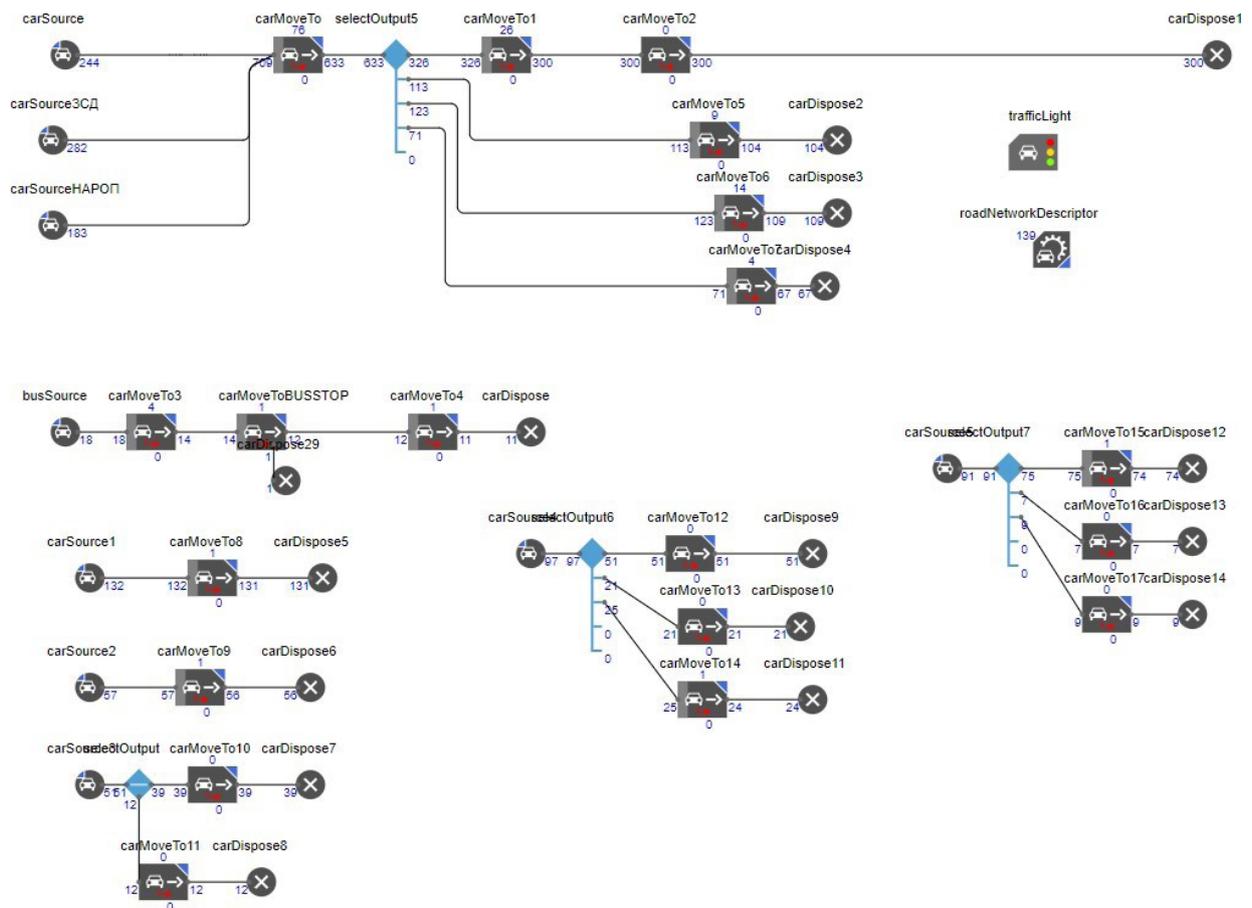


Fig. 2. Transport network before the change

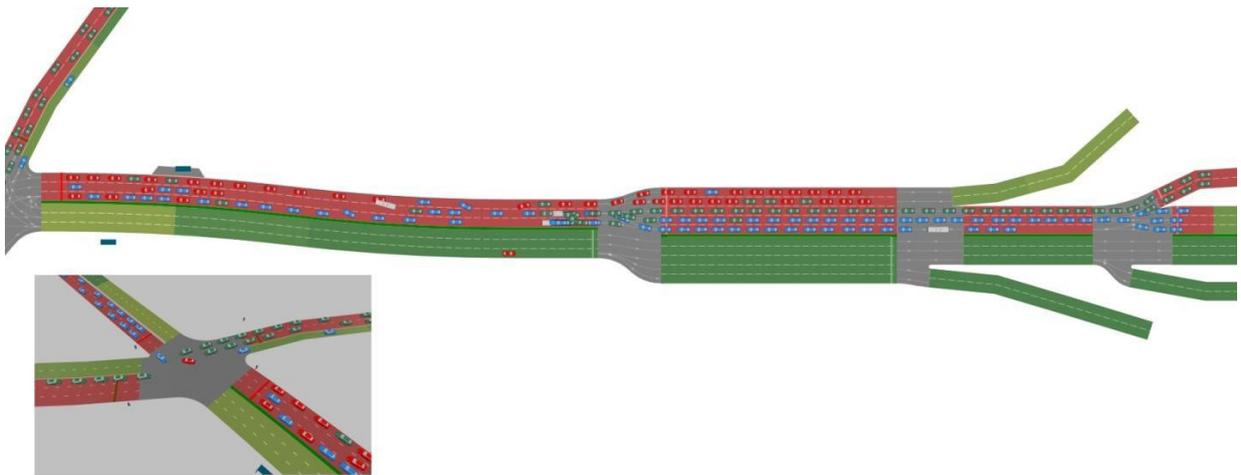


Fig. 3. Congestion on the model before the change

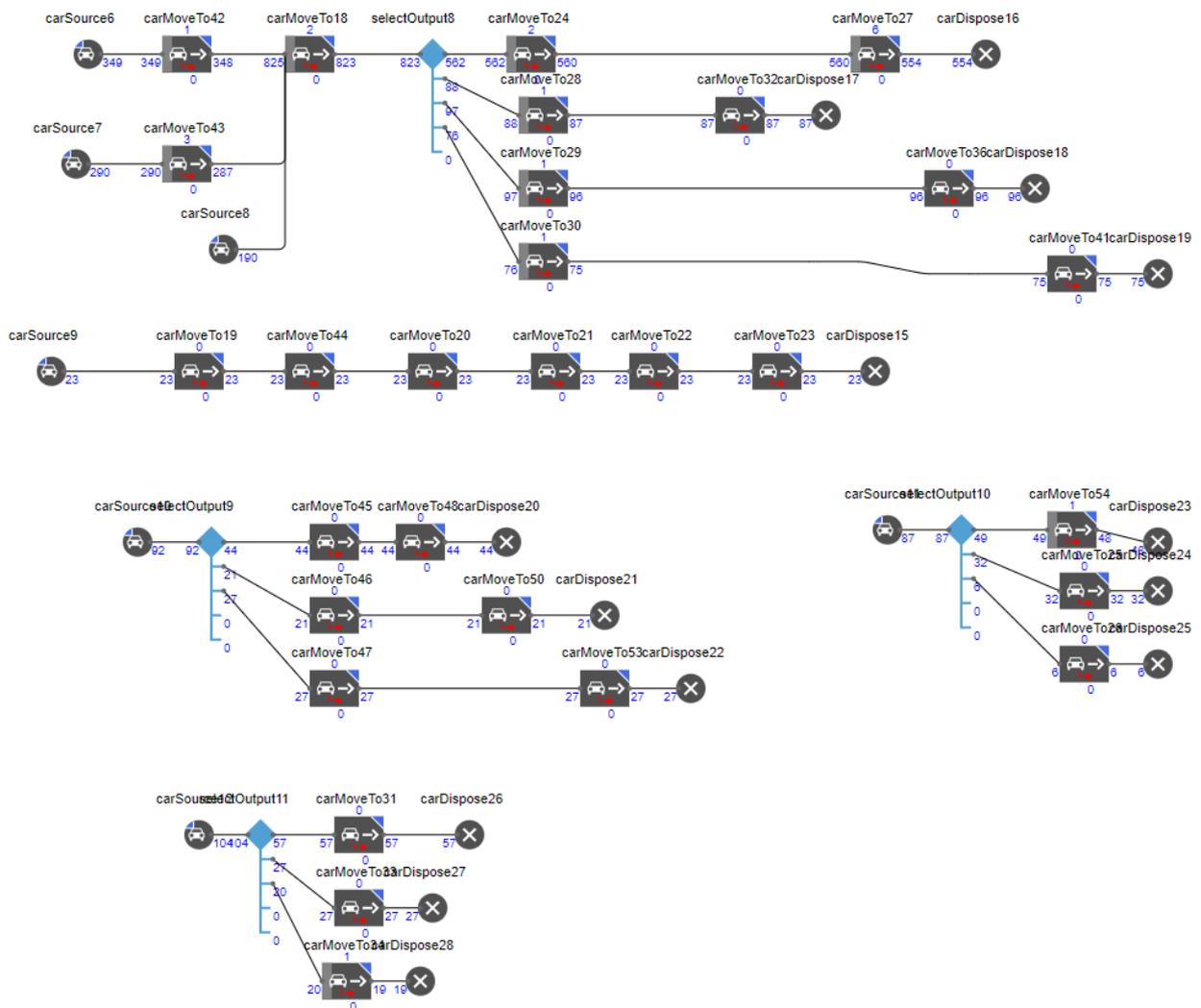


Fig. 4. Transport network after change (with roundabouts)

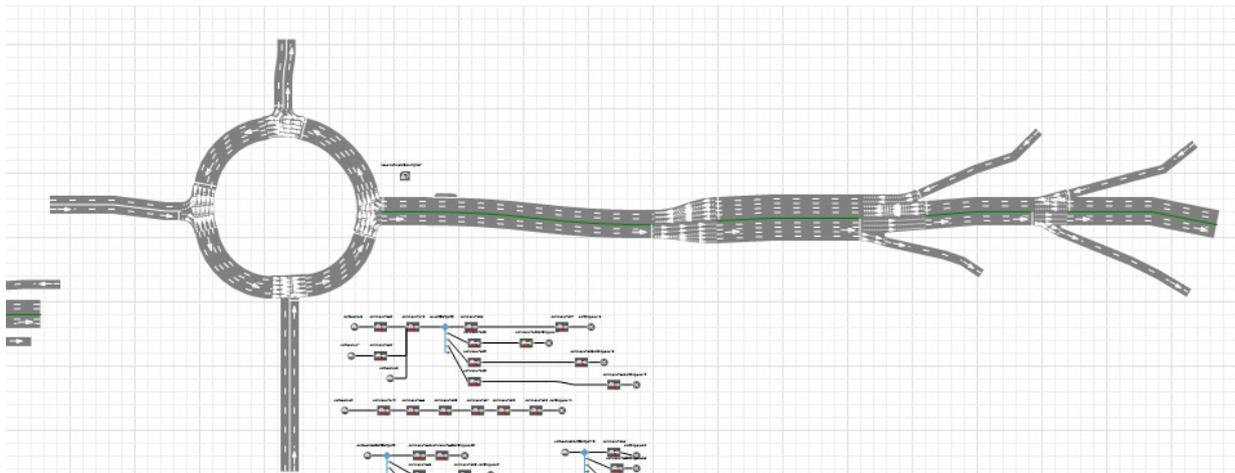


Fig. 5. Model after modification

lights from the traffic pattern. The implementation of the new version of the model is shown in Fig. 4 and 5.

Research results and discussion

The final comparative analysis of both models, with the same input data per unit time (1 hour), showed the following results, presented in Table 2.

Table 2

Final comparative analysis of models

| Without circular traffic, pcs. | | With circular traffic, pcs. | |
|--------------------------------|------------|-----------------------------|-------------|
| Buses treated | 11 | Buses treated | 23 |
| From road No. 3 | 580 | From road No. 3 | 812 |
| Towards road No. 3 | 99 | Towards road No. 3 | 103 |
| To road No. 4 | 96 | To road No. 4 | 92 |
| From road No. 4 | 90 | From road No. 4 | 86 |
| TOTAL | 865 | TOTAL | 1093 |

The results of the modelling show that the capacity of the motorway in the main direction has increased by 40 %, which allows a significant increase in traffic flow without the risk of congestion (Fig. 6). The final model has received an increase of 27 %. Accordingly, the roundabout-based scheme can handle around 1,500 vehicles per hour.

Thus, when performing a computer experiment on the model, it is possible to determine the degree of congestion of sections of the metropolitan route network in each period of time. In addition, the model allows varying the input parameters such as vehicle schedules, type and number of vehicles on the route, route structure, vehicle speed, etc., and analyzing changes in the situation to the benefit of effective decision-making [15, 16].

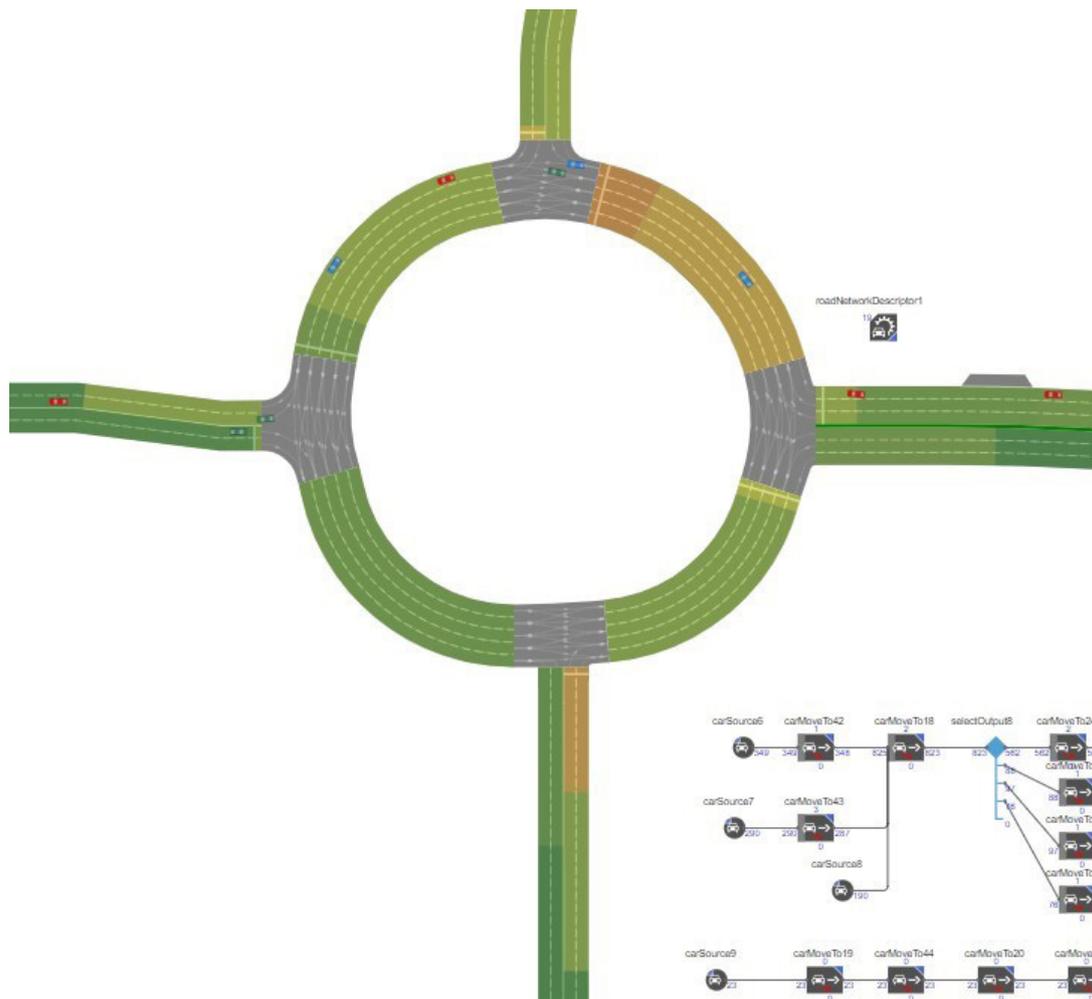


Fig. 6. Model after the computer experiment

Conclusion

The analysis shows the promise of using the proposed simulation model to study real traffic flow management processes in order to study their behavior. This model can be further improved by introducing more complex parameters and increasing the size of the transport network under study.

The model proposed in the article has the following advantages:

1. The model of the actual transport system is based on the object-oriented approach.
2. Visualization of the model allows identifying the busiest sections of the city's transport network that require redistribution of traffic flows.

The application of the developed model and the analysis of the data obtained from the experiment based on its use will improve the quality of transport services to the population, will contribute to the reduction of tension on the metropolitan motorways and reduce the number of road accidents.

