Philosophical foundations of technology and society

Мировоззренческие основы технологии и общества

UDC 378
https://doi.org/10.32362/2500-316X-2022-10-4-86-92

RESEARCH ARTICLE

Applying a reproducible research approach to distance education

Michail A. Eremeev,
Oleg V. Trubienko [@],
Ivan I. Zakharchuk

MIREA – Russian Technological University, Moscow, 119454 Russia
[@] Corresponding author, e-mail: trubienko@mirea.ru

Abstract

Objectives. Emerging as a response to the global threats presented by the COVID-19 pandemic**,** the changing nature of problem-solving in the field of information technology associated with economic globalization, including possibilities of remote working, imposes new requirements on the competencies and skills of future professionals. This, in turn, requires adjustments to the higher education process. Agile project management methodologies such as Scrum, along with Infrastructure-as-Code approaches in information and telecommunication infrastructure management, and Documentation-as-Code approaches in documentation development, aim to present design, development, testing, and documentation as short cycle iterative processes to permit the rapid and transparent addition of new product value in discrete portions. Applied to the education sphere, this approach implies new knowledge and practical skills of students that can be easily and transparently measured in the process of mastering a discipline. The present paper aims to develop methods of applying modern software development techniques to training students of technical specialties.

Methods. The use of reproducible research methods and agile design practices while organizing and managing practical tasks for students is proposed.

Results. Contemporary tools used in software development based on Git hosting services (GitLab and GitHub) are presented alongside reproducible research paradigms in distance education using the R Markdown format by RStudio.

Conclusions. In addition to increasing the involvement of students in the process of practical tasks, the proposed approach can be used to reduce the workload of teachers when checking and evaluating student working results.

Keywords: reproducible research, R Markdown in education, Git in education, Agile in education

• **Submitted:** 09.11.2021 • **Revised:** 15.05.2022 • **Accepted:** 27.06.2022

For citation: Eremeev M.A., Trubienko O.V., Zakharchuk I.I. Applying a reproducible research approach to distance education. *Russ. Technol. J.* 2022;10(4):86–92. https://doi.org/10.32362/2500-316X-2022-10-4-86-92

Financial disclosure: The authors have no a financial or property interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Применение подхода воспроизводимых исследований в процессе дистанционного обучения

М.А. Еремеев,
О.В. Трубиенко [@],
И.И. Захарчук

МИРЭА – Российский технологический университет, Москва, 119454 Россия
[@] Автор для переписки, e-mail: trubienko@mirea.ru

Резюме

Цели. Изменение характера решения задач в сфере информационных технологий, связанное с глобализацией экономики, возможностями удаленной работы, а также глобальными угрозами пандемии COVID-19, накладывает новые требования к компетенциям и навыкам будущих специалистов, что, в свою очередь, требует внесения корректив и в процесс подготовки специалистов в высших учебных заведениях. В настоящее время широко известны различные методики проектного управления Agile и Scrum (часто еще называемые «методологиями гибкой разработки»), подходы в управлении информационно-телекоммуникационной инфраструктурой Infrastructure-as-Code, а также подходы в разработке документации Documentation-as-Code. Их общая цель – представить процесс проектирования, разработки, тестирования и документирования в итеративном виде с короткими периодами цикла, позволяющего прозрачно добавлять новую ценность продукта небольшими порциями. В области образования этой ценностью являются новые знания, умения и практические навыки обучающихся, которые можно легко и прозрачно измерить в процессе освоения ими дисциплины. Цель исследования – разработать способы применения современных методик разработки программного обеспечения в процессе обучения студентов технических специальностей.

Методы. Использовались методы воспроизводимых исследований (reproducible research) и практик гибкого проектирования при организации и руководстве выполнением обучающимися практических работ.

Результаты. Представлен подход к использованию современных инструментов, применяемых при разработке программного обеспечения на базе онлайн-сервисов git-хостинга (GitLab и GitHub), а также парадигмы «воспроизводимых исследований» в процессе дистанционного обучения с использованием формата R Markdown компании RStudio.

Вывод. Применение предложенного подхода позволяет, помимо увеличения вовлеченности обучающихся в процесс выполнения практических заданий, снизить нагрузку на преподавателя по проверке и оцениванию результатов работы студентов.

