



МЕНЕДЖМЕНТ ОБРАЗОВАТЕЛЬНЫХ ОРГАНИЗАЦИЙ В ИНФОРМАЦИОННУЮ ЭПОХУ

MANAGEMENT OF EDUCATIONAL INSTITUTIONS IN THE INFORMATION ERA

DOI: 10.22363/2312-8631-2023-20-2-150-158

EDN: KEROFN

UDC 373

Research article / Научная статья

Learning analytics in Russia and abroad: level of development, trends and prospects

Tatiana A. Kustitskaya  , Mikhail V. Noskov 

Siberian Federal University, Krasnoyarsk, Russian Federation

 tkustitskaya@sfu-kras.ru

Abstract. *Problem statement.* Learning analytics is an emerging scientific field, which studies learners and learning process based on data from digital environment. The aim of the study – to observe the development of learning analytics, its prospects and limitations and detecting the state of art of this scientific field in Russia. *Methodology.* The study is based on context analysis of scientific articles on the topic in the public domain. Special attention is given to reviewing scientific publications of Russian-speaking authors devoted to analytics of education data and the implementation of learning analytics tools in the educational process. *Results.* The research detects the global directions of learning analytics development and its problematic aspects. It provides the quantitative and qualitative analysis of scientific publications of Russian-speaking authors and identifies the most popular research questions in the learning analytics field. It proposes the author’s vision of the hierarchy of directions for learning analytics development, consisting of the research aspect, the environment transformation aspect and the legal regulation aspect. The national initiatives in the digitalization of education are briefly discussed. *Conclusion.* A certain lag in the level of development of learning analytics in Russia from the global one is revealed. At the same time, there is a noticeable increase in interest to this area among individual researchers, educational institutions and at the state level, which allows us to count on positive changes.

Keywords: educational data mining, digitalization of education, data-driven learning management

Acknowledgements and Funding. This work is supported by the Russian Science Foundation under grant No 22-28-00413 “Digital academic history and its application in early warning systems for education”.

Article history: received 3 November 2022; revised 5 December 2022; accepted 12 January 2023.

© Kustitskaya T.A., Noskov M.V., 2023



This work is licensed under a Creative Commons Attribution 4.0 International License
<https://creativecommons.org/licenses/by-nc/4.0/legalcode>

For citation: Kustitskaya TA, Noskov MV. Learning analytics in Russia and abroad: level of development, trends and prospects. *RUDN Journal of Informatization in Education*. 2023;20(2):150–158. <http://doi.org/10.22363/2312-8631-2023-20-2-150-158>

Учебная аналитика в России и мире: уровень развития, основные тренды и перспективы

Т.А. Кустицкая  , М.В. Носков 

Сибирский федеральный университет, Красноярск, Российская Федерация

 tkustitskaya@sfu-kras.ru

Аннотация. *Постановка проблемы.* Учебная аналитика – относительно новое научное направление, исследующее учебный процесс и самих обучающихся по данным из цифровой образовательной среды. Цель исследования – определить текущий уровень, перспективы и ограничения развития аналитики обучения, выявить состояние этой научной области в России. *Методология.* Проведен контент-анализ научных публикаций по направлению исследования, находящихся в открытом доступе. Особое внимание уделено русскоязычным публикациям, посвященным анализу образовательных данных и внедрению инструментов учебной аналитики в образовательный процесс. *Результаты.* Определены и описаны общемировые тенденции и проблемные аспекты развития учебной аналитики. Представлены результаты количественного и качественного анализа научных публикаций русскоязычных авторов, определены наиболее популярные исследовательские вопросы в данной сфере. Предложено авторское видение направлений развития учебной аналитики, включающее научно-исследовательский аспект, аспекты преобразования образовательной среды и правового регулирования. Рассмотрены национальные инициативы в сфере цифровизации образования. *Заключение.* Выявлено некоторое отставание уровня развития учебной аналитики в России от общемирового. Вместе с тем наблюдается заметный рост интереса к этой сфере среди отдельных исследователей, образовательных учреждений и на государственном уровне, что позволяет рассчитывать на позитивные изменения.

Ключевые слова: анализ образовательных данных, цифровизация образования, управление обучением, данные

Благодарности и финансирование. Исследование выполнено при финансовой поддержке гранта Российского научного фонда, проект № 22-28-00413 «Цифровая образовательная история и ее применение в системах поддержки успешности обучения».