REFERENCES

1. **Adrianova A., Laskin M., Svistunova A.** Simulation of the route network section of Pulkovo and Domodedovo airports with anylogic software. *9th All-Russian Scientific and Practical Conference on Simulation and its Application in Science and Industry*, 2019, Pp. 347–351. (rus)

2. **Balamirzoev A.G., Balamirzoeva E.R., Gadzhiyeva A.M.** Using the simulation in the optimization of transport systems management. *Proceedings of the 3rd International Scientific and Practical Conference on Current Problems and Prospects for the Development of the Road Transport Complex*, Makhachkala, 2017. Pp. 121. (rus)
3. **Volodares M.V., Maslenikova A.G.** To the question of imitation modeling of transport systems. *Bulletin of the Donetsk Academy of Automobile Transport*, 2017, No. 4, Pp. 21–28. (rus)
4. **Iskanderov Y., Pautov M.** Heterogeneous engineering in intelligent logistics. *Advances in Intelligent Systems and Computing*, 2021, Vol. 1179, Pp. 83–91.
5. **Lupin S., Shein T., Lin K.K., Davydova A.** Modelling of transport systems in a competitive environment. *Proceedings of the 5th International Conference on Internet Technologies and Applications*, ITA, 2013, Pp. 41–48.
6. **Pogarskaya A.S.** Application of simulation modeling in planning the operation of transport systems. *Operation of Sea Transport*, 2020, No. 3 (96), Pp. 32–41. (rus). DOI: 10.34046/aumsuomt96/5
7. **Svistunova A., Khasanov D., Kravets D.** Simulation of passenger flow at a railway station in anylogic software. *Conference SAEC-2020*, Pp. 274–282. (rus). DOI: 10.18720/SPBPU/2/id20-176
8. **Seliverstov S.A., Seliverstov Y.A., Tarantsev A.A., Grigoriev V.A., Elyashevich A.M., Muksimova R.R.** Elaboration of intelligent development system of megalopolis transportation. *2017 IEEE II International Conference on Control in Technical Systems (CTS)*, 2017, Pp. 211–215. DOI: 10.1109/CTS.2017.8109528
9. **Seliverstov S., Seliverstov Y.** Developing principles for building transport networks of conflict-free continuous traffic. *Transportation Research Procedia*, No. 36, Pp. 689–699. DOI: 10.1016/j.trpro.2018.12.12
10. **Seliverstov Y.A., Seliverstov S.A., Lukomskaya O.Y., Nikitin K.V., Grigoriev V.A., Vydrina E.O.** The method of selecting a preferred route based on subjective criteria. *Proceedings of 2017 IEEE II International Conference on Control in Technical Systems (CTS)*, 2017, Pp. 126–130.
11. **Seliverstov Ya.A., Seliverstov S.A., Malygin I.G., Tarantsev A.A., Shatalova N.V., Lukomskaya O.Yu., Tishchenko I.P., Elyashevich A.M.** Development of management principles of urban traffic under conditions of information uncertainty. *Proceedings of the 2nd Conference on Creativity in Intelligent Technologies and Data Science (CIT&DS 2017)*, Volgograd, Russia, Sept. 12-14, 2017, Pp. 399–418. DOI: 10.1007/978-3-319-65551-2_29
12. **Shamlitskiy Y.I., Mironenko S.N., Kovbasa N.V., Bezrukova N.V., Tynchenko V.S., Kukartsev V.V.** Evaluation of the effectiveness of traffic control algorithms based on a simulation model in the AnyLogic. *Journal of Physics: Conference Series*, 2019, 1353 (1).
13. **Mishkurov P., Fridrikhson O., Lukyanov V., Kornilov S., Say V.** Simulated transport and logistics model of a mining enterprise. *TransSiberia*, 2020. (rus). DOI: 10.1016/j.trpro.2021.02.090
14. **Chandakas E., Leurent F., Cats O.** Applications and future developments: Modeling software and advanced applications. *Tracts on Transportation and Traffic*, 2016, No. 10. DOI: 10.1007/978-3-319-25082-3_9
15. **Svistunova A.S., Khasanov D.S.** Simulation of the passenger service process at the railway station. *Proceedings of the Jubilee International Scientific and Practical Conference. N.S. Solomenko Institute of Transport Problems of the Russian Academy of Sciences*, St. Petersburg, 2020, Pp. 28–32. (rus)
16. **Svistunova A.S., Khasanov D.S.** Estimation of the efficiency of service to passengers in the airport complex. *Proceedings of the Jubilee International Scientific and Practical Conference. N.S. Solomenko Institute of Transport Problems of the Russian Academy of Sciences*, St. Petersburg, 2020, Pp. 32–37. (rus)

Received 25.05.2021.

СПИСОК ЛИТЕРАТУРЫ

1. **Адрианова А.В., Ласкин М.Б., Свистунова А.С.** Имитационное моделирование участка маршрутной сети аэропортов «Пулково» и «Домодедово» в среде AnyLogic // IX Всерос. науч.-практ. конф. по имитационному моделированию и его применению в науке и промышленности. Екатеринбург, 2019. С. 347–351.

2. Баламирзоев А.Г., Баламирзоева Е.Р., Гаджиева А.М. Использование имитационного моделирования при оптимизации управления транспортными системами // Сб. трудов конф. Актуальные проблемы и перспективы развития дорожно-транспортного комплекса. Махачкала, 2017. С. 121.
3. Володарец Н.В., Масленикова А.Г. К вопросу имитационного моделирования транспортных систем // Вестник Донецкой академии автомобильного транспорта. 2017. № 4. С. 21–28.
4. Iskanderov Y., Pautov M. Heterogeneous engineering in intelligent logistics // Advances in Intelligent Systems and Computing. 2021. Vol. 1179. Pp. 83–91.
5. Lupin S., Shein T., Lin K.K., Davydova A. Modelling of transport systems in a competitive environment // Proc. of the 5th Internat. Conf. on Internet Technologies and Applications. 2013. Pp. 41–48.
6. Погарская А.С. Применение имитационного моделирования при планировании работы транспортных систем // Эксплуатация морского транспорта. 2020. № 3 (96). С. 32–41. DOI: 10.34046/aumsuomt96/5
7. Свистунова А.С., Хасанов Д.С., Кравец Д.М. Имитационное моделирование пассажиропотока на железнодорожном вокзале в программной среде AnyLogic // Системный анализ в проектировании и управлении. Ч. 2: сб. науч. трудов XXIV Междунар. науч. и уч.-практ. конф. 2020 [в 3 ч.]. С. 274–282. DOI: 10.18720/SPBPU/2/id20-176
8. Seliverstov S.A., Seliverstov Y.A., Tarantsev A.A., Grigoriev V.A., Elyashevich A.M., Muksimova R.R. Elaboration of intelligent development system of megalopolis transportation // Proc. of the 2017 IEEE II Internat. Conf. on Control in Technical Systems. 2017. Pp. 211–215. DOI: 10.1109/CTSUS.2017.8109528
9. Seliverstov S., Seliverstov Y. Developing principles for building transport networks of conflict-free continuous traffic // Transportation Research Procedia. 2018. No. 36. Pp. 689–699. DOI: 10.1016/j.trpro.2018.12.12
10. Seliverstov Y.A., Seliverstov S.A., Lukomskaya O.Y., Nikitin K.V., Grigoriev V.A., Vydrina E.O. The method of selecting a preferred route based on subjective criteria // Proc. of the 2017 IEEE II Internat. Conf. on Control in Technical Systems. 2017. Pp. 126–130.
11. Seliverstov Ya.A., Seliverstov S.A., Malygin I.G., Tarantsev A.A., Shatalova N.V., Lukomskaya O.Yu., Tishchenko I.P., Elyashevich A.M. Development of management principles of urban traffic under conditions of information uncertainty. Creativity in intelligent technologies and data science // Proc. of the 2nd Conf. on CIT&DS 2017. Volgograd, 2017. Pp. 399–418. DOI: 10.1007/978-3-319-65551-2_29
12. Shamlitskiy Y.I., Mironenko S.N., Kovbasa N.V., Bezrukova N.V., Tynchenko V.S., Kukartsev V.V. Evaluation of the effectiveness of traffic control algorithms based on a simulation model in theAnyLogic // J. of Physics: Conf. Series. 2019. 1353 (1).
13. Mishkurov P., Fridrikhson O., Lukyanov V., Kornilov S., Say V. Simulated transport and logistics model of a mining enterprise // The Internat. Siberian Transport Forum TransSiberia. 2020. DOI:10.1016/j.trpro.2021.02.090
14. Chandakas E., Leurent F., Cats O. Applications and future developments: Modeling software and advanced applications // Tracts on Transportation and Traffic. 2016. 10. DOI: 10.1007/978-3-319-25082-3_9
15. Свистунова А.С., Хасанов Д.С. Имитационное моделирование процессов обслуживания пассажиров на железнодорожном вокзале // Матер. Юбилейной междунар. науч.-практ. конф. Транспорт России: проблемы и перспективы. СПб., 2020. С. 28–32.
16. Свистунова А.С., Хасанов Д.С. Оценка эффективности обслуживания пассажиров в аэровокзальном комплексе // Матер. Юбилейной междунар. науч.-практ. конф. Транспорт России: проблемы и перспективы. СПб., 2020. С. 32–37.

Статья поступила в редакцию 25.05.2021.

THE AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Svistunova Aliaksandra S.

Свистунова Александра Сергеевна

E-mail: svistunova_alexandra@bk.ru

Khasanov Dmitry S.

Хасанов Дмитрий Салимович

E-mail: dkhasanovsuai@yandex.ru

© Санкт-Петербургский политехнический университет Петра Великого, 2021



Circuits and Systems for Receiving, Transmitting and Signal Processing

DOI: 10.18721/JCSTCS.14304
УДК 621.37

SEMI-NATURAL MODELING FOR GNSS INTEGRITY MONITORING ALGORITHM

A.P. Rachitskaya

Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

The paper considers the suboptimal version of GNSS integrity monitoring algorithm involving multichannel signal processing. This algorithm was examined in terms of probability-based characteristics obtained during semi-natural modeling. Such modeling assumes that multichannel snapshots are getting from real channels of multichannel GNSS receiver including antenna array when all subsequent procedures are implemented in Matlab later. Probability-based characteristics obtained in such way consequently checked with similar characteristics obtained by Matlab simulation ideal model, which ignored probable effects of signal transmission and reception in real environment. It was shown the level of similarity between characteristics of both types, and also clarified the conditions when the characteristics are close to each other, and the conditions when the difference between them is significant. The main reason of such difference was found out empirically.

Keywords: generalized maximum likelihood ratio test, probability-based characteristics, confidence interval, non-identity of channels, multichannel receiver.

Citation: Rachitskaya A.P. Semi-natural modeling for GNSS integrity monitoring algorithm. Computing, Telecommunications and Control, 2021, Vol. 14, No. 3, Pp. 43–55. DOI: 10.187-21/JCST-CS.14304

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

ПОЛУНАТУРНОЕ МОДЕЛИРОВАНИЕ АЛГОРИТМА КОНТРОЛЯ ЦЕЛОСТНОСТИ НАВИГАЦИОННОГО ПОЛЯ ГНСС

А.П. Рачицкая

Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

Рассмотрен алгоритм контроля целостности навигационного поля (КЦНП), синтезированный в соответствии с обобщенным критерием отношения правдоподобия для прямой многоканальной обработки сигналов с элементов антенной решетки (АР). Произведена оценка эффективности алгоритма с помощью полунатурного моделирования, с использованием записи реальных навигационных сигналов с элементов АР и их обработкой в среде Matlab в соответствии с рассматриваемым алгоритмом КЦНП. Проведено сравнение результатов измерения вероятностных характеристик алгоритма, полученных на основе такого моделирования, с аналогичными результатами идеализированного моделирования в среде Matlab, когда формирование сигналов производится искусственно без учета возможных факторов (на пути распространения сигналов или при их приёме), возникающих при работе алгоритма в реальных условиях. Выявлена степень соответствия результатов обоих типов моделирования. Определены условия, когда получаемые результаты оказываются близки и когда расхождения значительны. Экспериментальным путем установлена одна из возможных причин найденных различий между характеристиками.

Ключевые слова: обобщенный критерий максимального правдоподобия, вероятностные характеристики, доверительный интервал, неидентичность каналов, многоканальный приёмник.

Ссылка при цитировании: Rachitskaya A.P. Semi-natural modeling for GNSS integrity monitoring algorithm // Computing, Telecommunications and Control. 2021. Vol. 14. No. 3. Pp. 43–55. DOI: 10.18721/JCSTCS.14304

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>).

Introduction

Various faults occurring in the process of receiving and processing signals of the global navigation satellite system (GNSS) [2–5] can lead to significant errors in the navigation object (NO) devices. If the errors exceed a certain permissible level (for example, the level of normal errors caused by the thermal noise of the NO receiver), it is interpreted as a breach of the integrity of the navigation field (NF) [1, 6–9]. To detect such situations, specialists develop methods of GNSS integrity monitoring [9–16]. Each of the methods is usually optimized for specific types of faults. In addition, the effectiveness of the developed methods of GNSS integrity monitoring aimed at identifying critical failure caused by the false navigation signal sources (FNSS) in a number of cases either proves to be unacceptably low or inapplicable for all NO types [13, 15, 17–19]. Failures caused by spoofing are especially hazardous as the NO may receive signals identical to the signals of navigation satellites (NS) at the input. However, they have different values of time-frequency parameters, which results in bias errors in the results of the NO devices operation [11, 13, 20–22].

There are various approaches that consider the features of such faults [13, 14, 20, 23, 24, 26], but in the majority of cases, the task of GNSS integrity monitoring can be reduced to decision-making on whether or not there is any spoofing involved. To improve the effectiveness of such a check, it is possible to use statistical synthesis of a decision-making algorithm based on the *a priori* known data on the position of GNSS satellites at the orbit. Moreover, the best effectiveness is achieved by “direct” analysis of the processes observed in the antenna array (AA) elements [27, 28].

Direct suboptimal algorithm is of special practical interest. It is obtained as a result of statistical synthesis taking into account several simplifications, in particular, an assumption of significant difference between the NO coordinates estimates and its actual position in case of FNSS impact, as well as some other conditions (see below) [28]. The feasibility of using this algorithm is due to its low computational complexity in comparison with other similar algorithms (the number of multiplications omitted can reach up to one order or more) combined with its high level of effectiveness, which is proved with modeling in the Matlab environment [27, 28]. At the same time, this modeling was conducted in ideal conditions disregarding possible factors emerging in real conditions (features of satellite and FNSS signals propagation, nonidentity of NO device channels, etc.). Therefore, the conclusions drawn on the basis of such an ideal (hereafter referred to as “conventional”) modelling cannot be considered fully objective. We can make the results obtained before more accurate by means of semi-natural modeling, when we use the records of actual navigation signals from AA elements and process them in the Matlab environment according to the GNSS integrity monitoring algorithm under study.

We evaluated the effectiveness of the considered GNSS integrity monitoring algorithms [27] based on the analysis of such characteristics as:

- 1) the probability of a false decision that there is interference, while it is actually absent (P_{FA} probability of false alarm);
- 2) the probability of a false decision that there is no interference, while it is actually present (P_{MD} probability of missed detection).

Both probabilities are calculated, if there are GNSS signals present. In addition, according to the Neyman – Pearson criterion [25], the considered algorithms were optimized by means of minimizing the P_{MD} probability at a fixed value of the P_{FA} probability, which determines the value of decision-making threshold Λ_0 .

Direct suboptimal algorithm of GNSS integrity monitoring

Let us consider a direct suboptimal algorithm of GNSS integrity monitoring obtained on the basis of a comparison between the likelihood ratio and the threshold for the processes snapshots in the AA elements.

The likelihood functions for column-vector $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T$ of $x_m(t)$ processes snapshots on each m^{th} of M antenna elements is formed separately if one of the two possible hypothesis H_0 and H_1 is valid in case of observing a constellation from L satellites and the presence of additive white Gaussian noise (AWGN) in the AA elements. H_0 corresponds to the situation when the FNSS influence is absent, while H_1 takes place when there is FNSS influence present, moreover, we consider a single FNSS emitting all L radio-navigation signals from one point [11]. The parameters of the satellite radio-navigation signals (initial phases, amplitudes), as well as the NO parameters (its coordinates \mathbf{P}_{NO} , velocity, attitude angles $\alpha_1, \alpha_2, \alpha_3$) are considered unknown. We assume the FNSS parameters (coordinates \mathbf{P}_S of the FNSS itself), false coordinates \mathbf{P}' and velocity vector \mathbf{v}' of the NO generated by the FNSS to be unknown as well. We exclude the indicated unknown (“interfering”) parameters in accordance with the generalized likelihood ratio test [25].

The test involves maximization (either analytical directly during the algorithm synthesis or numerical during the consequent algorithm running) with respect to the values of these parameters. The “initial” direct optimal algorithm obtained in this manner requires considerable computations as the majority of the maximization procedures with respect to unknown parameters can be solved numerically only in the process of running the considered algorithm. We can significantly simplify it by a conversion into a suboptimal algorithm: to reduce the computing costs, we use a justified assumption on a significant deviation of the NO coordinates from the actual position if the H_1 hypothesis is valid [29]. We can achieve additional simplification by means of substituting the numerical maximization with respect to unknown values of \mathbf{P}' coordinates and \mathbf{v}' velocity vector with a numerical maximization directly with respect to time-frequency parameters $\boldsymbol{\tau}'_1, \Delta\boldsymbol{\omega}'_d$ of the FNSS radio-navigation signals. Here, $\boldsymbol{\tau}'_1 = [\tau_1^{(1)}, \tau_1^{(2)}, \dots, \tau_1^{(L)}]$ is a vector the elements of which comprise the propagation times of l^{th} FNSS signal to the AA antenna element chosen as a reference one, while $\Delta\boldsymbol{\omega}'_D = [\Delta\omega_D^{(1)}, \Delta\omega_D^{(2)}, \dots, \omega_D^{(L)}]$ is a vector of Doppler shifts of the FNSS signals frequencies.

The “direct suboptimal algorithm” obtained this way presupposes a numerical maximization with respect to angular directions μ_s and η_s (azimuth and incline respectively) in the FNSS [28]:

$$\frac{1}{ME_0N_0} \left\{ \max_{\substack{\boldsymbol{\tau}'_1, \Delta\boldsymbol{\omega}'_d \\ \mu_s, \eta_s}} \sum_{l=1}^L |\mathbf{V}_l'^T \mathbf{H}_s|^2 \right\} \begin{matrix} H_1 \\ > \\ < \\ H_0 \end{matrix} \Lambda_0, \quad (1)$$

where $E_0 = \int_{(T_a)} \left(\sum_{l=1}^L C_0^{(l)}(t) \right)^2 dt$; $N_0/2$ is the spectral density of the average AWGN power at the input of the receiver channels (the channels are assumed to be identical); $C_0^{(l)}(t)$ is a pseudorandom sequence modulating signal of the l^{th} satellite; $\mathbf{H}_s = [e^{j\delta\phi_1}, e^{j\delta\phi_2}, \dots, e^{j\delta\phi_M}]^T$ is the directing vector column for a pos-

sible direction μ_s, η_s in the FNSS, where $\delta\varphi'_m = \frac{\omega_0^{(l)}}{c} (\mathbf{k}_0^s(\mu_s, \eta_s))^T \tilde{\mathbf{P}}_m$; $\mathbf{k}_0^{(s)}(\mu_s, \eta_s) = [x_{k0}^{(s)}, y_{k0}^{(s)}, z_{k0}^{(s)}]^T$; $x_{k0}^{(s)} = \sin \mu_s \cos \eta_s$; $y_{k0}^{(s)} = \cos \mu_s \cos \eta_s$; $z_{k0}^{(s)} = \sin \eta_s$; c is the speed of light in vacuum, $l = 1 \dots L$, $\omega_0^{(l)}$ is the carrier frequency of the radio signal emitted by the l^{th} satellite; $\tilde{\mathbf{P}}_m = [\tilde{x}_m, \tilde{y}_m, \tilde{z}_m]^T$ are the coordinates of the AA elements in the local coordinates system of the NO [2, 3];

$$\mathbf{V}'_l = [V_1^{(l)}, V_2^{(l)}, \dots, V_M^{(l)}]^T, \tag{2}$$

where $V_m^{(l)} = \int_{(T_a)} F_{xm}(t) C_0^{*(l)}(t - \tau_1^{(l)}) e^{-j\Delta\omega_D^{(l)}t} e^{-j\Delta\omega_0^{(l)}t} dt$ and integration are performed in the analysis interval the T_a value of which is defined by the duration of the pseudorandom sequence $C_0^{(l)}(t)$; $F_{xm}(t)$ is a complex envelope of the $x_m(t)$ snapshot; $\Delta\omega_0^{(l)} = \omega_0^{(l)} - \omega_0$ is a deviation of the l^{th} satellite radio signal frequency from the carrier receiver frequency ω_0 .