Ключевые слова: воспроизводимые исследования, R Markdown в обучении, Git в обучении, Agile в обучении

• **Поступила:** 09.11.2021 • **Доработана:** 15.05.2022 • **Принята к опубликованию:** 27.06.2022

Для цитирования: Еремеев М.А., Трубиенко О.В., Захарчук И.И. Применение подхода воспроизводимых исследований в процессе дистанционного обучения. *Russ. Technol. J.* 2022;10(4):86–92. https://doi.org/10.32362/2500-316X-2022-10-4-86-92

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

Evolving information technologies have the possibility to transform conventional approaches to solving various tasks and problems arising in the course of human life. The sudden increase in risks to humanity associated with the COVID-19 pandemic stimulates the search for new approaches to improve the efficiency of activities including those in the education sphere. In addition to the convenience and flexibility of managing the learning process, the technical ability to support teacher–student interaction remotely through communication channels acquires a fundamentally new meaning.

A major trend in development and rationalization of modern education based on the use of electronic and mobile learning tools is the so-called collaborative learning approach, comprising the independent work of students in collaboration with other students with the participation of a teacher [1, 2]. The transition of Russian education system to a competence-based approach implies creating conditions for students to master a set of competencies to determine graduate readiness for sustainable professional activities [3–6]. The idea of applying the technologies used in software development to other spheres is not new. Various Agile project management methodologies[1] such as Scrum[2], along with Infrastructure-as-Code (IaC)[3] approaches in information and telecommunication infrastructure management, and Documentation-as-Code (Docs as Code)[4] approaches in documentation development are widely known. The use of proven tools, such as the Git[5] version control system, Wiki[6] and Markdown[7] markup languages, as well as task management systems such as Jira[8], Asana[9], and Trello[10], allows:

1) creation of a unified storage for the common knowledge product and provision of simultaneous access for all participants in the process of its creation;
2) the possibility of parallel work on competing solutions to the same problem and experiments that allow the possibility of reverting to any initial state;
3) the use of the knowledge and experience of all team members when making decisions, not just those directly involved in the project;
4) recording decisions for further retrospective analysis with a view to developing best practices in this area and disseminating them;
5) ensuring asynchronous interaction of the participants;
6) the possibility to evaluate the contribution of each participant to the result more objectively.

The present paper summarizes the experience of conducting practical and laboratory works using R Markdown interactive electronic report technologies [7] along with the Git version control system. Here, the main aim is to develop educational technologies that develop skills associated with the use of contemporary tools that support individual, independent learning paths at the same time as demonstrating the possibility of collective work in a unified information space. The possibility of combining these seemingly contradictory tendencies focuses exclusively on software products used in IT industry without the need to deploy special education software. By instilling engineering skills involved in building new information architectures, this eliminates the impression of "artificiality" in works performed by students, as well as increasing their commitment and activity. In addition to improving the quality of training, the use of standard technologies should increase the attractiveness of the educational institution for applicants [8].

In the aftermath of the COVID-19 pandemic, issues of organizing remote work and learning are becoming more relevant than ever before. However, the main focus has been on teleconferencing technologies (e.g., Zoom[11], Microsoft Teams[12], and Cisco Webex[13]), which preserve the traditional structure of teacher–student interaction. While preserving this important element, the approach proposed in the present paper also aims to facilitate the active interaction of students by introducing mechanisms of gamification and competition.

Of particular note here is the property of asynchronous interaction in the learning process, which allows all process participants to manage their individual time resources flexibly, including supporting interactions from anywhere in the world regardless of time zone. By encouraging students to schedule and organize their work independently, synchronous interaction removes formal restrictions concerning the timing and length of classes: only synchronization points (i.e., reporting deadlines) are required to be observed by university students.