История статьи: поступила в редакцию 3 ноября 2022 г.; доработана после рецензирования 5 декабря 2022 г.; принята к публикации 12 января 2023 г.

Для цитирования: *Kustitskaya T.A., Noskov M.V. Learning analytics in Russia and abroad: level of development, trends and prospects // Вестник Российского университета дружбы народов. Серия: Информатизация образования. 2023. Т. 20. № 2. С. 150–158. <http://doi.org/10.22363/2312-8631-2023-20-2-150-158>*

Problem statement. Learning analytics is a relatively new multi-disciplinary field of science located at the intersection of pedagogy, computer science and psychology. It is “the measurement, collection, analysis, and reporting of data about learners and their contexts, for the purposes of understanding and optimizing

learning and the environments in which it occurs” [1]. In literature, one can find the terms Educational Data Mining (EDM) and Learning Analytics (LA). Although these branches differ in the methods used, they have the same object of study and the same ultimate goal – to improve the quality of education. Therefore, in the context of this work, we will consider EDM and LA synonymous.

Due to the uneven distribution of the level of digitalization of world economies, the level of IT technologies, the peculiarities of national policies in the field of education, the current state of art in the field of learning analytics differs.

Our study aims at reporting about the current level of LA development and implementation of its tools to learning process, detecting global trends and challenges. We pay special attention to the level of scientific discourse on LA in Russia and national initiatives that can be drivers of LA enhancement.

Methodology. The paper presents the results of context analysis of scientific articles and research reports on the world’s best practices of LA and EDM. As learning analytics is an in multidisciplinary field we analyzed several aspects of its development – methodological, technological and application-oriented. Leading positions, challenges.

To assess the level of scientific discourse on learning analytics in Russia we reviewed the publications of Russian-speaking authors on the topic. For this purpose, on the electronic platform eLIBRARY.ru, we selected publications in Russian (as well as in English with Russian-language annotations) using the keywords related to the field of study, and performed a quantitative analysis of the dynamics of research interest to LA and EDM. The qualitative analysis the open sources was made to detect main directions of academic interest and find successful cases of LA applications to educational process.

Results and discussion. *Global trends and challenges.* The formation of learning analytics as an independent scientific discipline is associated with the digitalization of education and the accumulation of large volumes of data on the educational process and students. EDM and LA emerged in the 2000s: in 2006 appeared the first monograph on EDM (C. Romero and S. Ventura), in 2008 and 2011 were held the first conferences on EDM and LA.

To date, at least three specialized journals are being published – the International Journal of Learning Analytics and Artificial Intelligence for Education, the Journal of Educational Data Mining (indexed in Scopus) and the Journal of Learning Analytics (in Web of Science and Scopus Q1). There are several large actively working analyst communities on the field – the most known are the Society for learning analytics research (SoLAR) and the International EDM society.

Educational institutes of the USA are among the most active users of LA tools: they use educational dashboards (for example, the student monitoring panel of Nottingham Trent University, “My Learning Analytics” dashboard in the University of Michigan), early warning systems preventing students’ dropouts (Course Signals system of Purdue University). Recommendation systems based on educational data analysis are implemented at an institutional level (Degree Compass system at Austin Peay University). Universities actively implement data-driven adaptive learning systems (Arizona State University places students into computerized learning courses running on the software of adaptive-learning start-up Knewton).

In terms of publication activity, the leading positions are occupied by the United States, Spain, the United Kingdom, Australia, Germany, Canada, India, the Netherlands, Japan, and China [2]. A significant increase in research and publication activity on the LA topics can be expected in European countries due to the development of national policies, the launch of joint projects on a European level and the creation of competence centers [3].

Initially, the most intensive development was observed in such areas of learning analytics as predictive analysis of learning outcomes and monitoring students' learning progress. These topics are still trending and in several countries, their development is supported by national policies. For example, in the United Kingdom, higher educational institutes are required to produce detailed metrics for the government-established Office for Students [4].

There are numerous research articles devoted to the application of machine learning methods to forecasting student success, learning outcomes or dropouts [5]. Learning Analytics Dashboards become more and more widespread especially at Higher Educational institutions.

At the same time, new important directions for the development of LA are gaining importance, such as data-driven analysis and support of self-regulated learning, provision of personalized feedback and support for learners, development of recommendation and early warning systems. The effectiveness of the implementation of the mentioned LA instruments in a real educational process has recently become the main challenge in LA.