Semi-natural modeling

In the course of the semi-natural modeling, a GPS simulator of L1 range (1575.42 MHz) emits L signals of FNSS. An experimental simulation device of GNSS integrity monitoring includes a multichannel signal recorder, 6-element AA receiver (layout presented in Fig. 1), and a standard navigation receiver (Ublox-M8T) for the monitoring of radio navigation signals present. The multichannel signal recorder makes snapshots of the processes from the AA elements via a multichannel radio frequency (RF) receive path and digitizes the respective $x_m(t)$ processes using an analog-to-digital converter. The reference values of the $F_{xm}(t)$ complex envelope of the $x_m(t)$ processes digitized at the sampling rate of 2.046 MHz are stored as .dat-files.

In the course of the semi-natural modeling, the GPS simulator emitted the FNSS signals with the power level significantly exceeding the radio navigation signals of the satellite (the interference/signal ratio was $\gamma \approx 6$ dB). The signals were received and recorded in real conditions in the presence of trees, urban buildings, moving objects, etc. Fig. 2a shows the satellite group observed in the process of the experiment.

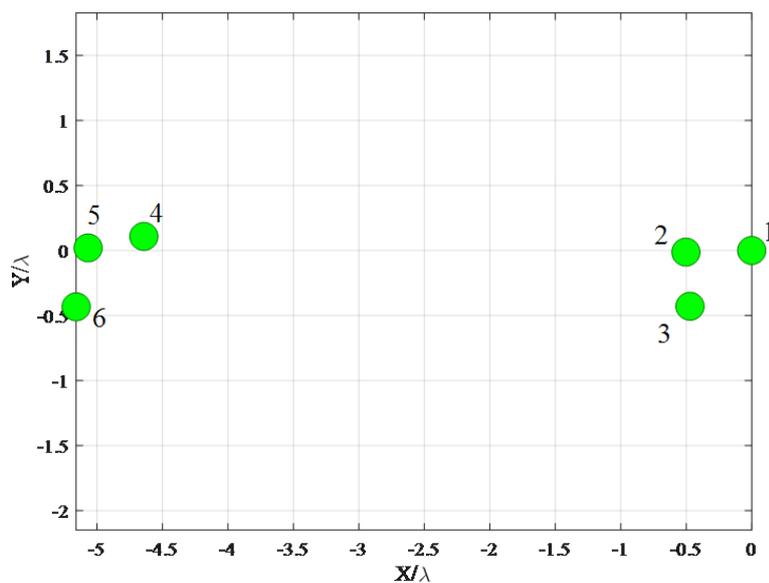


Fig. 1. Configuration of the 6-element AA used in the study

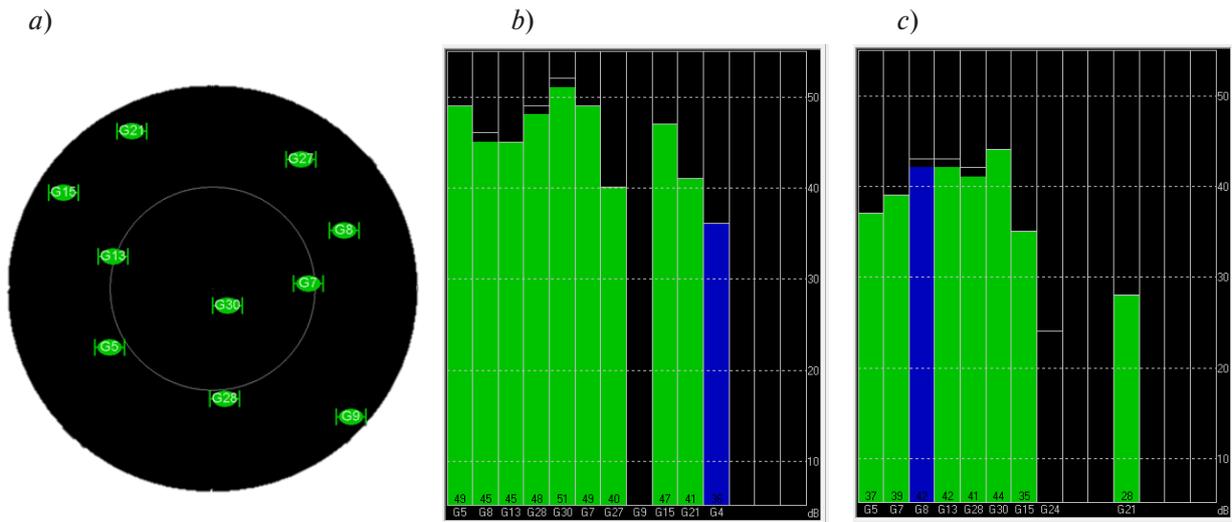


Fig. 2. View of the observed satellite constellation (a), visual representation of the detected radio navigation signals in the absence of FNSS signals (b) and in their presence (c) in the indicator of the Ublox-M8T receiver

Simultaneously with recording the signals from the AA elements, we were monitoring the readings of the standard navigation receiver connected to one of the AA elements. Fig. 2c shows an example of the results of processing legitimate GNSS satellite group signals in the interface of the standard navigation receiver Ublox-M8T in case there is no GNSS integrity failure. Each column depicted in Fig. 2b corresponds to the observed radio navigation signal the number of which (satellite number) is located in the bottom; the height of a column is proportional to the value of the C_0/N_0 ratio of C_0 carrier power of the radio navigation signal under consideration to the doubled spectral power density of the AWGN. Fig. 2b shows an example of indicating the results of processing the radio navigation signals by the same navigation receiver under the influence of FNSS. There is an obvious absence of any identifying attributes of the fact that in the second case the coordinates were measured according to the FNSS signals with significant mistakes. At the same time, in the presence of the FNSS the measured values of the NO coordinates considerably differ from the coordinates measured in case the FNSS influence is absent.

Thus, Table 1 shows an example of the measurements for the coordinates in both cases under study obtained in one of the experiments. The measured coordinates clearly differ by approximately 1600 m.

Table 1

Results of processing navigation signals by the Ublox-M8T receiver in the absence and in the presence of FNSS signals

| Presence of FNSS signals | Latitude, deg | Longitude, deg | Altitude, m | Standard deviation of measurements caused by AWGN, m | Absolute deviation from the actual NO position caused by FNSS, m |
|--------------------------|---------------|----------------|-------------|--|--|
| – | 59.9937998 | 30.3795120 | 56.5 | 3..7 | – |
| + | 59.9874327 | 30.3536052 | 2.7 | 4..8 | 1611.5 |

Results of the semi-natural modeling

We used Matlab algorithm model (1) for the semi-natural modeling, while the recordings of the real signals from the AA elements were incoming to the input of the model for analysis. The obtained charac-

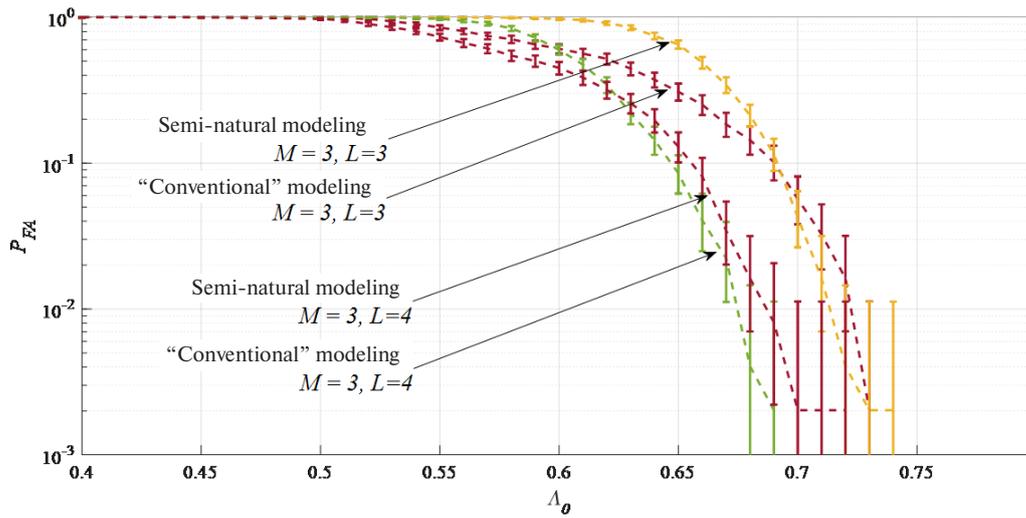


Fig. 3. Probability of false alarm at $M = 3, L = 3, 4$ and $C_0/N_0 = 45...50$ dB·Hz

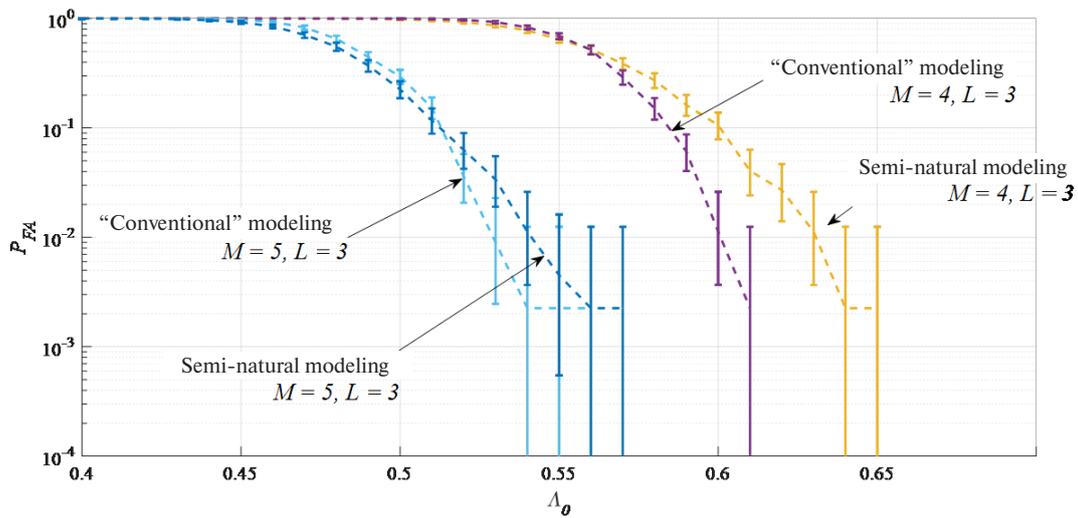


Fig. 4. Probability of false alarm at $L = 3, M = 4, 5$ and $C_0/N_0 = 45...50$ dB·Hz

teristics (P_{FA} and P_{MD}) of algorithm (1) were compared with the similar characteristics resulting from the “conventional” modeling in the identical conditions, when $C_0/N_0 = 40...50$ dB·Hz, $\gamma = 6$ dB. Fig. 3 presents dependencies of the P_{FA} probability on the decision-making threshold Λ_θ for $M = 3, L = 3$ (satellites no. 5, 7 and 8 in Fig. 2).

A comparison of the dependencies of the P_{FA} probability on the decision-making threshold obtained in the process of the semi-natural modeling with the results of the “conventional” modeling revealed a certain divergence (less than a half an order of magnitude). Similar conclusions can be also drawn while using the signals of 4 satellites ($L = 4$ in Fig. 3), as well as with a larger number of elements ($M = 4, 5$ in Fig. 4).

At the same time, when calculating the P_{MD} probability, we registered strong influence of the AA design on the degree of proximity of the characteristics obtained by two types of modeling under study. Thus, we found little difference in the P_{MD} values obtained in the course of the “conventional” and semi-natural modeling at the small number of the AA elements ($M = 3$ in Fig. 5); at the larger numbers ($M = 4$ in Fig. 6), this difference was growing. Therefore, at $M = 5$ (Fig. 7), the difference in the P_{MD} characteristics obtained by means of two different ways of modeling exceeded several orders.

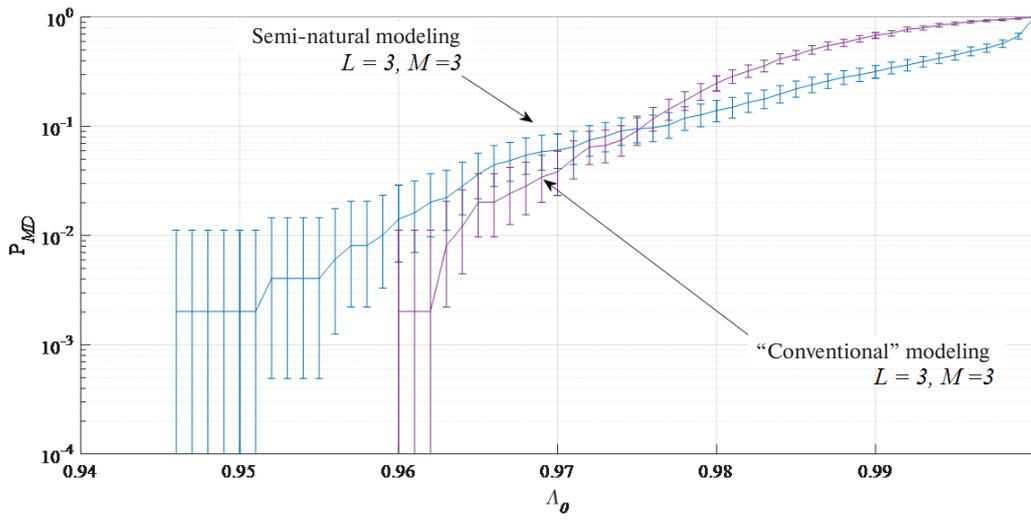


Fig. 5. Probability of missed detection at $M = 3, L = 4, C_0/N_0 = 45...50$ dB·Hz and $\gamma = 6$ dB

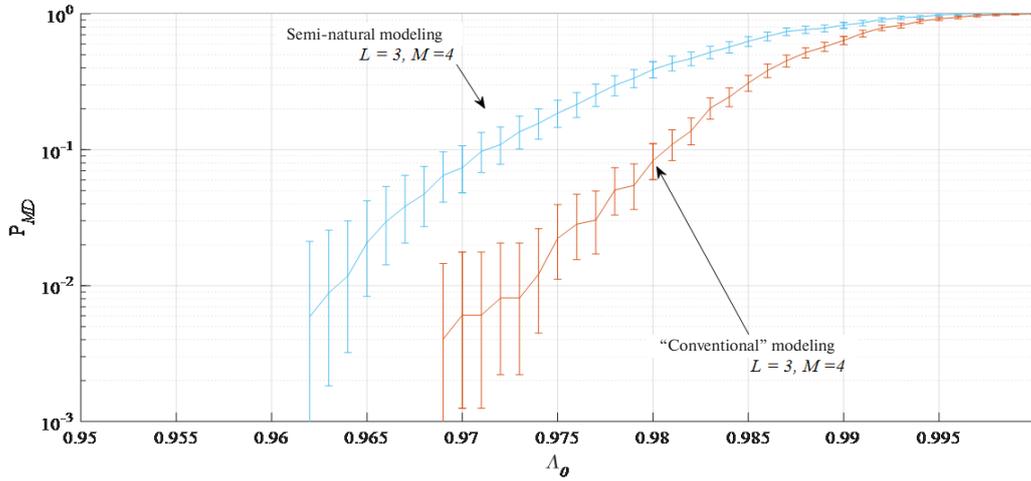


Fig. 6. Probability of missed detection at $M = 4, L = 3, C_0/N_0 = 45...50$ dB·Hz and $\gamma = 6$ dB

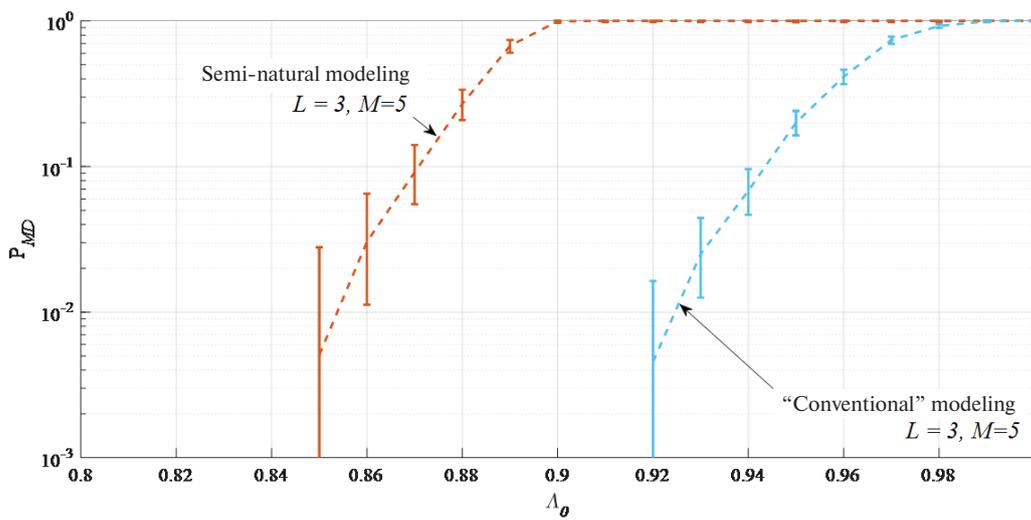


Fig. 7. Probability of missed detection at $M = 5, L = 3, C_0/N_0 = 45...50$ dB·Hz and $\gamma = 6$ dB

We can assume that the cause of the discovered divergence between the probability characteristics obtained by both the above mentioned types of modeling is a possible nonidentity of the channels of the receive path used for recording of the signals from the respective AA elements. Testing this assumption is of interest.

Influence of nonidentity of the receive path channels on the efficiency of the GNSS integrity monitoring system

As additional researched results [30] showed, the semi-natural modeling employed a multichannel receiver with a significant difference in the gain characteristics between the channels, so the phase difference $\Delta\varphi_m$ between the channels reached 2π rad and more (Table 2).

Table 2

Change in the unknown $\Delta\varphi_m$ phase tune-outs between the AA channels

| | Channel 1 | Channel 2 | Channel 3 | Channel 4 | Channel 5 | Channel 6 |
|-------------------------|-----------|-----------|-----------|-----------|-----------|-----------|
| $\Delta\varphi_m$, rad | 0 | 1.699130 | -1.650571 | 0.985480 | -1.128182 | -0.255712 |

By compensating the phase tune-outs (according to Table 2) during the semi-natural modeling, we can improve the characteristics of the GNSS integrity monitoring algorithm and make them approach the respective characteristics obtained by means of the “conventional” modeling (Fig. 8–10). The improvement of the probability characteristics after compensation is especially significant at $M > 3$, so at $M = 5$ it reaches 3 or more orders. Thus, we can confirm the assumption that the main reason for the divergence in the results of the semi-natural modeling from the “conventional” modeling lies in the nonidentity of the receiver channels.