[1] https://www.atlassian.com/ru/agile/manifesto. Accessed May 23, 2021 (in Russ.).
[2] https://www.scrum.org/. Accessed May 23, 2021.
[3] https://docs.microsoft.com/en-us/devops/deliver/what-is-infrastructure-as-code. Accessed May 23, 2021.
[4] https://openpracticelibrary.com/practice/docs-as-code/#:~:text=What%20Is%20Docs%20As%20Code,you%20can%20expect%20to%20see.&text=A%20culture%20of%20adaptation%20and,%2C%20and%20processes%2C%20over%20time. Accessed May 23, 2021.
[5] https://git-scm.com/. Accessed May 23, 2021.
[6] http://www.xwiki.org/xwiki/bin/view/Main/. Accessed May 23, 2021.
[7] https://daringfireball.net/projects/markdown/. Accessed May 23, 2021.
[8] https://www.atlassian.com/software/jira#. Accessed May 23, 2021.
[9] https://asana.com/. Accessed May 23, 2021 (in Russ.).
[10] https://trello.com/. Accessed May 23, 2021.
[11] https://zoom.us/. Accessed May 23, 2021.
[12] https://www.microsoft.com/ru-ru/microsoft-teams/group-chat-software. Accessed May 23, 2021 (in Russ.).
[13] https://www.webex.com/. Accessed May 23, 2021.

R MARKDOWN TECHNOLOGY FOR CREATING INTERACTIVE DIGITAL REPORTS

The product of contemporary scientific research increasingly involves the participation of entire teams, with new knowledge emerging from long chains of smaller studies based on each other's the results. For this reason, the costs of mistakes made at any particular stages can be significant. The problem of reproducibility of scientific research is not only acute in natural and engineering sciences, but also increasingly in the humanities.

The reproducibility of scientific research implies the possibility of repeating experiments, calculations, and carrying out modeling on basis of the same initial data used in the original experiment to obtain a similar result. To this end, all initial data should be made publicly available for studying along with a sufficiently detailed documentation of methods and materials used[14] [9]. In the case of modeling approaches, the source code of any software tools should be also presented. One of the most suitable options for realizing these requirements involves the use of R Markdown technology [10, 11].

By specifying the syntax of markup language, R Markdown presents a technology for processing specially generated documents, which include authoring text, formulas, and program code whose results are automatically inserted into a generated report. The results of executing such code may be tables, graphs, charts, diagrams, geographical maps, heat maps, etc.

As well as the automated generation of reports or articles, the advantages of this approach include minimization of error probability when transferring results to the report. In addition, since the program code is contained in the report itself, other specialists can check and reproduce repeatedly performed calculations. Moreover, the use of already written program code facilitates the reuse of the performed calculations, developed algorithms and methods in other in other works and spheres.

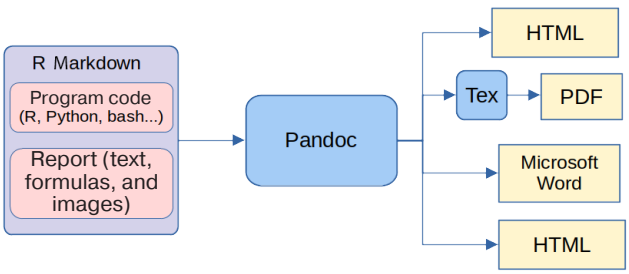

Figure. A scheme for the resulting document generation

[14] Zaitsev V.S. *Modern pedagogical technologies: textbook*. In 2 books. Chelyabinsk: ChSPU; 2012. 411 p. (in Russ.).

The report can be generated as a MS Word file, a MS PowerPoint presentation or Reveal.js, in TeX typesetting system format, in PDF format, or as an interactive Web application (see figure).

GIT VERSION CONTROL SYSTEM

Created in 2005 by Linus Torvalds, the founder of the Linux operating system, as a replacement for similar closed source software, Git was originally intended as the distributed version control system for software source code. Since then, it has quickly gained popularity due to its distributed version control structure and simple architecture. The following may be also mentioned among Git's main advantages:

1) support for a non-linear process of writing the source code of a software product with the possibility of branching, working on each branch in parallel, and merging;
2) use of existing widespread protocols (SSH, FTP, and HTTPS) to provide communication channels between repositories;
3) high working speed;
4) cryptographic protection of change history integrity.

With Git, the history of source code changes consists of sequential "checkpoints" (commits). This sequence branches out when working on different code versions in parallel and converges when the changes are merged.