As it is mentioned in [5] although data-driven optimization of the learning process is the final stage of the learning analytics cycle, most of the research is still devoted to developing predictive models rather than to designing LA interventions.

There are several possible reasons why the data-driven interventions into the learning process are still not widespread, which are related both to students and course instructors. On the one hand, if the interventions are aimed at supporting at-risk students they can face the problem of weak metacognitive skills and self-regulation of such students [6]. On the other hand, implementation of supportive interventions means an additional workload for course instructors and teachers, which is usually not regulated.

Another challenge is due to the dual nature of LA. As methods and tools of LA come from IT and data science LA instruments are usually designed by IT specialists not much familiar with pedagogical and psychological aspects of the learning process and peculiarities of the studied disciplines. This results in unjustified simplification when modelling such objects as “proper learning behaviour”, “learning style”, “student online interactions” and even “learning performance”. In [7] authors argue that to be effective LA instruments should be grounded in robust and clearly articulated theories of learning. Unfortunately, educators who are competent in the disciplines they teach and more or less familiar with learning theories usually have difficulty developing data analysis tools on their own.

However, the technological aspect of LA itself is also problematic. Big Data techniques well-proven for other fields cannot be automatically applied to an educational context. For instance, in physical sciences data is collected in highly controlled, carefully designed and replicable experiments, while educational data is collected incidentally and in circumstances that differ for every single user-

resource or user-user interaction [7]. The replicability of such experiments is not only doubtful but also contradicts the overall purpose of LA which is to optimize the learning environment. The data mining technique from other fields should be carefully adapted to LA taking into account both the subjectiveness of education data and goals of data collection and analysis.

Learning analytics in Russian Federation. In Russia, data-driven educational analytics have started to develop actively in the last decade due to the later spread of e-learning methods, the introduction of LMS in the learning process of educational institutions, and the emergence of massive online courses.

To assess the level of LA development and implementation of LA instruments in Russia, we highlight the following aspects of this process:

The research aspect that concerns the development of new or adaptation of already existing methods for educational data collection and analysis. The research topics include identifying predictors of learning success, building student models, designing learners digital twins, creating and testing methods for managing the educational process, designing supportive data-driven interventions, monitoring learners track and performance in e-learning courses, etc.

The environment transformation aspect, which means digitalization of the learning process in the educational institutions, the introduction of information systems for collecting educational data with in-built analytical tools at the level of universities and above, organizing access for all stakeholders of the educational process to monitoring data and the results of their tracks' analysis, the transformation of institutional policies to ensure the possibility of automatizing some processes, conducting targeted corrective measures based on the analysis of the student's educational footprint, etc.

The legal regulation aspect that implies modernization of the state policy in the field of education: changing educational standards towards increasing its flexibility, developing regulatory norms for conducting analytics and making data-driven decisions in the educational field.

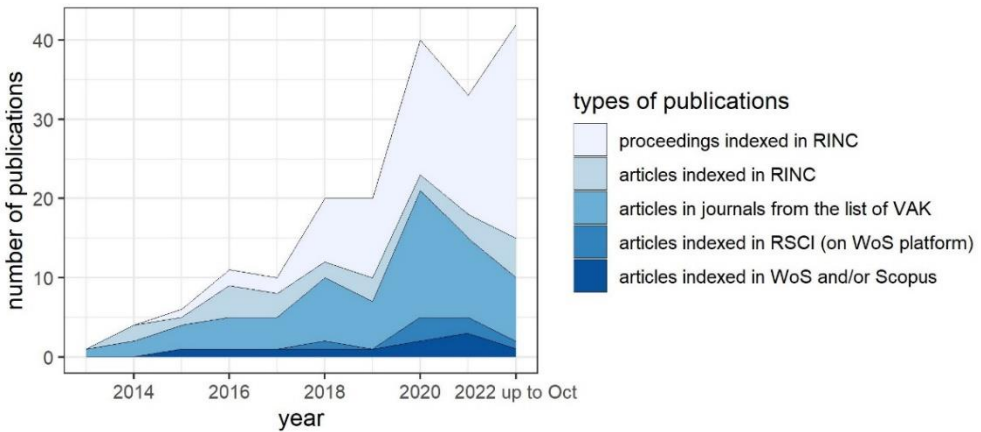
Although development can proceed simultaneously in all directions, we believe that any administrative and state initiatives can be successful only if academics understand the goals and methods of transformations, and if empirical research is carried out and actively discussed.