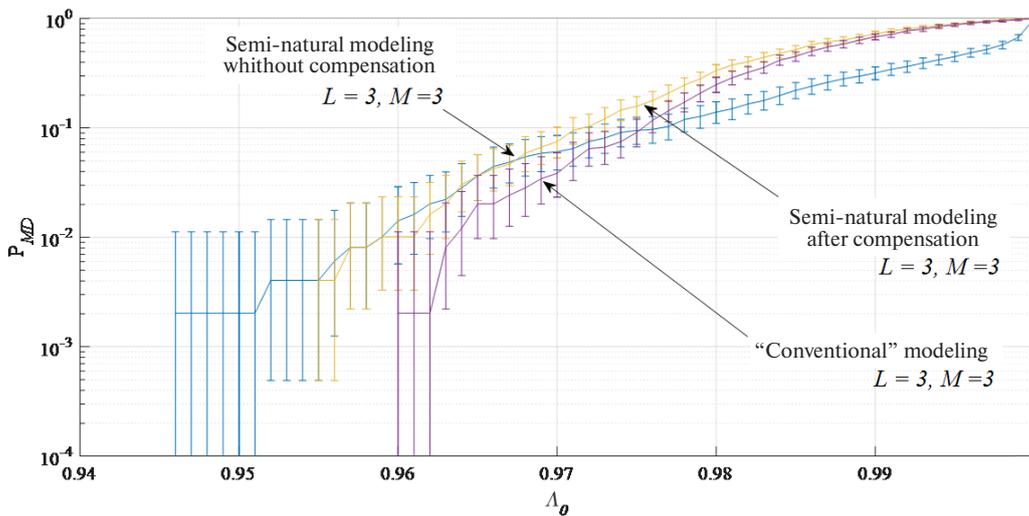


Fig. 8. Probability of missed detection at $M = 3, L = 3, C_0/N_0 = 45...50$ dB·Hz and $\gamma = 6$ dB

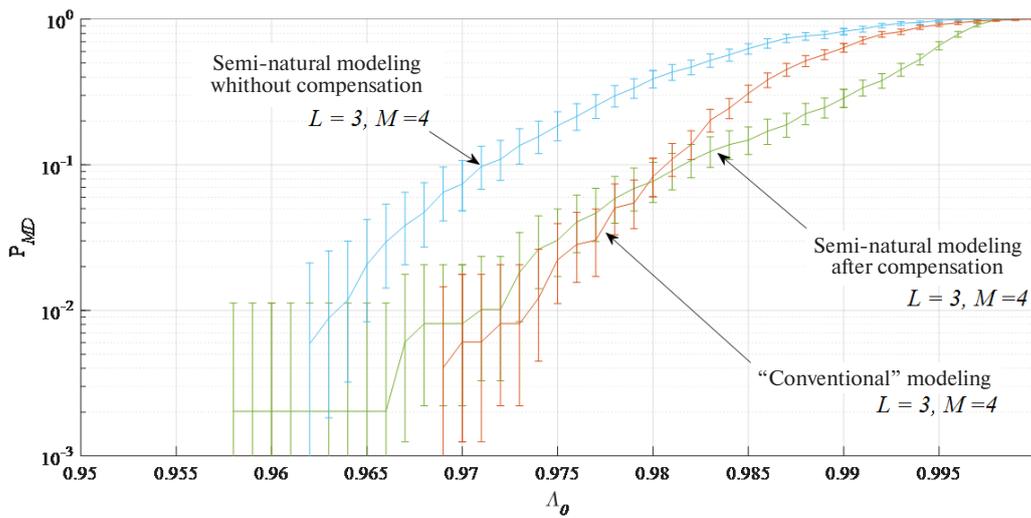


Fig. 9. Probability of missed detection at $M = 4$, $L = 3$, $C_0/N_0 = 45...50$ dB·Hz and $\gamma = 6$ dB

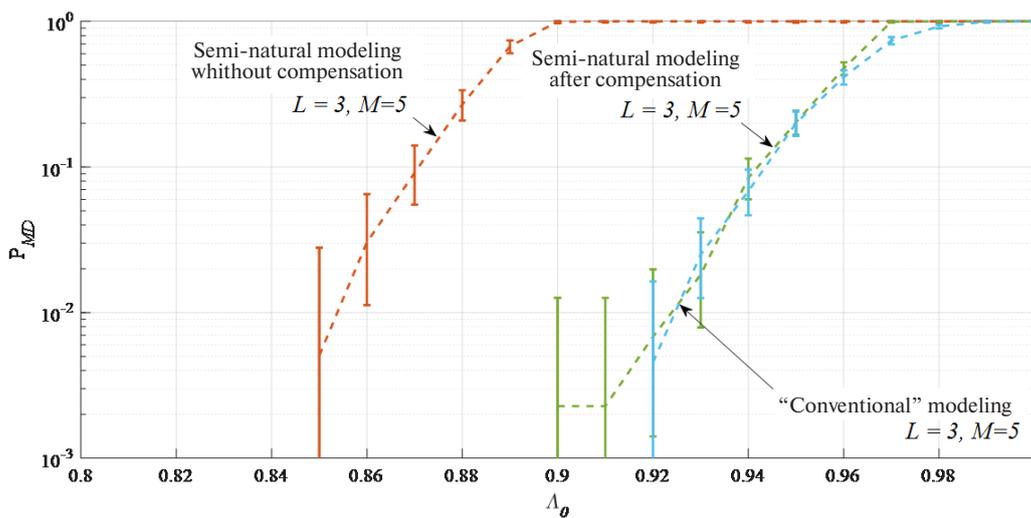


Fig. 10. Probability of missed detection at $M = 5$, $L = 3$, $C_0/N_0 = 45...50$ dB·Hz and $\gamma = 6$ dB

Conclusions

In the process of the conducted research, we showed that the “conventional” modeling in ideal conditions (rectilinear propagation of satellite signals, absence of the multipath propagation effect, identity of the receive path channels, etc.) allows us to evaluate the considered GNSS integrity monitoring algorithm for the simplest antenna arrays (2–3 elements) quite accurately in real receiving conditions even in case of a considerable nonidentity of the receive path channels. On the other hand, when more complex antenna arrays (4 or more elements) are involved, the results of the “conventional” modeling adequately reflect the efficiency of the GNSS integrity monitoring algorithm built only on the basis of radio receive path with a compensation of phase tune-outs between the channels.

The considered method of the semi-natural modeling obviously allows evaluating the efficiency of the GNSS integrity monitoring algorithm with various geometric properties of the GNSS groups and in different conditions of receiving signals from satellites and false sources.

REFERENCES

1. **Teunissen P.J.G., Montenbruck O.** (eds.). *Springer handbook of global navigation satellite systems*. Springer, 2017.
2. **Cheng J., Wang J., Zhao L.** A direct attitude determination approach based on GPS double-difference carrier phase measurements. *Journal of Applied Mathematics*, 2014, No. 6, Pp. 1–6. DOI: 10.1155/2014/548083
3. **Daneshmand S., Sokhandan N., Lachapelle G.** Precise GNSS attitude determination based on antenna array processing. *Proceedings of the 27th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2014)*, Tampa, FL, USA, 2014, Vol. 812.
4. **Niu X., Yan K., Zhang T., Zhang Q., Zhang H., Liu J.** Quality evaluation of the pulse per second (PPS) signals from commercial GNSS receivers. *GPS Solutions*, 2015, No. 19(1), Pp. 141–150. DOI: 10.1007/s10291-014-0375-7
5. **Karutin S.N., Lerner D.V., Kharisov V.N.** Synthesis of synchronization algorithms based on the retransmission of navigation signals from a ground station [Synthesis of algorithms based on relaying navigation signals from a ground station]. *Radiotekhnika*, 2016, No. 9, Pp. 88–96. (rus)
6. **Ochieng W.Y., Sauer K.** GPS integrity and potential impact on aviation safety. *Journal of Navigation*, 2003, Vol. 56, No. 1, Pp. 51–65.
7. **Milner C., Macabiau C., Thevenon P.** Bayesian inference of GNSS failures. *Journal of Navigation*, 2016, No. 69(2), Pp. 277–294. DOI: 10.1017/S0373463315000697
8. **Heng L., Gao G.X., Walter T., Enge P.** Statistical characterization of GLONASS broadcast ephemeris errors. *Proceedings of the 24th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2011)*, Portland, OR, 2011, Pp. 3109–3117.
9. **Ahn J., Lee Y.J., Won D.H., Jun H.S., Yeom C., Sung S., Lee J.O.** Orbit ephemeris failure detection in a GNSS regional application. *International Journal of Aeronautical and Space Sciences*, 2015, No. 16(1), Pp. 89–101.
10. **Kulnev V., Mikhailov S.** Analiz napravleniy i sostoyaniya razrabotok funktsionalnykh dopolneniy k sputnikovym radionavigatsionnym sistemam [Analysis of state of the art of development of augmentation aids to satellite radio navigation systems]. *Besprovodnyye tekhnologii [Wireless Technologies]*, 2006, No. 4, Pp. 61–69. (rus)
11. **Amin M.G., Closas P., Broumandan A., Volakis J.L.** Vulnerabilities, threats, and authentication in satellite-based navigation systems. *Proceedings of the IEEE*, 2016, No. 104(6), Pp. 1169–1173. DOI: 10.1109/JPROC.2016.2550638
12. **Bao L., Wu R., Wang W., Lu D.** Spoofing mitigation in Global Positioning System based on C/A code self-coherence with array signal processing. *Journal of Communications Technology and Electronics*, 2017, No. 62(1), Pp. 66–73.
13. **Broumandan A., Jafarnia-Jahromi A., Daneshmand S., Lachapelle G.** Overview of spatial processing approaches for GNSS structural interference detection and mitigation. *Proceedings of the IEEE*, 2016, No. 104(6), Pp. 1246–1257. DOI: 10.1109/JPROC.2016.2529600
14. **Wesson K.D., Gross J.N., Humphreys T.E., Evans B.L.** GNSS signal authentication via power and distortion monitoring. *IEEE Transactions on Aerospace and Electronic Systems*, 2018, No. 54(2), Pp. 739–754.
15. **Zhang Z., Zhan X.** GNSS spoofing network monitoring based on differential pseudorange. *Sensors*, 2016, Vol. 16, No. 10, 1771. DOI: 10.3390/s16101771
16. **Wang F., Li H., Lu M.** GNSS spoofing countermeasure with a single rotating antenna. *IEEE Access*, 2017, Vol. 5, Pp. 8039–8047.
17. **Ochin E.** *Detection of spoofing using differential GNSS*. Zeszyty Naukowe Akademii Morskiej w Szczecinie, 2017.
18. **Khanafseh S., Roshan N., Langel S., Chan F.C., Joerger M., Pervan B.** GPS spoofing detection using RAIM with INS coupling. *Proceedings of the Position, Location and Navigation Symposium*, 2014, Vol. 2014.

19. Hewitson S., Wang J. Extended receiver autonomous integrity monitoring (E-RAIM) for GNSS/INS integration. *Journal of Surveying Engineering*, 2010, Vol. 136, No. 1, Pp. 13–22.
20. Ioannides R.T., Pany T., Gibbons G. Known vulnerabilities of global navigation satellite systems, status, and potential mitigation techniques. *Proceedings of the IEEE*, 2016, No. 104(6), Pp. 1174–1194.
21. Van der Merwe J.R., Zubizarreta X. Classification of spoofing attack types. *2018 European Navigation Conference (ENC)*, IEEE, 2018, Pp. 91–99.
22. Stubberud S.C., Kramer K.A. Threat assessment for GPS navigation. *Proceedings of the 2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications*, IEEE, 2014, Pp. 287–292. DOI: 10.1109/INISTA.2014.6873632
23. Jovanovic A., Botteron C., Fariné P.A. Multi-test detection and protection algorithm against spoofing attacks on GNSS receivers. *Position, Location and Navigation Symposium (PLANS) 2014*, 2014 IEEE/ION, Pp. 1258–1271.
24. Amin M.G., Closas P., Broumandan A., Volakis J.L. Vulnerabilities, threats, and authentication in satellite-based navigation systems. *Proceedings of the IEEE*, 2016, No. 104(6), Pp. 1169–1173. DOI: 10.1109/JPROC.2016.2550638
25. Van Trees H.L. *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2004.
26. Montgomery P.Y., Humphreys T.E., Ledvina B.M. Receiver-autonomous spoofing detection: Experimental results of a multi-antenna receiver defense against a portable civil GPS spoofer. *Proceedings of the 2009 International Technical Meeting of the Institute of Navigation*, 2009, Pp. 124–130.
27. Melikhova A.P., Tsikin I.A. Optimum array processing with unknown attitude parameters for GNSS anti-spoofing integrity monitoring. *Proceedings of the 2018 41st International Conference on Telecommunications and Signal Processing (TSP)*, IEEE, 2018, Pp. 1–4.
28. Rachitskaya A.P., Tsikin I.A. GNSS integrity monitoring in case of a priori uncertainty about user's coordinates. *Proceedings of the 2018 IEEE International Conference on Electrical Engineering and Photonics (EExPolytech)*, IEEE, 2018, Pp. 83–87.
29. Tippenhauer N.O., Pöpper C., Rasmussen K.B., Capkun S. On the requirements for successful GPS spoofing attacks. *Proceedings of the 18th ACM Conference on Computer and Communications Security*, ACM, 2011, Pp. 75–86. DOI: 10.1145/2046707.2046719
30. Tsikin I., Shcherbinina E. GPS antenna array calibration for attitude determination based on reference phase difference method. *Proceedings of the 2016 39th International Conference on Telecommunications and Signal Processing (TSP)*, IEEE, 2016, Pp. 174–177. DOI: 10.1109/TSP.2016.7760853

Received 17.05.2021.

СПИСОК ЛИТЕРАТУРЫ

1. Teunissen P.J.G., Montenbruck O. (eds.). Springer handbook of global navigation satellite systems. Springer, 2017.
2. Cheng J., Wang J., Zhao L. A direct attitude determination approach based on GPS double-difference carrier phase measurements // *J. of Applied Mathematics*. 2014. No. 6. Pp. 1–6. DOI: 10.1155/2014/548083
3. Daneshmand S., Sokhandan N., Lachapelle G. Precise GNSS attitude determination based on antenna array processing // *Proc. of the 27th Internat. Technical Meeting of the Satellite Division of the Institute of Navigation*. Tampa, FL, USA, 2014. Vol. 812.
4. Niu X., Yan K., Zhang T., Zhang Q., Zhang H., Liu J. Quality evaluation of the pulse per second (PPS) signals from commercial GNSS receivers // *GPS Solutions*. 2015. No. 19 (1). Pp. 141–150. DOI: 10.1007/s10291-014-0375-7

5. **Карутин С.Н., Лернер Д.В., Харисов В.Н.** Синтез алгоритмов синхронизации на основе ретрансляции навигационных сигналов с наземной станции // *Радиотехника*. 2016. № 9. С. 88–96.
6. **Ochieng W.Y., Sauer K.** GPS integrity and potential impact on aviation safety // *J. of Navigation*. 2003. Vol. 56. No. 1. Pp. 51–65.
7. **Milner C., Macabiau C., Thevenon P.** Bayesian inference of GNSS failures // *J. of Navigation*. 2016. No. 69 (2). Pp. 277–294. DOI: 10.1017/S0373463315000697
8. **Heng L., Gao G.X., Walter T., Enge P.** Statistical characterization of GLONASS broadcast ephemeris errors // *Proc. of the 24th Internat. Technical Meeting of the Satellite Division of the Institute of Navigation*. Portland, OR, 2011. Pp. 3109–3117.
9. **Ahn J., Lee Y.J., Won D.H., Jun H.S., Yeom C., Sung S., Lee J.O.** Orbit ephemeris failure detection in a GNSS regional application // *Internat. J. of Aeronautical and Space Sciences*. 2015. No. 16 (1). Pp. 89–101.
10. **Кульнев В., Михайлов С.** Анализ направлений и состояния разработок функциональных дополнений к спутниковым радионавигационным системам // *Беспроводные технологии*. 2006. № 4. С. 61–69.
11. **Amin M.G., Closas P., Broumandan A., Volakis J.L.** Vulnerabilities, threats, and authentication in satellite-based navigation systems // *Proc. of the IEEE*. 2016. No. 104 (6). Pp. 1169–1173. DOI: 10.1109/JPR-OC.2016.2550638
12. **Bao L., Wu R., Wang W., Lu D.** Spoofing mitigation in Global Positioning System based on C/A code self-coherence with array signal processing // *J. of Communications Technology and Electronics*. 2017. No. 62 (1). Pp. 66–73.
13. **Broumandan A., Jafarnia-Jahromi A., Daneshmand S., Lachapelle G.** Overview of spatial processing approaches for GNSS structural interference detection and mitigation // *Proc. of the IEEE*. 2016. No. 104 (6). Pp. 1246–1257. DOI: 10.1109/JPROC.2016.2529600
14. **Wesson K.D., Gross J.N., Humphreys T.E., Evans B.L.** GNSS signal authentication via power and distortion monitoring // *IEEE Transactions on Aerospace and Electronic Systems*. 2018. No. 54 (2). Pp. 739–754.
15. **Zhang Z., Zhan X.** GNSS spoofing network monitoring based on differential pseudorange // *Sensors*. 2016. Vol. 16. No. 10. 1771. DOI: 10.3390/s16101771
16. **Wang F., Li H., Lu M.** GNSS spoofing countermeasure with a single rotating antenna // *IEEE Access*. 2017. Vol. 5. Pp. 8039–8047.
17. **Ochin E.** Detection of spoofing using differential GNSS. *Zeszyty Naukowe Akademii Morskiej w Szczecinie*, 2017.
18. **Khanafseh S., Roshan N., Langel S., Chan F.C., Joerger M., Pervan B.** GPS spoofing detection using RAIM with INS coupling // *Proc. of the Position, Location and Navigation Symp.* 2014. Vol. 2014.
19. **Hewitson S., Wang J.** Extended receiver autonomous integrity monitoring (E-RAIM) for GNSS/INS integration // *J. of Surveying Engineering*. 2010. Vol. 136. No. 1. Pp. 13–22.
20. **Ioannides R.T., Pany T., Gibbons G.** Known vulnerabilities of global navigation satellite systems, status, and potential mitigation techniques // *Proc. of the IEEE*. 2016. No. 104 (6). Pp. 1174–1194.
21. **Van der Merwe J.R., Zubizarreta X.** Classification of spoofing attack types // *Proc. of the 2018 European Navigation Conf. IEEE*, 2018. Pp. 91–99.
22. **Stubberud S.C., Kramer K.A.** Threat assessment for GPS navigation // *Proc. of the 2014 IEEE Internat. Symp. on Innovations in Intelligent Systems and Applications. IEEE*, 2014. Pp. 287–292. DOI: 10.1109/INISTA.2014.6873632
23. **Jovanovic A., Botteron C., Fariné P.A.** Multi-test detection and protection algorithm against spoofing attacks on GNSS receivers // *Position, Location and Navigation Symp.* 2014. IEEE/ION, 2014. Pp. 1258–1271.
24. **Amin M.G., Closas P., Broumandan A., Volakis J.L.** Vulnerabilities, threats, and authentication in satellite-based navigation systems // *Proc. of the IEEE*. 2016. No. 104 (6). Pp. 1169–1173. DOI: 10.1109/JPROC.2016.2550638

25. **Van Trees H.L.** Optimum array processing: Part IV of detection, estimation, and modulation theory. John Wiley & Sons, 2004.
26. **Montgomery P.Y., Humphreys T.E., Ledvina B.M.** Receiver-autonomous spoofing detection: Experimental results of a multi-antenna receiver defense against a portable civil GPS spoofer // Proc. of the 2009 Internat. Technical Meeting of the Institute of Navigation. 2009. Pp. 124–130.
27. **Melikhova A.P., Tsikin I.A.** Optimum array processing with unknown attitude parameters for GNSS anti-spoofing integrity monitoring // Proc. of the 2018 41st Internat. Conf. on Telecommunications and Signal Processing. IEEE, 2018. Pp. 1–4.
28. **Rachitskaya A.P., Tsikin I.A.** GNSS integrity monitoring in case of a priori uncertainty about user's coordinates // Proc. of the 2018 IEEE Internat. Conf. on Electrical Engineering and Photonics (EExPolytech). IEEE, 2018. Pp. 83–87.
29. **Tippenhauer N.O., Pöpper C., Rasmussen K.B., Capkun S.** On the requirements for successful GPS spoofing attacks // Proc. of the 18th ACM Conf. on Computer and Communications Security. ACM, 2011. Pp. 75–86. DOI: 10.1145/2046707.2046719
30. **Tsikin I., Shcherbinina E.** GPS antenna array calibration for attitude determination based on reference phase difference method // Proc. of the 2016 39th Internat. Conf. on Telecommunications and Signal Processing. IEEE, 2016. Pp. 174–177. DOI: 10.1109/TSP.2016.7760853

Статья поступила в редакцию 17.05.2021.