Today, there are many online source code hosting services, GitHub.com and GitLab.com being the best known. They are web applications allowing not only setting up Git code repositories but also performing a wide range of design, code writing, testing, documentation, integration, and software delivery tasks.

ORGANIZING THE EDUCATIONAL PROCESS USING R MARKDOWN AND GIT

When combining these technologies, the following main objectives are pursued:

1) mastering student teamwork and interaction skills at all stages of project implementation, from planning to report writing;
2) familiarity with modern technologies as well as analytical and software development tools [12];
3) increase motivation of students for high-quality performance of tasks due to their interest in the possible reuse of developments, including in other disciplines [13];
4) controlling the performance history of the assigned tasks and evaluating the contribution of each participant;
5) facilitating validation of presented calculations and excluding "fitting" data to the desired result, which stimulates students to study the topic deeper and more detailed [14].

During training sessions, trainees are divided into teams (crews) of 4–6 people. Each team creates its own central repository (using GitHub or GitLab services), in which participants can organize their work on a task or project. The repository is added by the report template and initial data. Team members may add any materials related to the assigned task on their own: references, links to literature and Internet resources, schematics, etc. Thus, the repository becomes a logical focal point for student interaction, helping them to acquire and strengthen remote interaction skills using modern information technologies.

The working results of each team are used to write the R Markdown report in RMD format for subsequent output in HTML or PDF format. The main requirement for the resulting report is its reproducibility, i.e., the ability to compile an identical report document on the teacher's PC using the input data. During this process, an automatic validation of the methods and tools used by students is carried out. According to the commits (checkpoints) in the group repository, the teacher may evaluate the efficiency of each group member along with his or her contribution to the resulting product. Each team member must participate in the task and report-writing process by creating at least one meaningful commit. This encourages participants to get involved as early as possible doing the most obvious work on the assignment. Conversely, low activity in the course of doing the work forces participants to supplement and improve the already posted working results: to comment and design the program code, study and describe alternatives for the task, and improve the design of charts, i.e., perform work requiring more detailed and in-depth study of material.

Among the advantages of this approach to training, it is worth noting asynchronous interaction of the teacher with students and students with each other. This allows each student to organize and schedule study time individually, while maintaining process interactivity and a sense of involvement in the common work. Thus, an individual educational path may be built while preserving the main "reference points" of curriculum subjects.

Combining organizational tools for distributed work with those for documenting and "packaging" knowledge in a reproducible format helps to increase the efficiency of learning and mastering competencies while maintaining flexibility in managing the student time management.

CONCLUSIONS

By organizing practical and laboratory works according to the approach proposed in the paper, a number of advantages may be realized as compared to the conventional approach of dividing students into teams. Firstly, the overall activity of students increases; they get interested in using contemporary means of management, time-tracking, as well as management of documents and program code. Secondly, the work yields a reproducible result that excludes the possibility of fitting the solution to the desired result or searching for a ready-made solution due to the need to perform all search and decision-making stages; thus, the overall credibility of the system for evaluating student success is greatly enhanced. Thirdly, evaluating student works is facilitated for the teacher, as well as presenting wide opportunities for the automated evaluation of the learning process and creation of its results. This trend appears a very promising direction for further research.

Authors' contributions

M.A. Eremeev—research of the distance learning processes.

O.V. Trubienko—study of the possibility of using software tools to organize the distance learning process.

I.I. Zakharchuk—research idea, the development of aims and objectives, the formulation of conclusions.

REFERENCES

1. Altukhov A.I., Bagretsov S.A., Karpinchuk N.A., Cheburkov M.A. Methodology of assessment of time spending on learning a course in the academic discipline with the use of automated training systems. *Izvestiya SPbSTU LETI = Proceedings of Saint Petersburg Electrotechnical University*. 2016;7:32–37 (in Russ.).
2. Vasil'eva N.V., Kunturova N.B., Prokof'eva A.L. Educational tools of information and communication technologies in the educational process of the military institution. *Trudy Voenno-kosmicheskoi akademii im. A.F. Mozhaiskogo = Proceedings of the Mozhaisky Military Aerospace Academy*. 2018;661:207–214 (in Russ.).