The level of scientific discourse on learning analytics in Russia. On the electronic platform eLIBRARY.ru, we selected publications in Russian (or with Russian-language annotations) using the keywords (in Russian): “educational analytics”, “learning analytics”, “analysis of educational data”, “educational data mining”, “methods (methodology) for educational data analysis”, “methods for educational data mining”, “methodology and technology for educational analytics”, “Big Data in education”, “learning success prediction”, “learning success modelling”, “student digital footprint”.

Then duplications and irrelevant results were removed, and the remaining 187 publications from the resulting collection were divided into five categories: 1) articles indexed in WoS and/or Scopus; 2) articles indexed in RSCI (Russian Science Citation Index – on the Web of Science Platform) not included in category 1; 3) articles in journals included in the list of the Higher Attestation Commission (VAK) not included categories 1 and 2; 4) articles in journals indexed in

the RINC (Russian Index of Scientific Citation), not included in categories 1, 2, and 3; 5) proceedings of conferences, indexed in the RINC.

The first publication on the LA in Russian appeared in 2013, then the number of publications increased and in 2020 it resulted already in 40 publications per year (Figure). In the period January – October 2022, 42 articles have already been published. We think that a significant increase in the number of publications in 2020–2022 is due to the acute problem of monitoring the educational process in the period of COVID-19 restrictions.



Dynamics of the number of publications on LA on eLIBRARY.ru from 2013 to October 2022

Source: compiled by the authors.

After reviewing the content of the open-source publications, we found that most of the papers are just announcing the possibility of using learning analytics in education, providing basic information about its goals, objectives and tools. For an initial acquaintance with the essence of educational analytics, the article [8] seems to be the most useful for a novice researcher. The most complete overview of foreign experience in the use of LA tools can be found in [9; 10].

The most popular among Russian researchers LA problem is the prediction of learning success. In our list we found 14 papers, describing cases on solving this problem by means of various methods of statistical analysis and machine learning. The authors predict academic performance at the level of an academic course or MOOC (Massive Open Online Course), the success of university graduation, and the level of competencies development. At the same time, in 8 works, researchers provide full descriptions of empirical studies and determine the most effective predictive methods. We consider the papers [11; 12] to be the most informative on the topic.

Several studies are devoted to identifying patterns in the learning behaviour of students and describing their learning strategies with a further purpose to use the results for designing individual learning trajectories. In this case, various clustering algorithms are used and compared, as, for example, in [13].

Only in 6 publications authors describe learning management tools built on the basis of already designed and implemented in the educational process LA instruments. The described tools are used for data collection in the e-learning environment and for analysis of the collected data. These tools are implemented either at the level of individual academic courses [14] or at the university level [15].

Most of the articles describe, discuss, or explore various aspects of educational data mining as applied to higher education. Only 4 publications from our list are devoted to studies of educational data in school (for example, in [16]). Five publications deal with MOOC learning analytics (for example, in [17]).

Russian-speaking researchers also pay attention to the methodological and algorithmic basis of educational analytics. They introduce new quality metrics, approaches to the adaption of data analyses and modelling methods to the analysis of the educational process and learners. For instance, in [18] authors propose a methodology for the automated assessment of programming tasks in an electronic environment that is based on statistical analysis. In [19] various probabilistic methods are used for student success modelling, as well as for learning activities assessment.

Transformation of educational environments and national policy. Digitalization of education has been declared to be one of the strategic tasks of the Russian Federation as an aspect of the digitalization of economics.

Recent years there appeared several national initiatives aimed at development and establishment of regulatory rules and guidelines for digitalization and usage data for educational research. One of such initiatives (University-2035) announces developing approaches to data-driven educational management as one of its priority tasks. One of the important products of University-2035 in the LA field is a Digital Footprint Standard, which covers the rules for sampling, processing and use of the digital footprint in education.¹

In 2019 there was launched the Consortium “Evidence-based digitalization for student success” which states providing analytical support for digitalization for making informed and effective management decisions based on research data as its main goal. About 20 universities have joined the consortium at the moment. One of the Consortium projects, Learning Analytics Research Community, unites academics, education centers’ managers and independent researchers at a platform for communication, sharing educational data and LA-algorithms.

Hopefully, the mentioned initiatives will result in significant modernization of policies or in national level research projects.

Conclusion. The importance of learning analytics is constantly growing together with digitalization of education, providing educators promising prospects of the data-driven learning process optimization. However, there are several problematic aspects of its development – pedagogical, technological and ethical.