THE AUTHOR / СВЕДЕНИЯ ОБ АВТОРЕ

Rachitskaya Antonina P.
Рачицкая Антонина Павловна
E-mail: antonina_92@list.ru

© Санкт-Петербургский политехнический университет Петра Великого, 2021

DOI: 10.18721/JCSTCS.14305
УДК 621.396

INFLUENCE OF TRANSIENTS IN THE INFORMATION PROCESSING CHANNEL OF THE AIRPORT AUTOMATIC RADIO DIRECTION FINDER ON THE DIRECTION FINDING ACCURACY

G.K. Aslanov¹, T.G. Aslanov¹, R.B. Kazibekov², R.R. Musaibov¹

¹ Daghestan State Technical University,
Makhachkala, Republic of Daghestan, Russian Federation;

² Joint-stock Company Derbent Scientific Research Institute "Volna",
Derbent, Republic of Daghestan, Russian Federation

The article is devoted to the study of the transients influence in the information processing channel of an aerodrome automatic direction finder (ADF) on the accuracy of direction finding. This task is of great importance, since improving the accuracy of navigation equipment allows setting the separation standards higher and improving the aircraft flights safety. When the phase difference between the signals on the neighboring vibrators of the antenna system (AS) is equal to 180°, a signal loss is observed in the low-frequency filter of the ADF due to the amplitudes equality of the component signals from the k^{th} and $k + 1^{\text{th}}$ vibrators, that leads to failures (the appearance of abnormal errors) in the ADF operation. When using the radio direction finders operation to find speech-modulated signals (an amplitude-modulated signal), certain gaps emerge in the direction-finding signal due to the operation of automatic gain control in the ADF receiver, which leads to the accuracy deterioration of the direction-finding. We propose methods of reducing the transients influence on the accuracy of radio direction finding. To eliminate the influence of transients caused by the operation of the Automatic Gain Control, taking into account the fact that the processing of direction finding information in the ADF is carried out on a channel microprocessor, it is necessary to assign weight coefficients to the directions calculated for eight switching cycles of the AS elements and calculate the value of the weighted average finding.

Keywords: radio direction finder, antenna system, transients, finding accuracy, radio direction finder failures.

Citation: Aslanov G.K., Aslanov T.G., Kazibekov R.B., Musaibov R.R. Influence of transients in the information processing channel of the airport automatic radio direction finder on the direction finding accuracy. Computing, Telecommunications and Control, 2021, Vol. 14, No. 3, Pp. 56–63. DOI: 10.18721/JCST-CS.14305

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

ВЛИЯНИЕ ПЕРЕХОДНЫХ ПРОЦЕССОВ В КАНАЛЕ ОБРАБОТКИ ИНФОРМАЦИИ АЭРОДРОМНОГО АВТОМАТИЧЕСКОГО РАДИОПЕЛЕНГАТОРА НА ТОЧНОСТЬ ПЕЛЕНГОВАНИЯ

Г.К. Асланов¹, Т.Г. Асланов¹, Р.Б. Казибеков², Р.Р. Мусаибов¹

¹ Дагестанский государственный технический университет,
г. Махачкала, Республика Дагестан, Российская Федерация;

² Дербентский НИИ «Волна»,
г. Дербент, Республика Дагестан, Российская Федерация

Статья посвящена исследованию влияния переходных процессов в канале обработки информации аэродромного автоматического радиопеленгатора (АРП) на точность пеленгования. Эта задача большой важности, так как повышение точности навигационного оборудования позволяет ужесточить нормы эшелонирования и повысить безопасность полетов воздушных судов. При разности фаз между сигналами на соседних вибраторах АС, равной 180° , в низкочастотном фильтре АРП наблюдается пропадание сигнала из-за равенства амплитуд составляющих сигналов с пеленгования. Для устранения влияния переходных процессов, вызываемых работой АРУ (Автоматическая регулировка усиления), с учетом того, что обработка пеленгационной информации в АРП осуществляется на канальном микропроцессоре, необходимо к вычисленным за восемь циклов коммутации элементов АС пеленгам присвоить весовые коэффициенты и вычислить значение средне-взвешенного пеленга.

Ключевые слова: радиопеленгатор, антенная система, переходные процессы, точность пеленгования, сбои в работе радиопеленгатора.

Ссылка при цитировании: Aslanov G.K., Aslanov T.G., Kazibekov R.B., Musaibov R.R. Influence of transients in the information processing channel of the airport automatic radio direction finder on the direction finding accuracy // Computing, Telecommunications and Control. 2021. Vol. 14. No. 3. Pp. 56–63. DOI: 10.18721/JCSTCS.14305

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>).

Introduction

Automatic direction finders (ADF) are widely used in navigation, radio monitoring, monitoring of rare animals movement, in seismic hazard assessment, etc., [1–6].

The growth of interregional and international traffic leads to the increase in the intensity of air traffic, which requires further improvement of radio equipment accuracy for flight support, in particular airfield automatic direction finders.

In this regard, priority attention is given to the issues of ensuring the accuracy characteristics of the ADF [7–14].

One of the reasons that reduce the ADF direction finding accuracy is the presence of transients in the information processing channel of the radio direction finder.

In the ADF, the direction-finding signal, in order to transfer to the phase relations stable frequency that occur in the antenna system (AS) in the radio direction finders (ADF-75, ADF-80K), is modulated at a frequency of 4200 Hz, and in the radio direction finders ADF AS (“Platan”, DF-2000, “Pikhta”) at a frequency of 5550 Hz [1].

Subsequently the signal is demodulated and filtered. The values of the phase differences between the filtered and reference signals determine the direction finding on the radiation source [16].

The narrowing of the filter bandwidth allows reducing the interference influence on the direction finding accuracy which simultaneously leads to the increase in the transients duration caused by switching the AS vibrators.

Fig. 1 explains the switching effect of the ADF AS vibrators on the accuracy of the direction finding. Fig. 1a shows the U_{in} signals induced on the k and $k + 1$ ADF AS vibrators when they are switched sequentially (the moment of switching the vibrators on the time axis is marked with a dot). Fig. 1b shows the signal induced on the k vibrator after it passes through the bandpass filter. Fig. 1c shows the signal at the bandpass filter output when the $k + 1$ vibrator is turned on.

After switching the vibrator, the filter output has a total signal U_{out} (Fig. 1d), equal to:

$$U_{out} = U_k + U_{k+1}. \quad (1)$$

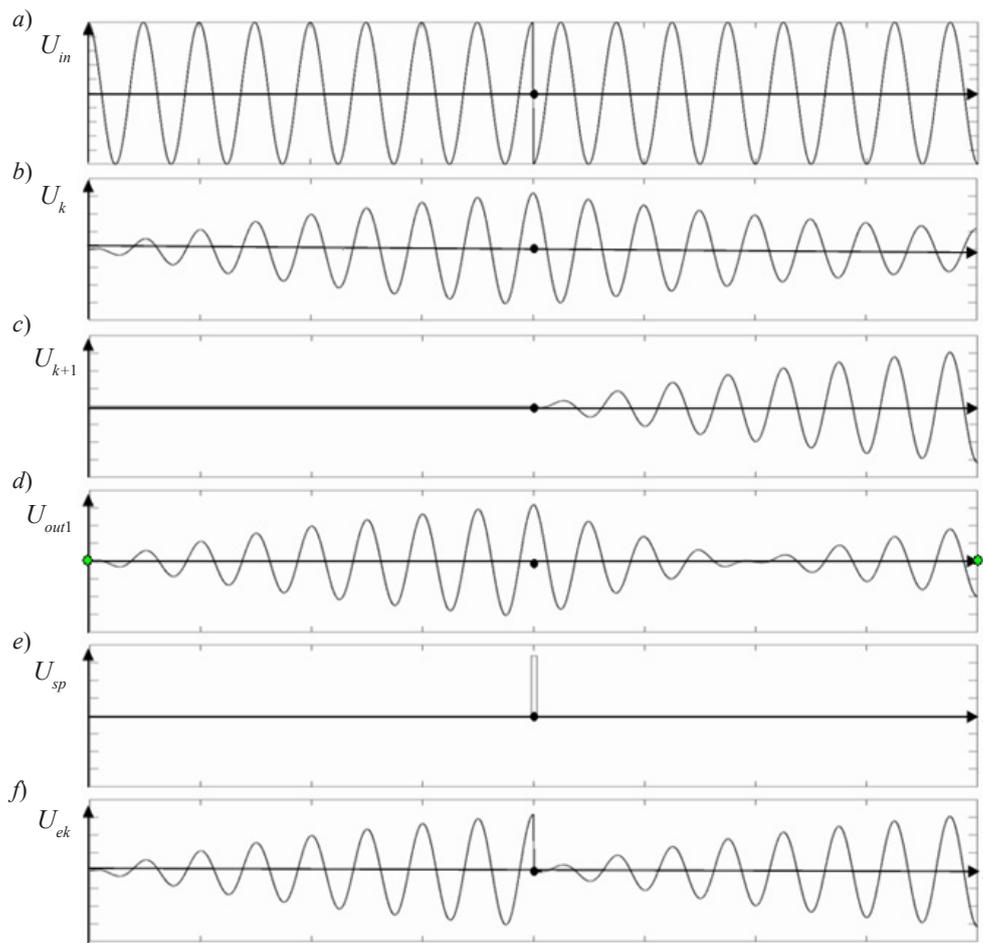


Fig. 1. The effect of radio direction finder vibrators switching on its accuracy

The Fig. 1f shows the case when the phase difference between the signals from the k and $k + 1$ vibrators is 180° , and if the signals are equal their sum is zero.

The phase difference $\varphi_{k+1} - \varphi_k$ is determined by the expression:

$$\varphi_{k+1} - \varphi_k = \frac{2\pi R}{\lambda} \cos \beta \left[\cos \left(\theta - \frac{2\pi k}{N} \right) - \cos \left(\theta - \frac{2\pi(k-1)}{N} \right) \right], \quad k = 1, \dots, N, \quad (2)$$

where R – AS radius; β – angle of the position on the radiation source relative to the ADF AS; λ – direction-finding signal wavelength; θ – azimuth (direction finding) on the radiation source; N – AS elements number.

In accordance with expression (2), for example, between 4 and 5 vibrators at a frequency of 300 MHz, a phase difference of 180° occurs at a direction finding of 0° and a position angle of 48° .

In addition, due to the vibrators switching the phase of the direction-finding signal is distorted, since its initial phase at the filter output will always be zero regardless of the input signal initial phase, which is confirmed by Fig. 1c,d. The use of such phases in direction finding calculation leads to significant errors and even to failures.

The accuracy of the bearing is also affected by the transients caused by the automatic gain control (AGC) operation of the ADF receiver.

The AGC leads to the gaps in the direction-finding high-frequency signal.

Fig. 2 shows the output signal waveform of radio station “Baklan” receiving channel, which explains the effect of the AGC on the transition process [15].

The figure shows that at the moment of the signal appearance, amplitude jumps appear at the input of the receiving device, and therefore the signal phases, which is explained by the sharp change in the AGC gain.

Reducing the influence of transients in the ADF information processing channel on the direction finding accuracy

Failures caused by switching of the AS vibrators can be eliminated by discharging the bandpass filter reactive elements at the moment of the ADF antenna system vibrators switching [17].

Fig. 3 shows the operating principle of such a filter.

Here, when switching the vibrators, a short pulse of U_{sp} is generated from the switching signal (Fig. 1e), which is fed to the electronic key. When a short pulse is received the electronic key discharges the capacitors.

Thanks to this, the filter is damped and the parasitic component of the signal from the previous vibrator quickly fades out.

After damping, the output signal of the U_{ek} filter has no interference component from the previous vibrator signal.

During the switching time of one AS vibrator pass through the filter eight periods allocated by the filter of the low-frequency signal, the phase of which is restored by the end of the vibrator switching time, so in the processing of direction finding information, it is necessary to use the averaged values of the phase differences of the seventh and eighth period signals [18].

In the ADF in operation since 2000, the direction finding information is processed in a channel microcomputer.

This allows reducing the direction finding error caused by the AGC operation by pre-processing the array of phase differences of the ARP AS vibrators signals.

It is experimentally established (Fig. 2) that the duration of the transient process in the receiving device can reach 0.2–0.3 sec. with the required ADF speed of 0.5 sec.

During operation, ADF switches 8 cycles of the antenna system elements in 0.5 sec., i.e., the direction finding on the radio source is calculated eight times. Due to the AGC operation at the starting moment of

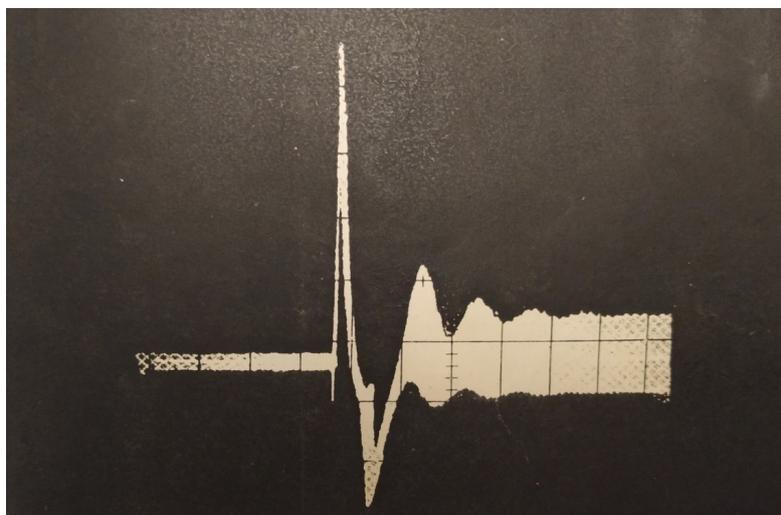


Fig. 2. Waveform of the output signal of radio station “Baklan” receiving channel

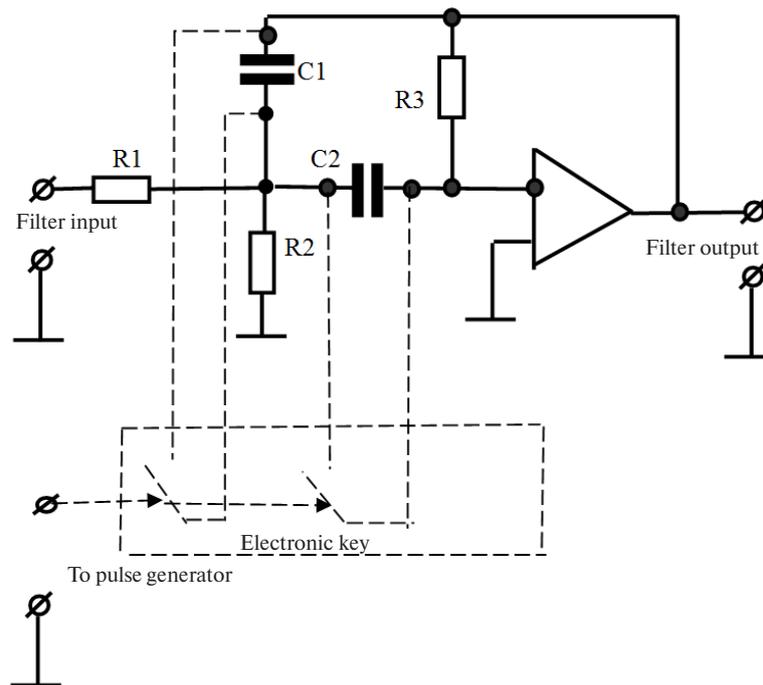


Fig. 3. Electronic key filter

the direction-finding onboard radio station operation, the signals phases from the AS vibrators have large jumps, hence the direction findings are calculated with large errors [19]. Therefore, it is necessary to assign weight coefficients to the calculated direction findings and calculate the value of the weighted average direction finding. At the same time the first direction findings should have minimum coefficients and the last ones should have maximum coefficients.

Conclusion

Transients that occur in the automatic radio direction finder when the direction-finding signal passes through the information processing path are one of the reasons for the occurrence of abnormal errors in the radio direction finder operation.

ADF failures caused by AS vibrators switching can be eliminated by discharging the reactive elements of the bandpass filter at the moments of ARP antenna system vibrators switching.

Processing of direction finding information in the ADF is carried out on a channel microprocessor which allows using software tools that significantly reduce errors caused by transients in the radio direction finder.