СПИСОК ЛИТЕРАТУРЫ

1. Алтухов А.И., Багрецов С.А., Карпинчук Н.А., Чебурков М.А. Методика оценивания временных затрат на изучение курса учебной дисциплины с применением автоматизированных обучающих систем. *Известия СПбГТУ ЛЭТИ*. 2016;7:32–37.
2. Васильева Н.В., Кунтурова Н.Б., Прокофьева А.Л. Образовательные средства информационно-коммуникационных технологий в учебном процессе военного ВУЗа. *Труды Военно-космической академии им. А.Ф. Можайского*. 2018;661:207–214.
3. Калинин В.Н. Основные направления формирования и критерии оценивания общекультурных и профессиональных компетенций в цикле математических и естественнонаучных дисциплин. *Информационный бюллетень Военно-космической академии им. А.Ф. Можайского*. 2014;121:19–27.

3. Kalinin V.N. The main directions of formation and criteria for assessing general cultural and professional competencies in the cycle of mathematical and natural science disciplines. *Informatsionnyi byulleten' Voenno-kosmicheskoi akademii im. A.F. Mozhaiskogo = Information Bulletin of the Mozhaisky Military Aerospace Academy*. 2014;121:19–27 (in Russ.).
4. Altukhov A.I., Golovina V.V., Kalinin V.N. Forming and criterions of competence estimation in subjects of mathematical and naturally scientific cycle. *Trudy Voenno-kosmicheskoi akademii im. A.F. Mozhaiskogo = Proceedings of the Mozhaisky Military Aerospace Academy*. 2014;642:210–215 (in Russ.).
5. Kuzhekin N.S., Chikurov V.A., Zinov'ev S.V. The assessment tools fund as a component of the basic professional educational program. *Trudy Voenno-kosmicheskoi akademii im. A.F. Mozhaiskogo = Proceedings of the Mozhaisky Military Aerospace Academy*. 2015;648:202–209 (in Russ.).
6. Altukhov A.I., Kalinin V.N., Cheburkov M.A. The experience of formation and assessment of competencies in the disciplines of the professional cycle in the system of military education. *Trudy Voenno-kosmicheskoi akademii im. A.F. Mozhaiskogo = Proceedings of the Mozhaisky Military Aerospace Academy*. 2016;650:204–210 (in Russ.).
7. Xie Y., Dervieux C., Riederer E. *R Markdown Cookbook*. Chapman and Hall & CRC Press; 2020. 360 p.
8. Rogova V.A., Shamin R.V. Optimization procedures in the problem of marketing educational services at the stage of forming a policy for recruiting applicants to universities. *Russian Technological Journal*. 2020;8(5):91–102 (in Russ.). https://doi.org/10.32362/2500-316X-2020-8-5-91-102
9. Panchenko L.F., et al. PhD student training: principles and implementation. *J. Phys.: Conf. Ser.* 2021;1840(1):12056. https://doi.org/10.1088/1742-6596/1840/1/012056
10. Estrellado R.A., Freer E.A., Motsipak J., Rosenberg J.M., Velásquez I.C. *Data Science in Education Using R*. London: Routledge; 2020. 304 p. https://doi.org/10.4324/9780367822842
11. Ellerbrock C.R., Abbas B., Dicicco M., Denmon J.M., Sabella L., Hart J. Relationships: The fundamental R in education. *Phi Delta Kappan*. 2015;96(8):48–51. https://doi.org/10.1177/0031721715583963
12. Nguyen S. Scholarly Git Experiences Part II: Git for Education. *IASGE*; 2019. Available from URL: https://investigating-archiving-git.gitlab.io/updates/git-for-education/
13. Kelleher J. Employing Git in the classroom. In: *World Congress on Computer Applications and Information Systems* (*WCCAIS*). 2014. http://dx.doi.org/10.13140/2.1.4202.0801
14. Wilson G. *Teaching Tech Together: How to Make Your Lessons Work and Build a Teaching Community around them*. Routledge & CRC Press; 2019. 260 p.