Learning analytics in Russia is gaining popularity. There are some examples of successful implementation of some LA tools in the educational process, but in general, the scientific and pedagogical community in Russia is just getting acquainted with this field. Nevertheless, there are several encouraging signals coming both from individual researchers and at the state level. The number of Russian-language publications on the topic is rapidly increasing. Research and educational communities focused on the development of technology-enhanced education have a good chance of receiving funding and other support due to the concerns about the distance education effectiveness that are currently strong in both academic society and the authorities.

¹ Standard.2035.university. Moscow: ANO National Technology Initiative University. Available from: <https://standard.2035.university/>

References

- [1] Siemens S, Gašević D. Special issue on learning and knowledge analytics. *Educational Technology & Society*. 2012;15(3):1–163.
- [2] Tsai YS, Rates D, Moreno-Marcos PM, Muñoz-Merino PJ, Jivet I, Scheffel M, Gašević D. Learning analytics in European higher education – trends and barriers. *Computers & Education*. 2020;155. <http://doi.org/10.1016/j.compedu.2020.103933>
- [3] Nouri J, Ebner M, Ifenthaler D, Sqr M, Malmberg J, Khalil M, Berthelsen UD. Efforts in Europe for data-driven improvement of education – a review of learning analytics research in six countries. *International Journal of Learning Analytics and Artificial Intelligence for Education*. 2019;1(1). <http://doi.org/10.3991/ijai.v1i1.11053>
- [4] Olney T, Walker S, Wood C, Clarke A. Are we living in LA (P) LA Land? Reporting on the practice of 30 STEM tutors in their use of a learning analytics implementation at the open university. *Journal of Learning Analytics*. 2021;8(3):45–59. <http://doi.org/10.18608/jla.2021.7261>
- [5] Wong BT, Li KC. A review of learning analytics intervention in higher education (2011–2018). *Journal of Computers in Education*. 2020;7(1):7–28. <http://doi.org/10.1007/s40692-019-00143-7>
- [6] Wise AF. Designing pedagogical interventions to support student use of learning analytics. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge* (24–28 March 2014). Indianapolis; 2014. p. 203–211.
- [7] Wilson A, Watson C, Thompson TL, Drew V, Doyle S. Learning analytics: challenges and limitations. *Teaching in Higher Education*. 2017;22(8):991–1007. <http://doi.org/10.1080/13562517.2017.1332026>
- [8] Belonozhko PP, Karpenko AP, Khramov DA. Analysis of educational data: directions and prospects of application. *Bulletin of Eurasian Science*. 2017;9(4). (In Russ.) Available from: <http://naukovedenie.ru/PDF/15TVN417.pdf> (accessed: 20.10.2022).
Белоножко П.П., Карпенко А.П., Храмов Д.А. Анализ образовательных данных: направления и перспективы применения // Науковедение. 2017. Т. 9. № 4. URL: http://naukovedenie.ru/PDF/15TVN417.pdf (дата обращения: 20.10.2022).
- [9] Nazhmidinov KA. Analysis of the best practices of the use of learning analytics in the foreign universities. *Proceedings of the 12th International Scientific-Practical Conference SITU (25 February – 1 March 2019)*. Ekaterinburg; 2019. p. 172–178. (In Russ.)
Нажмидинов Х.А. Анализ лучших практик применения образовательной аналитики в зарубежных вузах // Наука. Информатизация. Технологии. Образование: материалы XII Международной научно-практической конференции, Екатеринбург, 25 февраля – 1 марта 2019 г. Екатеринбург, 2019. С. 172–178.
- [10] Vilkova KA, Zakharova US. Learning analytics in conventional education: its role and outcomes. *University Management: Practice and Analysis*. 2020;24(3):59–76. <http://doi.org/10.15826/umpa.2020.03.026>
- [11] Ozerova GP, Pavlenko GF. Prediction of student performance in blended learning utilizing learning analytics data. *Science for Education Today*. 2019;9(6):73–87. (In Russ). <http://doi.org/10.15293/2658-6762.1906.05>
Озерова Г.П., Павленко Г.Ф. Прогнозирование успешности студентов при смешанном обучении с использованием данных учебной аналитики // Science for Education Today. 2019. Т. 9. № 6. С. 73–87. http://doi.org/10.15293/2658-6762.1906.05
- [12] Bystrova T, Larionova V, Sinitsyn V, Tolmachev A. Learning analytics in massive open online courses as a tool for predicting learner performance. *Educational Studies*. 2018;4:139–166. <http://doi.org/10.21125/iceri.2018.1033>
- [13] Aldunin DA. Application of the adaptive content concept for an e-learning resource. *Business Informatics*. 2016;(4):27–34. <http://doi.org/10.17323/1998-0663.2016.4.27.34>
- [14] Kustitskaya TA, Karnaukhova OA. Developing an early warning system to detect at-risk students: a feedback mechanism. *Proceedings of the 5th International Conference*