REFERENCES

1. Aslanov T.G., Gasanov O.I., Kazibekov R.B., Musaibov R.R. Pelengator ionosferynykh predvestnikov zemletryaseniya [Direction finder of ionospheric earthquake precursors]. *Inzhenernyy Vestnik Dona*, 2020, No. 4. (rus). Available: ivdon.ru/ru/magazine/archive/n4y2020/6398
2. Saidov A.S., Tagilayev A.R., Aliyev N.M., Aslanov G.K. *Design of phase automatic radio direction finders*. Moscow: Radio i svyaz Publ., 1997, 160 p. (rus)
3. Vasin A.A., Ponomareva L.I., Cheremisin O.P. Vysokotochnoye pelengovaniye proizvolno korrelirovannykh mnogoluchevykh signalov s ispolzovaniyem tsifrovyykh antenykh reshetok [High-precision direction

finding of arbitrarily correlated multipath signals using digital antenna arrays]. *Radiotekhnika i Elektronika* [Radio Engineering and Electronics], 2015, No. 12 (60), Pp. 12–37. (rus). DOI: 10.7868/S0033849415120190

4. **Kukes I.S., Starik M.Ye.** *Osnovy radiopelengatsii* [Fundamentals of radio direction finding]. Moscow: Sov. Radio Publ., 1964. (rus)

5. **Ziskind I., Bar-Ness Y.** Direction finding of narrowband autoregressive sources by antenna arrays. *Antennas and Propag.: Int. Symp. Dig. Merg. Technol. 90's, Dallas, Tex., May 7-11, 1990*, Vol. 4. Piscataway (N.J.), 1990.

6. **Johnson J.** R&S direction finders for Her Majesty's Coastguard. *News from Rohde & Schwarz*, 1985, No. 109, Pp. 36–37.

7. **Aslanov G.K., Aslanov T.G., Tetakayev U.R., Kazibekov R.B.** Otsenka oshibok, vyzyvayemykh vykhodom iz stroya elementov antennoy sistemy aerodromnogo avtomaticheskogo radiopelengatora [Evaluation of errors caused by the failure of the aerodrome automatic direction finder antenna system elements]. *Vestn. Dagestanskogo Tekhnicheskogo Universiteta. Tekhnicheskkiye Nauki*, 2018, Vol. 45 (2), Pp. 94–103. (rus)

8. **Belyavskiy L.S., Chutkiy I.P.** K voprosu ob otsenke vliyaniya otrazheniy ot mestnykh predmetov na tochnost radiopelengovaniya [On the question of assessing the impact of reflections from local objects on the accuracy of radio direction finding]. *Radiotekhnicheskoye oborudovaniye aeroportov i vozdushnykh trass. Mezhvuz. sb. nauch. tr. Kiyev: Kniga Publ.*, 1981, Pp. 54–59. (rus)

9. **Dubrovin A.V.** Otsenka srednekvadraticheskoy oshibki izmereniya pelenga putem vychisleniya nevyazok faz [Estimation of the root-mean-square error of direction finding measurement by calculating phase residuals]. *Radiotekhnika i Elektronika* [Radio Engineering and Electronics], 2019, No. 8 (64), Pp. 796–799. (rus). DOI: 10.1134/S0033849419070076

10. **Ivanov N.M.** Adaptivnyye metody obnaruzheniya i pelengovaniya signalov [Adaptive methods of signal detection and direction finding]. *Radiotekhnika i Elektronika* [Radio Engineering and Electronics], 2016, No. 10 (61), Pp. 979–983. (rus). DOI: 10.7868/S0033849416100156

11. **Wang Fengzhen.** Direction-of-arrival estimation for narrow band coherent and incoherent sources in the presence of unknown noise fields. *Res. IEEE Int. Radar Conf., Arlington, Va, May 7-10, 1990*. New York, 1990.

12. **Marushchak A.I., Rasin A.M.** Sposoby povysheniya tochnosti pelengovaniya aerodromnykh UKV radiopelengatorov [Ways to improve the accuracy of the direction finding of airfield VHF radio direction finders] Ways to improve the accuracy of the direction finding of airfield VHF radio direction finders]. *Trudy NII Grazhdanskoy Aviatsii*, 1997, Vol. 136. (rus)

13. **Morozov R.O., Devit D.V.** Metody obrabotki navigatsionnoy informatsii v tselyakh povysheniya tochnosti [Methods of processing navigation information to improve accuracy]. *Inzhenernyy Vestnik Dona*, 2018, No. 1. (rus). Available: ivdon.ru/ru/magazine/archive/n2y2018/5027

14. **Holbrook J.G.** An analysis of errors in long range radio direction finder systems. *Proceedings of the IRE*, Dec. 1953, Vol. 41, Iss. 12, Pp. 1747–1749. DOI: 10.1109/JRPROC.1953.274360

15. **Aslanov G.K., Gasanov O.I.** Analiz prichin voznikoveniya anomalnykh oshibok v kvazidoplerovskikh avtomaticheskikh radiopelengatorakh [Analysis of the causes of anomalous errors in quasi-Doppler automatic radio direction finders]. *Nauchno-Tekhnicheskkiye Vedomosti SPbGPU, Informatika. Telekommunikatsii. Upravleniye*, 2009, No. 2, Pp. 87–93. (rus)

16. Izmereniye fazovogo sdviga. (rus). Available: <http://www.vevivi.ru/best/Izmerenie-fazovogo-sdviga-ref2-27640.html>.

17. **Koltik Ye.D.** *Fazosdvygayushchiye ustroystva* [Phase-shifting devices]. Moscow: Izd-vo standartov Publ., 1981, 164 p. (rus)

18. Kwartsevyye filtry s peremennoy polosoy propuskaniya [Quartz filters with variable bandwidth]. (rus). Available: <http://ra3ggi.qrz.ru/UZLY/r230682.htm#>

19. **Zhalnerauskas V.** *Uzkopolosnyye kvartsevyye filtry na odinakovykh rezonatorakh* [Narrow-band quartz filters on identical resonators]. *Radio*, 1982, No. 1, 2. (rus)

Received 25.04.2021.

СПИСОК ЛИТЕРАТУРЫ

1. **Асланов Т.Г., Гасанов О.И., Казибеков Р.Б., Мусаилов Р.Р.** Пеленгатор ионосферных предвестников землетрясений // Инженерный вестник Дона. 2020. № 4 // URL: ivdon.ru/ru/magazine/archive/n4y2020/6398
2. **Саидов А.С., Тагилаев А.Р., Алиев Н.М., Асланов Г.К.** Проектирование фазовых автоматических радиопеленгаторов. М.: Радио и связь, 1997. 160 с.
3. **Васин А.А., Пономарева Л.И., Черемисин О.П.** Высокоточное пеленгование произвольно коррелированных многолучевых сигналов с использованием цифровых антенных решеток // Радиотехника и электроника. 2015. № 12 (60). С. 12–37. DOI: 10.7868/S0033849415120190
4. **Кукес И.С., Старик М.Е.** Основы радиопеленгации. М.: Сов. радио, 1964.
5. **Ziskind I., Bar-Ness Y.** Direction finding of narrowband autoregressive sources by antenna arrays // Antennas and Propag.: Int. Symp. Dig. Merg. Technol. 90's, Dallas, Tex., May 7-11, 1990. Vol. 4. Piscataway (N.J.), 1990.
6. **Johnson J.** R&S direction finders for Her Majesty's Coastguard // News from Rohde & Schwarz. 1985. No. 109. Pp. 36–37.
7. **Асланов Г.К., Асланов Т.Г., Тетакаев У.Р., Казибеков Р.Б.** Оценка ошибок, вызываемых выходом из строя элементов антенной системы аэродромного автоматического радиопеленгатора // Вестн. Дагестанского технического университета. Технические науки. 2018. Вып. 45 (2). С. 94–103.
8. **Белявский Л.С., Чуткий И.П.** К вопросу об оценке влияния отражений от местных предметов на точность радиопеленгования // Радиотехническое оборудование аэропортов и воздушных трасс. Межвуз. сб. науч. тр. Киев: Книга, 1981. С. 54–59.
9. **Дубровин А.В.** Оценка среднеквадратической ошибки измерения пеленга путем вычисления невязок фаз // Радиотехника и электроника. 2019. № 8 (64). С. 796–799. DOI: 10.1134/S0033849419070076
10. **Иванов Н.М.** Адаптивные методы обнаружения и пеленгования сигналов // Радиотехника и электроника. 2016. № 10 (61). С. 979–983. DOI: 10.7868/S0033849416100156
11. **Wang Fengzhen.** Direction-of-arrival estimation for narrow band coherent and incoherent sources in the presence of unknown noise fields // Res. IEEE Int. Radar Conf., Arlington, Va, May 7-10, 1990. New York, 1990.
12. **Марущак А.И., Расин А.М.** Способы повышения точности пеленгования аэродромных УКВ радиопеленгаторов // Труды НИИ гражданской авиации. 1997. Вып. 136.
13. **Морозов Р.О., Девит Д.В.** Методы обработки навигационной информации в целях повышения точности // Инженерный вестник Дона. 2018. № 1 // URL: ivdon.ru/ru/magazine/archive/n2y2018/5027
14. **Holbrook J.G.** An analysis of errors in long range radio direction finder systems // Proc. of the IRE. Dec. 1953. Vol. 41. Iss. 12. Pp. 1747–1749. DOI: 10.1109/JRPROC.1953.274360
15. **Асланов Г.К., Гасанов О.И.** Анализ причин возникновения аномальных ошибок в квазидоплеровских автоматических радиопеленгаторах // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. 2009. № 2. С. 87–93.
16. Измерение фазового сдвига // URL: <http://www.vevivi.ru/best/Izmerenie-fazovogo-sdviga-ref2276-40.html>.
17. **Колтик Е.Д.** Фазосдвигающие устройства. М.: Изд-во стандартов, 1981. 164 с.
18. Кварцевые фильтры с переменной полосой пропускания // URL: <http://ra3ggi.qrz.ru/UZLY/r230682.htm#>
19. **Жалнераускас В.** Узкополосные кварцевые фильтры на одинаковых резонаторах // Радио. 1982. № 1, 2.

Статья поступила в редакцию 25.04.2021.

THE AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Aslanov Gajdarbek K.
Асланов Гайдарбек Кадырбекович
E-mail: uits@dstu.ru

Aslanov Tagirbek G.
Асланов Тагирбек Гайдарбекович
E-mail: tabasik@gmail.com

Kazibekov Rustam B.
Казибеков Рустам Бидирханович
E-mail: kazibus11@mail.ru

Musaibov Rashid R.
Мусаибов Рашид Рагимханович
E-mail: rashid_musaibov@mail.ru

© Санкт-Петербургский политехнический университет Петра Великого, 2021

DOI: 10.18721/JCSTCS.14306

УДК 621.375

ZERO-DRIFT OPERATIONAL AMPLIFIERS

*A. Assim^{1,2}, E.V. Balashov²*¹ Salahaddin University, Erbil, Iraq;² Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation

This article considers the design and implementation of four different zero-drift operational amplifiers with 50 nm technology CMOS and compares their characteristics. The aim is minimizing input offset voltage and flicker noise. Offset voltage is unavoidable in operational amplifiers, because no two transistors can be identical. A small difference in their dimensions (length or width) gives rise to this undesirable effect, the value of offset voltage in common operational amplifiers is less than 10 mV. In this article, two major techniques of dynamic offset cancellation, chopping and auto-zeroing, are considered. The operational amplifier with chopping shows the best result among the four amplifiers.

Keywords: zero-drift operational amplifiers, auto-zeroing, chopper amplifier, offset voltage reduction, CMOS.

Citation: Assim A., Balashov E.V. Zero-drift operational amplifiers. Computing, Telecommunications and Control, 2021, Vol. 14, No. 3, Pp. 64–74. DOI: 10.18721/JCST-CS.14306

This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

ОПЕРАЦИОННЫЕ УСИЛИТЕЛИ С НУЛЕВЫМ ДРЕЙФОМ

*А. Ассим^{1,2}, Е.В. Балашов²*¹ Университет им. Салахаддина, Эрбиль, Ирак;² Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация

Рассмотрена разработка и реализация четырех различных операционных усилителей с нулевым дрейфом с использованием технологии КМОП 50 нм. Целью является минимизация входного напряжения смещения и фликкерного шума. Напряжение смещения неизбежно в операционных усилителях, так как невозможно изготовить два транзистора одинаковых размеров. Небольшая разница в их размерах (длина или ширина) вызывает этот нежелательный эффект, значение напряжения смещения в обычных операционных усилителях меньше 10 мВ. В статье описаны два основных метода: коррекция дрейфа нуля на основе модуляции/демодуляции сигнала и периодическая коррекция дрейфа. Операционный усилитель с техникой коррекции дрейфа нуля на основе модуляции/демодуляции сигнала показал наилучшие результаты среди четырех усилителей.

Ключевые слова: операционные усилители с нулевым дрейфом, периодическая коррекция дрейфа, коррекция дрейфа нуля на основе модуляции и демодуляции сигнала, снижение напряжения смещения, КМОП.

Ссылка при цитировании: Assim A., Balashov E.V. Zero-drift operational amplifiers // Computing, Telecommunications and Control. 2021. Vol. 14. No. 3. Pp. 64–74. DOI: 10.18721/JCSTCS.14306

Статья открытого доступа, распространяемая по лицензии CC BY-NC 4.0 (<https://creativecommons.org/licenses/by-nc/4.0/>).

Introduction

Operational amplifiers are among the most versatile building blocks that are used in many modern analog and digital systems including filters, active rectifiers, current-to-voltage converters, etc. [3–18]. Due to their robust performance and righteous characteristics, they mimic an ideal amplifier, namely high input resistance, low output resistance, high gain, and stability [17]. Like in any other device, getting ideal characteristics in practice is impossible; therefore, tradeoffs must be made between the parameters (noise, linearity, gain, supply voltage, voltage swings, speed, input/output impedance and power dissipation) according to the required application [3]. Indeed, such tradeoffs pose many challenges to the designer, requiring solid knowledge and experience to reach an acceptable compromise.

Offset voltage is a dominant error source for operational amplifiers, especially at low frequencies. It has direct proportionality with flicker noise, hence in this article offset reduction is measured through reduction in flicker noise. Offset voltage exists due to mismatch in transistor sizes. Amplifiers without offset compensation have flicker noise as high as $400 \text{ nV}/\sqrt{\text{Hz}}$ [1, 2]. There are three major techniques that are commonly used to reduce offset voltage and flicker noise: trimming, auto-zeroing, and chopping. Trimming is done during fabrication to eliminate offset. In order to correct the dimensions of the circuit elements, laser is used. It does not belong to the category of zero-drift operational amplifiers. Because of that, trimming is not considered in this article. Auto-zeroing's principle of operation is based on sampling. The offset voltage is captured in one clock phase (a capacitor is charged to such value, with an opposite polarity) and then subtracted in the next clock phase. Chopping is based on continuous-time modulation. The input signal is modulated by the first chopper to a higher frequency. The offset voltage is added to it, then this signal will be amplified, and the second chopper modulates the offset voltage and demodulates the modulated input signal, hence the offset is converted to a higher frequency. Since the offset voltage undergoes modulation in chopper amplifier, a ripple is observed at the amplifier's output. Both mentioned techniques are dynamic techniques that continuously reduce offset. They also reduce low frequency noise and offset drift as a function of temperature or time, in this paper, auto-zeroing and chopping methods are implemented practically with different architectures. These include a fundamental auto-zeroing amplifier, continuous-time auto-zeroing, chopper amplifier and a combination of chopping and auto-zeroing together.

Fundamental auto-zeroing amplifier

A principal auto-zeroing amplifier is provided in Fig. 1. It consists of an operational amplifier with 5 transistors that act as switches, the differential input signals are given on the input. Clock signals (C1 and C2) are out of phase by 180 degrees. The transistors (M1, M3 and M4) are controlled by clock signal C1, while M0 and M6 transistors are controlled by clock signal C2. The output signal is single-ended, labeled as (Vout). The proposed amplifier operates in the following manner. On the first cycle when the input clock C1 is on (C2 is off), both of the differential inputs (labeled as V⁻ and V⁺) are shorted. The feedback loop (R0 resistor and transistor M3) is closed and the offset that appears at the output is fed back into the input. Thus, the capacitor (C0 – 5 nF) is charged to the offset voltage value (–10 mV), as shown in Fig. 2. On the other cycle when C2 is on (C1 is off), the amplifier works as usual, meaning that the differential inputs appear at the amplifier's input (they are not shorted). At the same time the capacitor charge compensates the offset voltage, because they are opposite in sign. This will result in zero offset voltage at the input. The capacitor takes around $350.2 \mu\text{s}$ to charge and begins to compensate the offset voltage.

The fundamental auto-zeroing amplifier isn't suitable for continuous-time applications, therefore, another architecture is introduced below, namely "Continuous-time auto-zeroing".