4. Алтухов А.И., Головина В.В., Калинин В.Н. Формирование и критерии оценивания компетенций в цикле математических и естественнонаучных дисциплин. *Труды Военно-космической академии им. А.Ф. Можайского*. 2014;642:210–215.
5. Кужекин Н.С., Чикуров В.А., Зиновьев С.В. Фонд оценочных средств как компонент основной профессиональной образовательной программы. *Труды Военно-космической академии им. А.Ф. Можайского*. 2015;648:202–209.
6. Алтухов А.И., Калинин В.Н., Чебурков М.А. Об опыте формирования и оценивания компетенций по дисциплинам профессионального цикла в системе военного образования. *Труды Военно-космической академии им. А.Ф. Можайского*. 2016;650:204–210.
7. Xie Y., Dervieux C., Riederer E. *R Markdown Cookbook*. Chapman and Hall & CRC Press; 2020. 360 p.
8. Рогова В.А., Шамин Р.В. Оптимизационные процедуры в задаче маркетинга образовательных услуг на этапе формирования политики набора абитуриентов в вузы. *Российский технологический журнал*. 2020;8(5):91–102. https://doi.org/10.32362/2500-316X-2020-8-5-91-102
9. Panchenko L.F., et al. PhD student training: principles and implementation. *J. Phys.: Conf. Ser.* 2021;1840(1):12056. https://doi.org/10.1088/1742-6596/1840/1/012056
10. Estrellado R.A., Freer E.A., Motsipak J., Rosenberg J.M., Velásquez I.C. *Data Science in Education Using R*. London: Routledge; 2020. 304 p. https://doi.org/10.4324/9780367822842
11. Ellerbrock C.R., Abbas B., Dicicco M., Denmon J.M., Sabella L., Hart J. Relationships: The fundamental R in education. *Phi Delta Kappan*. 2015;96(8):48–51. https://doi.org/10.1177/0031721715583963
12. Nguyen S. Scholarly Git Experiences Part II: Git for Education. *IASGE*; 2019. URL: https://investigating-archiving-git.gitlab.io/updates/git-for-education/
13. Kelleher J. Employing Git in the classroom. In: *World Congress on Computer Applications and Information Systems* (*WCCAIS*). 2014. http://dx.doi.org/10.13140/2.1.4202.0801
14. Wilson G. *Teaching Tech Together: How to Make Your Lessons Work and Build a Teaching Community around them*. Routledge & CRC Press; 2019. 260 p.

About the authors

Michail A. Eremeev, Dr. Sci. (Eng.), Professor, Department of Applied Information Techonologies, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: eremeev_m@mirea.ru. Scopus Author ID 57188205500, RSCI SPIN-code 3609-5733, https://orcid.org/0000-0002-5511-4000

Oleg V. Trubienko, Cand. Sci. (Eng.), Head of the Department of Applied Information Techonologies, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: trubienko@mirea.ru. Scopus Author ID 57213680467, RSCI SPIN-code 1668-5640, https://orcid.org/0000-0002-6809-8732

Ivan I. Zakharchuk, Cand. Sci. (Eng.), Assistant Professor, Department of Applied Information Techonologies, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: zaharchuk@mirea.ru. Scopus Author ID 57222072783, RSCI SPIN-code 1634-5302

Об авторах

Еремеев Михаил Алексеевич, д.т.н., профессор кафедры «Прикладные информационные технологии» Института кибербезопасности и цифровых технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: eremeev_m@mirea.ru. Scopus Author ID 57188205500, SPIN-код РИНЦ 3609-5733, https://orcid.org/0000-0002-5511-4000

Трубиенко Олег Владимирович, к.т.н., заведующий кафедрой «Прикладные информационные технологии» Института кибербезопасности и цифровых технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: trubienko@mirea.ru. Scopus Author ID 57213680467, SPIN-код РИНЦ 1668-5640, https://orcid.org/0000-0002-6809-8732

Захарчук Иван Илларионович, к.т.н., доцент кафедры «Прикладные информационные технологии» Института кибербезопасности и цифровых технологий ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: zaharchuk@mirea.ru. Scopus Author ID 57222072783, SPIN-код РИНЦ 1634-5302

Translated from Russian into English by K. Naqarov
Edited for English language and spelling by Thomas Beavitt