on Informatization of Education and E-learning Methodology: Digital Technologies in Education (6–9 October 2020). Krasnoyarsk; 2020. p. 289–293.

- [15] Baranova EV, Shvetsov GV. Methods and tools for analysing students' digital footprint in the course of work under educational programmes. *Perspectives of Science and Education*. 2021;(2):415–430. (In Russ.) <http://doi.org/10.32744/pse.2021.2.29>
Баранова Е.В., Швецов Г.В. Методы и инструменты для анализа цифрового следа студента при освоении образовательного маршрута // *Перспективы науки и образования*. 2021. № 2 (50). С. 415–430. <http://doi.org/10.32744/pse.2021.2.29>
- [16] Alishev T, Gafarov F, Sabitova E. Academic performance and grade 9 milestone: primary data analysis in the system of "Electronic education in the Republic of Tatarstan". *Kazan Pedagogical Journal*. 2020;6:177–186.
- [17] Tolmachev AV, Sinitsyn EV, Astratova GV. Probability distributions of the academic performance of online courses's students as a tool for assessment of the quality of control materials. *Russian Journal of Resources, Conservation and Recycling*. 2020;7(3). <http://doi.org/10.15862/10INOR320>
- [18] Esin TE. Development of the metric of determination probabilistic distance to solution in difficult problem areas. *Modeling, Optimization and Information Technology*. 2021;9(1). (In Russ.) <http://doi.org/10.26102/2310-6018/2021.32.1.006>
Есин Т.Е. Разработка метрики определения вероятностного расстояния до решения в сложных проблемных областях // *Моделирование, оптимизация и информационные технологии*. 2021. Т. 9. № 1. <http://doi.org/10.26102/2310-6018/2021.32.1.006>
- [19] Ozerova GP. Design and assessment self-study tasks in blended learning environment. *Proceedings of Voronezh State University: Problems of Higher Education*. 2020;2:82–86. (In Russ.)
Озерова Г.П. Разработка и оценка заданий самостоятельной работы в рамках концепции смешанного обучения // *Вестник Воронежского государственного университета*. Серия: Проблемы высшего образования. 2020. № 2. С. 82–86.

Bio notes:

Tatiana A. Kustitskaya, Candidate of Physical and Mathematical Sciences, Associate Professor, Associate Professor of the Department of Applied Mathematics and Computer Security, Siberian Federal University, 79 Svobodnyi Prospekt, Krasnoyarsk, 660041, Russian Federation. ORCID: 0000-0001-9854-1259, eLIBRARY SPIN-code: 5202-8701. E-mail: tkustitskaya@sfu-kras.ru

Mikhail V. Noskov, Doctor of Physical and Mathematical Sciences, Professor, Professor of the Department of Applied Mathematics and Computer Security, Siberian Federal University, 79 Svobodnyi Prospekt, Krasnoyarsk, 660041, Russian Federation. ORCID: 0000-0002-4514-7925, eLIBRARY SPIN-code: 3957-7221. E-mail: MNoskov@sfu-kras.ru

Сведения об авторах:

Кустицкая Татьяна Алексеева, кандидат физико-математических наук, доцент, доцент кафедры прикладной математики и компьютерной безопасности, Сибирский федеральный университет, Российская Федерация, 660041, Красноярск, Свободный пр-кт, д. 79. ORCID: 0000-0001-9854-1259, eLIBRARY SPIN-код: 5202-8701. E-mail: tkustitskaya@sfu-kras.ru

Носков Михаил Валерианович, доктор физико-математических наук, профессор, профессор кафедры прикладной математики и компьютерной безопасности, Сибирский федеральный университет, Российская Федерация, 660041, Красноярск, Свободный пр-кт, д. 79. ORCID: 0000-0002-4514-7925, eLIBRARY SPIN-код: 3957-7221. E-mail: MNoskov@sfu-kras.ru