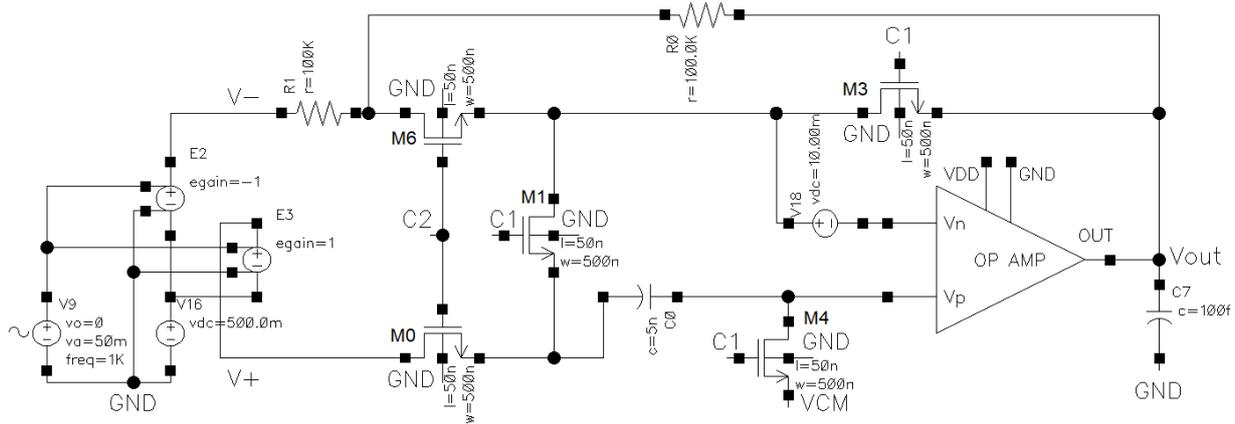


Fig. 1. Amplifier with auto-zeroing

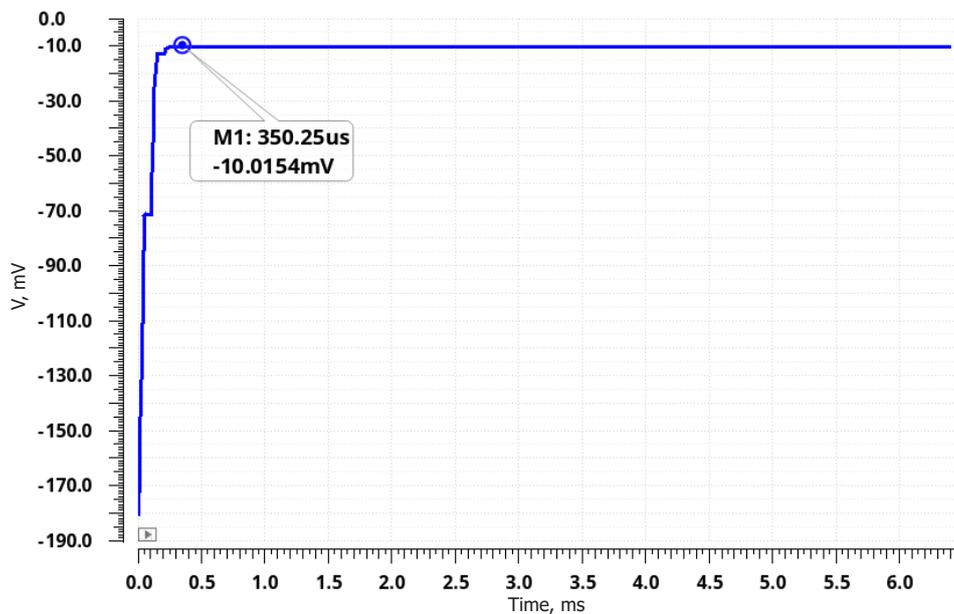


Fig. 2. Voltage on capacitor C0 versus time

Continuous-time auto-zeroing

The fundamental auto-zeroing amplifier was meant to be used in non-continuous applications. Except certain applications, it should not be used, when continuous-time signals are needed, as in voice amplifiers or analog-to-digital converters. A configuration exists that is known as continuous-time auto-zeroing amplifier (CTAZ or ping-pong amplifier). It is a general term that can be used for any amplifier that implements two identical sub-amplifiers with opposite clock pulses to achieve a continuous signal at the output. A realization of such amplifier is given in Fig. 3. It consists of two identical operational amplifiers, 2 feedback resistors (100 kΩ), 5 transistors that are controlled by two out of phase clock signals C1 and C2; the output is a single-ended signal taken from Vout pin. Its working principle can be summarized in two stages. The first stage, when C2 is high (C1 is low), the upper amplifier receives the signal from the differential inputs, amplifies it and feeds it to the output (Vout). The lower amplifier's inputs are shorted and it's in compensation mode. In the next stage, when C2 turns into low (C1 is high), the lower amplifier amplifies

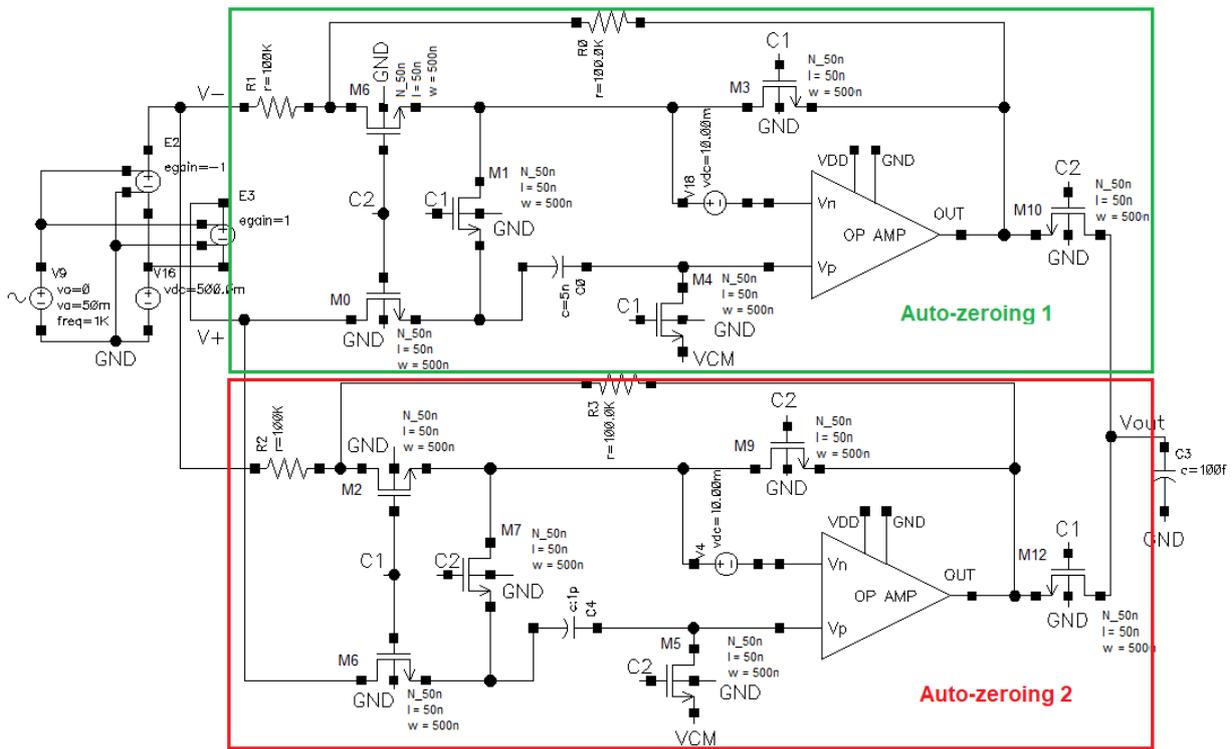


Fig. 3. Continuous-time auto-zeroing amplifier

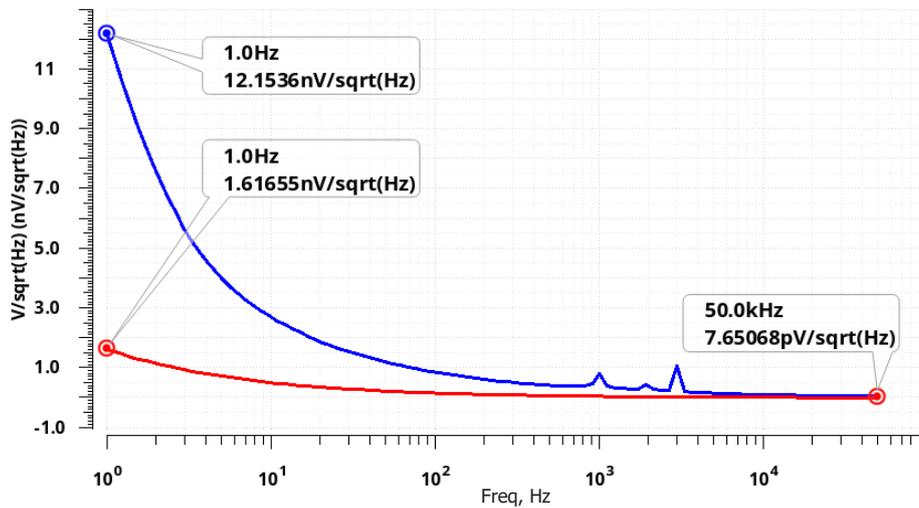


Fig. 4. PSS noise results for circuit in Fig. 3 before and after compensation (blue and red)

the signal, and the output voltage is taken from the lower amplifier. The upper amplifier compensates for offset, this results in a continuous signal at all times at the output. Fig. 4 shows flicker noise reduction.

Besides auto-zeroing, another popular offset reduction technique known as “chopping” exists. It has better flicker-noise reduction characteristics as demonstrated below.

Chopping

Chopping is a major offset-reduction technique. It is widely used to reduce offset voltage [4, 8–11, 20–24]. It is favorable in applications where a continuous-time signal is needed. Unlike auto-zeroing

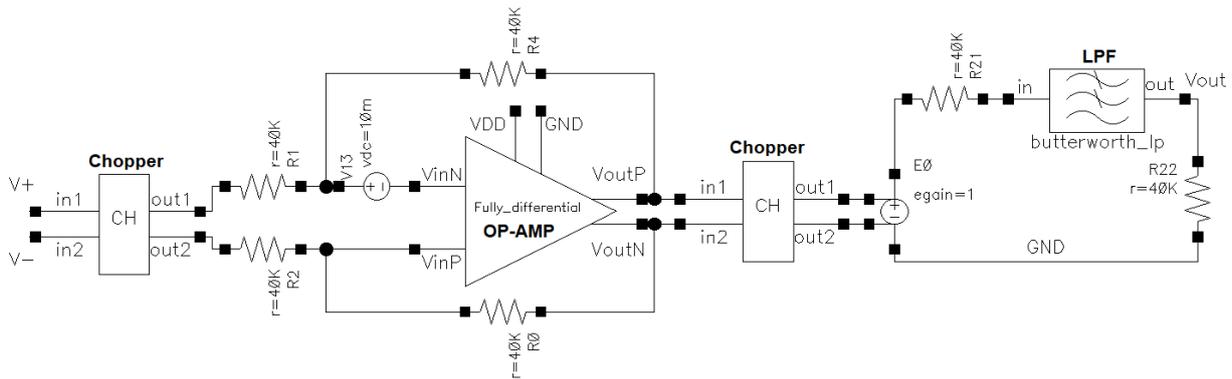


Fig. 5. Chopper amplifier circuit

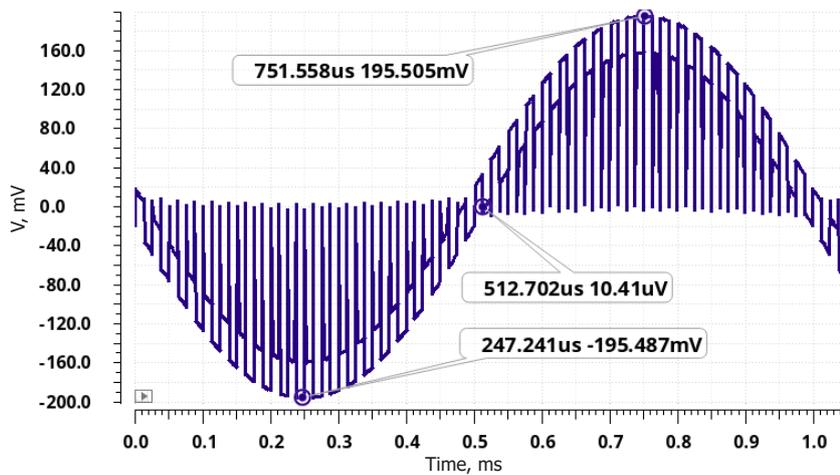


Fig. 6. Chopper amplifier's output signal

amplifiers, chopper amplifiers do not cause noise-folding. This method is based on modulation. The principle of operation is that the voltage V_{in} goes through the chopper that is driven by a clock at frequency f_{ch} . Hence, it is transformed to a modulated pulse voltage [11, 20–22]. Later, the modulated signal is amplified along with the input offset. The second chopper acts as a demodulator. It demodulates the input signal to a DC voltage, and concurrently modulates the offset to the odd harmonics of clock frequency that will be removed by a low-pass filter [13, 20–24]. In contrast to auto-zeroing, chopper amplifiers do not need any capacitor, they compensate offset voltage using modulation rather than charge compensation. A basic chopper amplifier is presented in Fig. 5. It consists of two modulators (choppers), a fully-differential op-amp, and a Butterworth LPF. As in the previous cases, the value of offset voltage is 10 mV, and it is added as a DC voltage source. The output signal is shown in Fig. 6. The signal looks like a sampled signal due to ripples.

Interestingly, both popular offset-voltage reduction techniques can be used together. This advanced architecture allows us to get the best of each technique.

Chopping & auto-zeroing

A combination of both chopping and auto-zeroing can be used to achieve better noise performance. That is a more sophisticated configuration despite its complexity, because auto-zeroing part gets rid of the voltage ripples caused by chopping, while chopping gets rid of the noise folding problem caused by auto-zeroing. An example of such topology is provided in Fig. 7.

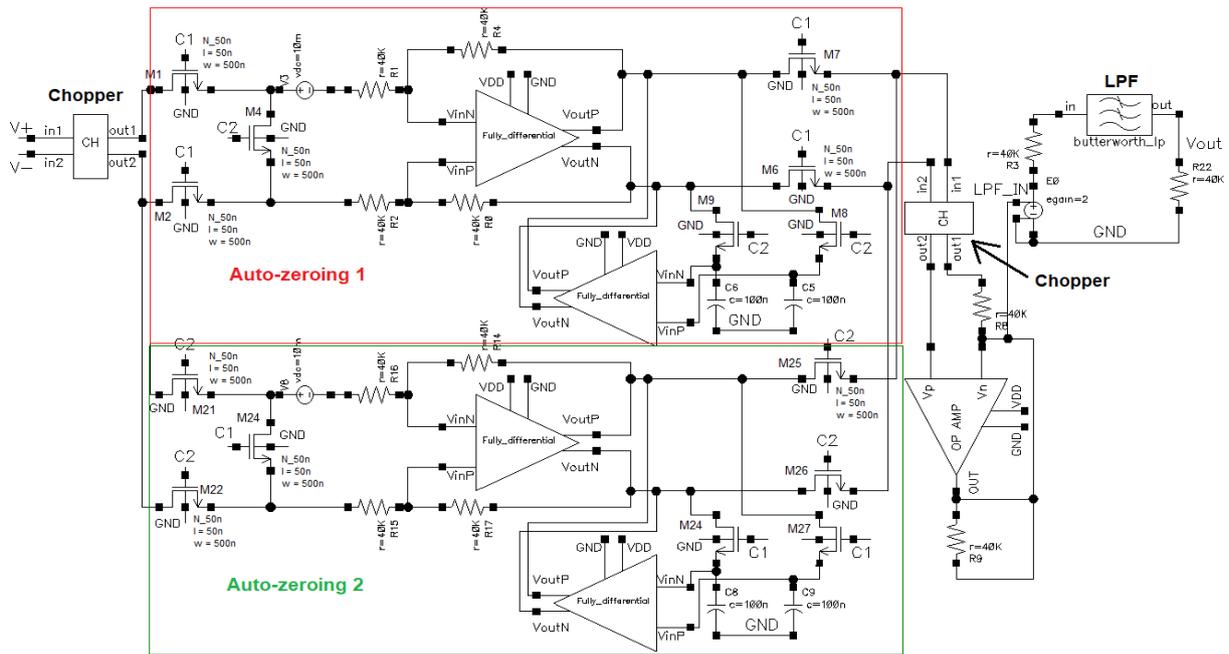


Fig. 7. Schematic of an amplifier using both chopping and auto-zeroing

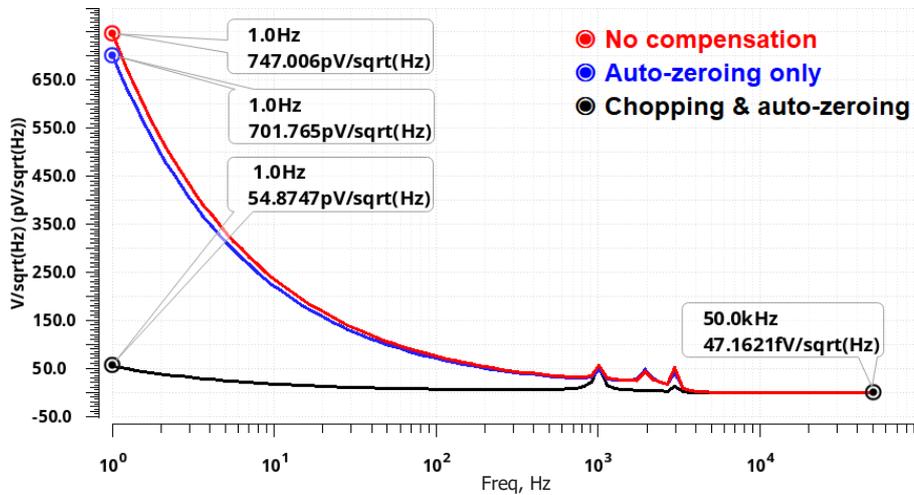


Fig. 8. Schematic of an amplifier using both chopping and auto-zeroing

The circuit consists of two choppers, they are placed at the input and output stages, the first chopper acts as a modulator, while the second chopper acts as a demodulator. Two auto-zeroing amplifiers (they are placed in red and green rectangles) are used between the choppers to further reduce flicker noise and reduce the ripples in the continuous-time signal; their operation can be explained simply by the two-phase nonoverlapping clock signals (C1 and C2). When C1 is one and C2 is zero, the upper auto-zeroing amplifier works in amplification mode while the lower auto-zeroing amplifier compensates the input offset of 10 mV. The feedback loop op-amp senses the voltage difference at the output of the main amplifier, then the capacitors are charged to this value. Later, they are amplified and subtracted from main amplifier's output. In the next clock period (when C1 is zero and C2 is one), the system works in a similar manner.

The lower amplifier operates in amplification mode and the upper amplifier compensates the offset. PSS noise analysis results for the circuit in Fig. 7 are provided in Fig. 8.

Table 1

Input-referred noise PSD, PSRR and CMRR

| Configuration | Input-referred noise PSD before compensation, nV/ $\sqrt{\text{Hz}}$ | Input-referred noise PSD before compensation, nV/ $\sqrt{\text{Hz}}$ | PSRR, dB | CMRR, dB |
|------------------------------|--|--|----------|----------|
| Auto-zeroing (AZ) | 9.3 | 2.5 | 104 | 98 |
| Continuous-time auto-zeroing | 9.4 | 0.02 | 103 | 97 |
| Chopping | 337 | 4.5 | 116 | 110 |
| Chopping & auto-zeroing | 0.74 | 0.054 | 112 | 106 |

Table 2

Comparison of the best performing circuit (chopping & auto-zeroing) with previous works

| Parameters | This work 2021 | [2] 2015 | [6] 2016 | [7] 2015 | [8] 2015 | [9] 2017 | [11] 2015 | [14] 2017 | [17] 2011 | [20] 2010 | [23] 2017 |
|---|----------------------|----------|--------------------|----------|----------------------|----------|-----------|-----------|-----------|-----------|-----------|
| Bandwidth, MHz | 47 | 2.8 | 10 | 10.3 | 3.1 | 43 | 0.329 | 1.5 | – | – | 30 |
| Bias current, μA | 4.6×10^{-5} | 420 | 5×10^{-5} | – | 1.5×10^{-5} | – | – | – | 175 | – | – |
| Chopping frequency, kHz | 20 | – | 150 | 10 | 800 | 500 | 500 | 1 | – | 10 | – |
| CMOS technology μm | 0.05 | 0.5 | 0.6 | 0.045 | 0.18 | 0.18 | 0.028 | 0.18 | 0.18 | 0.18 | 0.18 |
| CMRR, dB | 106.6 | 123 | > 120 | – | 145 | – | 70 | – | 90 | – | – |
| Voltage noise PSD, nV/ $\sqrt{\text{Hz}}$ | 0.05 | 90 | 6.5 | 2.54 | 6.8 | 3.5 | 27 | 200 | 15 | 179 | – |
| Maximum VOS, μV | 10.4 | 90 | – | – | – | – | – | – | 15 | – | 1296 |
| PSRR, dB | 112.6 | – | > 120 | – | 150 | – | 68 | 131 | 88 | – | – |
| Supply voltage, V | 1 | 5 | 5.5 | 1 | 60 | 5 | 0.9 | 1.8 | 3.3 | 1.8 | 1.2 |

Results

There are multiple metrics that are commonly mentioned in datasheets to evaluate the performance of operational amplifiers: power supply rejection ratio (PSRR), common-mode rejection ratio (CMRR) and input-referred noise power spectral density. These parameters are provided in Table 1. PSRR and CMRR can be obtained using these formulas:

$$\text{PSRR} = \frac{\Delta V_{DD}}{\Delta V_{OS}}, \tag{1}$$

$$\text{CMRR} = \frac{\Delta V_{CM}}{\Delta V_{OS}}, \quad (2)$$

where V_{CM} is the common-mode voltage; V_{OS} is the offset voltage and V_{DD} is the drain supply voltage (equal to 1 V in the proposed circuits). Additionally, the thermal performance was analyzed, the operational amplifiers work properly in the temperature range of $-40\text{ }^{\circ}\text{C}$ to $85\text{ }^{\circ}\text{C}$.

The results obtained in this work are compared with the existing works in Table 2.

Conclusion

Four zero-drift operational amplifiers were realized and compared. The operational amplifiers were simulated using Cadence Virtuoso software. The Periodic Steady State (PSS) analysis results showed that the proposed techniques are an effective way to reduce the input offset voltage and flicker (1/f) noise. Two operational amplifiers used in this work have a gain-bandwidth product (GBWP) of 47 MHz and 493 MHz, and an open-loop gain of 69.78 dB and 47 dB, respectively. The clock frequency of 20 kHz was chosen for all the circuits. In short, chopper amplifier and continuous-time auto-zeroing amplifier are reducing the flicker noise and input offset voltage more effectively. Chopper amplifier reduces flicker noise from $337\text{ nV}/\sqrt{\text{Hz}}$ to $4.5\text{ nV}/\sqrt{\text{Hz}}$, in other words by approximately 7500 %, while continuous-time auto-zeroing amplifier reduces flicker noise from $9.4\text{ nV}/\sqrt{\text{Hz}}$ to $0.02\text{ nV}/\sqrt{\text{Hz}}$, that is by 47000 %. But taking into account the simplicity, chopper amplifier is the best configuration. Nevertheless, choosing an operational amplifier for a certain application does not depend only on its flicker noise and offset voltage reduction capability. For instance, in continuous-time applications such as in analog-to-digital converters, continuous-time auto-zeroing must be used, while for amplifying a trigger signal from a sensor, a basic auto-zeroing amplifier can be used despite its humble qualities when compared to other more sophisticated amplifiers.

Acknowledgment

I would like to thank my parents for their never ending support. Special thanks to my dearest friend (Nikolai Kirichenko) for his help throughout my studies in St. Petersburg, Russia. Finally, I want to extend my thanks to the editors of this journal for taking time to review my work.

REFERENCES

1. Balashov E.V., Ivanov N.V., Korotkov A.S. SOI instrumentation amplifier for high-temperature applications. *2020 IEEE East-West Design and Test Symposium (EWDTS)*, Varna, Bulgaria, 2020. IEEE, 2020, Pp. 1–4. DOI: 10.1109/EWDTS50664.2020.9224893
2. Balashov E.V., Ivanov N.V., Akhmetov D.B., Korotkov A.S. High-temperature instrumentation amplifier. *Nanoindustry Journal*, 2020, Iss. 96 (3), Pp. 160–163. (rus)
3. Baker R.J. *CMOS: Circuit design, layout, and simulation*. 3rd ed., IEEE Press, 2010, Pp. 773–908.
4. Bortun N., Stan M.N., Brezeanu G. High precision bidirectional chopper amplifier with extended common mode input voltage range. *International Semiconductor Conference (CAS)*, Sinaia, Romania, 2015. IEEE, 2015, Pp. 297–300. DOI: 10.1109/SMICND.2015.7355237
5. Carusone T.C., Johns D.A., Martin K.W. *Analog integrated circuit design*. 2nd ed., Wiley, 2011, Pp. 242–293.
6. Duwe M., Chen T. Offset correction of low power, high precision op amp using digital assist for biomedical applications. *2012 IEEE International Symposium on Circuits and Systems (ISCAS)*, Seoul, South Korea, 2012. IEEE, 2012, Pp. 850–853. DOI: 10.1109/ISCAS.2012.6272175
7. Huijsing J. *Operational amplifiers theory and design*, 2nd ed., Springer Science+Business, Delft, 2011, Pp. 351–412.

8. **Ivanov V., Shaik M.** A 5.1 A 10 MHz-bandwidth 4 μ s-large-signal-settling 6.5 nV/ $\sqrt{\text{Hz}}$ -noise 2 μ V-offset chopper operational amplifier. *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, San Francisco, USA, 2016. IEEE, 2016, Pp. 88–89. DOI: 10.1109/ISSCC.2016.7417920
9. **Kuang X., Wang T., Fan F.** The design of low noise chopper operational amplifier with inverter. *2015 IEEE 16th International Conference on Communication Technology (ICCT)*, Hangzhou, China, 2015. IEEE, 2016, Pp. 568–571. DOI: 10.1109/ICCT.2015.7399903
10. **Kusuda Y.** A 60 V auto-zero and chopper operational amplifier with 800 kHz interleaved clocks and input bias current trimming. *IEEE Journal of Solid-State Circuits*, 2015, Vol. 50, Issue 12, Pp. 2804–2813. DOI: 10.1109/JSSC.2015.2456891
11. **Mai T., Schmid K., Rober J., Hagelauer A., Weigel R.** A fully differential operational amplifier using a new chopping technique and low-voltage input devices. *2017 24th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, Batumi, Georgia, 2017. IEEE, 2018, Pp. 74–77. DOI: 10.1109/ICECS.2017.8292081
12. **Pipino A., Pezzotta A., Resta F., De Matteis M., Baschirotto A.** A rail-to-rail-input chopper instrumentation amplifier in 28nm CMOS. *2015 IEEE International Conference on Electronics, Circuits, and Systems (ICECS)*, Cairo, Egypt, 2015. IEEE, 2016, Pp. 73–76. DOI: 10.1109/ICECS.2015.7440252
13. **Prokop R.** Dynamic input offset auto-compensation of continuously working opamp. *2013 36th International Conference on Telecommunications and Signal Processing (TSP)*, Rome, Italy, 2013. IEEE, 2013, Pp. 440–443. DOI: 10.1109/TSP.2013.6613970
14. **Prokop R., Novak P., Musil V.** Bulk driven offset compensation for continuous time opamp operation. *Proceedings of the 15th International Scientific and Applied Science Conference*, Sofia, Bulgaria, 2006. TU-SOFIA, 2006, Pp. 26–29.
15. **Raghuvver V., Balasubramanian K., Sudhakar S.** A 2 μ V low offset, 130 dB high gain continuous auto-zero operational amplifier. *2017 International Conference on Communication and Signal Processing (ICCS)*, Chennai, India, 2017. IEEE, 2018, Pp. 1715–1718. DOI: 10.1109/ICCS.2017.8286685
16. **Shukla G., Srivastava N., Shadab A.** DC offset voltage reduction in cascaded instrumentation amplifier. *2013 Students Conference on Engineering and Systems (SCES)*, Allahabad, India, 2013. IEEE, 2013, Pp. 1–5. DOI: 10.1109/SCES.2013.6547579
17. **Singh R., Audet Y., Gagnon Y., Savaria Y., Boulais É., Meunier M.** A laser-trimmed rail-to-rail precision CMOS operational amplifier. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2011, Vol. 58, No. 2, Pp. 75–79. DOI: 10.1109/TCSII.2010.2104011
18. **Wu R., Huijsin J.H., Makinwa K.A.A.** *Precision instrumentation amplifiers and read-out integrated circuits*. New York: Springer, 2013, Pp. 21–48.
19. **Xu J., Nguyen A.T., Luu D.K., Drealan M., Yang Z.** Noise optimization techniques for switched-capacitor based neural interfaces. *IEEE Transactions on Biomedical Circuits and Systems*, 2020, Vol. 14, No. 5, Pp. 1024–1035. DOI: 10.1109/TBCAS.2020.3016738
20. **Yang X., Yang J., Lin L., Ling C.** Low-power low-noise CMOS chopper amplifier. *2010 International Conference on Anti-Counterfeiting, Security and Identification*, Chengdu, China, 2010. IEEE, 2010, Pp. 83–84. DOI: 10.1109/ICASID.2010.5551831
21. **Yeh C., Huang J., Wu P., Tsai H., Juang Y.** A low power and low noise CMOS chopper amplifier for use in capacitive type accelerometer. *2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, Jeju, South Korea, 2016. IEEE, 2017, Pp. 642–645. DOI: 10.1109/APCCAS.2016.7804072
22. **Yong X., Fei Z., Zheng S., Yuanliang W.** Design of novel chopper stabilized rail-to-rail operational amplifier. *2015 IEEE 11th International Conference on ASIC (ASICON)*, Chengdu, China, 2015. IEEE, 2016, Pp. 1–4. DOI: 10.1109/ASICON.2015.7516988
23. **Zhong X., Bermak A., Tsui C.** A low-offset dynamic comparator with area-efficient and low-power offset cancellation. *2017 IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC)*, Abu Dhabi, United Arab Emirates, 2017. IEEE, 2017, Pp. 1–6. DOI: 10.1109/VLSI-SoC.2017.8203481

24. **Zhou Y., Zhao M., Dong Y., Wu X., Tang L.** A low-power low-noise biomedical instrumentation amplifier using novel ripple-reduction technique. *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. Cleveland, USA, 2018. IEEE, 2018. Pp. 1–4. DOI: 10.1109/BIOCAS.2018.8584744

Received 26.08.2021.

СПИСОК ЛИТЕРАТУРЫ

1. **Balashov E.V., Ivanov N.V., Korotkov A.S.** SOI instrumentation amplifier for high-temperature applications // 2020 IEEE East-West Design and Test Symp. Varna, Bulgaria, 2020. IEEE, 2020. Pp. 1–4. DOI: 10.1109/EWDTS50664.2020.9224893
2. **Балашов Е.В., Иванов Н.В., Ахметов Д.Б., Коротков А.С.** Высокотемпературный инструментальный усилитель // Наноиндустрия. 2020. Вып. 96 (3). С. 160–163.
3. **Baker R.J.** CMOS: circuit design, layout, and simulation. 3rd ed. IEEE Press, 2010. Pp. 773–908.
4. **Bortun N., Stan M.N., Brezeanu G.** High precision bidirectional chopper amplifier with extended common mode input voltage range // Internat. Semiconductor Conf. (CAS). Sinaia, Romania, 2015. IEEE, 2015. Pp. 297–300. DOI: 10.1109/SMICND.2015.7355237
5. **Carusone T.C., Johns D.A., Martin K.W.** Analog integrated circuit design. 2nd ed. Wiley, 2011. Pp. 242–293.
6. **Duwe M., Chen T.** Offset correction of low power, high precision op amp using digital assist for biomedical applications // 2012 IEEE Internat. Symp. on Circuits and Systems. Seoul, South Korea, 2012. IEEE, 2012. Pp. 850–853. DOI: 10.1109/ISCAS.2012.6272175
7. **Huijsing J.** Operational amplifiers theory and design. 2nd ed. Delft: Springer Science+Business, 2011. Pp. 351–412.
8. **Ivanov V., Shaik M.** A 5.1 A 10 MHz-bandwidth 4 μ s-large-signal-settling 6.5 nV/ $\sqrt{\text{Hz}}$ -noise 2 μ V-offset chopper operational amplifier // 2016 IEEE Internat. Solid-State Circuits Conf. San Francisco, USA, 2016. IEEE, 2016. Pp. 88–89. DOI: 10.1109/ISSCC.2016.7417920
9. **Kuang X., Wang T., Fan F.** The design of low noise chopper operational amplifier with inverter // 2015 IEEE 16th Internat. Conf. on Communication Technology. Hangzhou, China, 2015. IEEE, 2016. Pp. 568–571. DOI: 10.1109/ICCT.2015.7399903
10. **Kusuda Y.** A 60 V auto-zero and chopper operational amplifier with 800 kHz interleaved clocks and input bias current trimming // IEEE J. of Solid-State Circuits. 2015. Vol. 50. Iss. 12. Pp. 2804–2813. DOI: 10.1109/JSSC.2015.2456891
11. **Mai T., Schmid K., Rober J., Hagelauer A., Weigel R.** A fully differential operational amplifier using a new chopping technique and low-voltage input devices // 2017 24th IEEE Internat. Conf. on Electronics, Circuits and Systems. Batumi, Georgia, 2017. IEEE, 2018. Pp. 74–77. DOI: 10.1109/ICECS.2017.8292081
12. **Pipino A., Pezzotta A., Resta F., De Matteis M., Baschiroto A.** A rail-to-rail-input chopper instrumentation amplifier in 28nm CMOS // 2015 IEEE Internat. Conf. on Electronics, Circuits, and Systems. Cairo, Egypt, 2015. IEEE, 2016. Pp. 73–76. DOI: 10.1109/ICECS.2015.7440252
13. **Prokop R.** Dynamic input offset auto-compensation of continuously working opamp // 2013 36th Internat. Conf. on Telecommunications and Signal Processing. Rome, Italy, 2013. IEEE, 2013. Pp. 440–443. DOI: 10.1109/TSP.2013.6613970
14. **Prokop R., Novak P., Musil V.** Bulk driven offset compensation for continuous time opamp operation // Proc. of the 15th Internat. Scientific and Applied Science Conf. Sofia, Bulgaria, 2006. TU-SOFIA, 2006. Pp. 26–29.
15. **Raghuveer V., Balasubramanian K., Sudhakar S.** A 2 μ V low offset, 130 dB high gain continuous auto-zero operational amplifier // 2017 Internat. Conf. on Communication and Signal Processing. Chennai, India, 2017. IEEE, 2018. Pp. 1715–1718. DOI: 10.1109/ICCSP.2017.8286685

16. **Shukla G., Srivastava N., Shadab A.** DC offset voltage reduction in cascaded instrumentation amplifier // 2013 Students Conf. on Engineering and Systems. Allahabad, India, 2013. IEEE, 2013. Pp. 1–5. DOI: 10.1109/SCES.2013.6547579
17. **Singh R., Audet Y., Gagnon Y., Savaria Y., Boulais É., Meunier M.** A laser-trimmed rail-to-rail precision CMOS operational amplifier // IEEE Transactions on Circuits and Systems II: Express Briefs. 2011. Vol. 58. No. 2. Pp. 75–79. DOI: 10.1109/TCSII.2010.2104011
18. **Wu R., Huijsin J.H., Makinwa K.A.A.** Precision instrumentation amplifiers and read-out integrated circuits. New York: Springer, 2013. Pp. 21–48.
19. **Xu J., Nguyen A.T., Luu D.K., Drealan M., Yang Z.** Noise optimization techniques for switched-capacitor based neural interfaces // IEEE Transactions on Biomedical Circuits and Systems. 2020. Vol. 14. No. 5. Pp. 1024–1035. DOI: 10.1109/TBCAS.2020.3016738
20. **Yang X., Yang J., Lin L., Ling C.** Low-power low-noise CMOS chopper amplifier // 2010 Internat. Conf. on Anti-Counterfeiting, Security and Identification. Chengdu, China, 2010. IEEE, 2010. Pp. 83–84. DOI: 10.1109/ICASID.2010.5551831
21. **Yeh C., Huang J., Wu P., Tsai H., Juang Y.** A low power and low noise CMOS chopper amplifier for use in capacitive type accelerometer // 2016 IEEE Asia Pacific Conf. on Circuits and Systems. Jeju, South Korea, 2016. IEEE, 2017. Pp. 642–645. DOI: 10.1109/APCCAS.2016.7804072
22. **Yong X., Fei Z., Zheng S., Yuanliang W.** Design of novel chopper stabilized rail-to-rail operational amplifier // 2015 IEEE 11th Internat. Conf. on ASIC. Chengdu, China, 2015. IEEE, 2016. Pp. 1–4. DOI: 10.1109/ASICON.2015.7516988
23. **Zhong X., Bermak A., Tsui C.** A low-offset dynamic comparator with area-efficient and low-power offset cancellation // 2017 IFIP/IEEE Internat. Conf. on Very Large Scale Integration. Abu Dhabi, United Arab Emirates, 2017. IEEE, 2017. Pp. 1–6. DOI: 10.1109/VLSI-SoC.2017.8203481
24. **Zhou Y., Zhao M., Dong Y., Wu X., Tang L.** A low-power low-noise biomedical instrumentation amplifier using novel ripple-reduction technique // 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS). Cleveland, USA, 2018. IEEE, 2018. Pp. 1–4. DOI: 10.1109/BIOCAS.2018.8584744

Статья поступила в редакцию 26.08.2021.

THE AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Assim Ara Abdulsatar
Ассим Ара Абдулсатар
E-mail: araabdulsattar@gmail.com

Balashov Evgenii V.
Балашов Евгений Владимирович
E-mail: balashov_ev@mail.ru

© Санкт-Петербургский политехнический университет Петра Великого, 2021

ИНФОРМАТИКА, ТЕЛЕКОММУНИКАЦИИ И УПРАВЛЕНИЕ

COMPUTING, TELECOMMUNICATIONS AND CONTROL

Том 14, № 3, 2021

Учредитель — Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Санкт-Петербургский государственный политехнический университет»

Журнал зарегистрирован Федеральной службой по надзору в сфере информационных технологий и массовых коммуникаций (Роскомнадзор).
Свидетельство о регистрации ЭЛ № ФС77-77378 от 25.12.2019 г.

Редакция журнала

д-р техн. наук, профессор *А.С. Коротков* — главный редактор
Е.А. Калинина — литературный редактор, корректор
Г.А. Пышкина — ответственный секретарь, выпускающий редактор

Телефон редакции (812)552-62-16

E-mail: infocom@spbstu.ru

Компьютерная верстка *А.А. Кононова*

Перевод на английский язык *Д.Ю. Алексеева*

Лицензия ЛР № 020593 от 07.08.97

Дата выхода 30.09.2021. Формат 60×84 1/8

Санкт-Петербургский политехнический университет Петра Великого
Адрес университета и редакции: 195251, Санкт-Петербург, ул. Политехническая, д. 29.