

# The Formation Method of Recommendations in the Process of Scientific Peer Review of Mathematical Papers

A. M. Elizarov<sup>1, 2</sup>[0000-0003-2546-6897], S. M. Khaydarov<sup>1</sup>[0000-0002-6835-3289]  
and E. K. Lipachev<sup>1, 2</sup>[0000-0001-7789-2332]

<sup>1</sup> N. I. Lobachevskii Institute of Mathematics and Mechanics,  
<sup>2</sup> Higher School of Information Technologies and Intelligent Systems,  
Kazan (Volga Region) Federal University  
amelizarov@gmail.com, 15jkeee@gmail.com, elipachev@gmail.com

**Abstract.** A method for the automation of the scientific peer review workflow in a mathematical journal has been developed. For this, a database of experts in various areas of mathematics has been created in the journal's information system. For each expert, using automatic search methods for citation databases, a profile has been generated that includes various classification codes, in particular, Mathematics Subject Classification 2010 (MSC 2010). A system of coefficients has been developed that takes into account the professional interests of experts. Also, indicators of their workload in the review process are taken into account. The first step in forming recommendations for the examination of an article is to determine the scope of the research conducted in it. For this, the codes of mathematical subject classification presented in this article are distinguished. Further, according to these codes, a list of possible experts is formed. The developed method is implemented as plugins of Open Journal Systems (OJS) and tested in the information system of the scientific journal Lobachevskii Journal of Mathematics (<https://ljm.kpfu.ru>).

**Keywords:** scientific journal information system, Open Journal Systems, peer review workflow, automated reviewers selection, Mathematics Subject Classification 2010, Lobachevskii Journal of Mathematics.

## 1 Introduction

Currently, there has been an increase in the number of scientific publications in professional journals, and the growth rate is more than 5% per year (see, for example, [1]).

Today it is universally recognized that after acquiring new scientific knowledge, the time needed to bring it to an interested audience should be as short as possible. Therefore, the task of ensuring the life cycle of a scientific document is currently being solved at new organizational and information levels (see, for example, [2]). All these processes are based on the widespread use of information and communication

technologies (ICT). At the same time, there is practically no difference between electronic and print publications. Without information services and specialized programs, it became impossible to prepare a scientific work for publication, and without network communications to publish it. In addition, new forms of scientific publications have appeared. The infrastructure of modern scientific publications has changed significantly: today it is not so much about the forms and means of using ICTs, but about creating a software platform that implements a developed system of services for working with digital content. Thus, a modern scientific journal should be based on a software platform that implements a developed system of services for working with digital content, based on ICT.

This article discusses the issues of expanding the functionality of the Open Journal Systems (OJS) platform in order to automate the work processes of the scientific journal editorial staff. A method for the automated selection of reviewers, the distribution of notifications and the control of the review period is proposed.

## **2 Automation of Business Processes of a Scientific Publication using Information Systems**

Currently, almost all scientific journals use software tools in their work. The main features of the use of ICT in the process of publishing scientific journals were studied in [3–5]. There, a comparison was made of existing information systems in terms of automation of editorial processes. From the existing information systems, the open platform Open Journal Systems was singled out.

OJS is licensed under the GNU / GPL. Today this information system is used by more than 9 thousand scientific journals around the world, in which there are at least 10 publications per year (<https://pkp.sfu.ca/ojs/ojs-usage/ojs-stats/>). OJS platform is used by many Russian journals (see, for example, <https://neicon.ru>).

OJS supports a wide range of business models for periodicals, as well as settings for providing access from fully open access to resources to brief annotations and commercial subscriptions. This allows you to use this system as a single platform for managing a complex of electronic publications (for example, in a research or educational organization) [6].

OJS System is configured as a cloud-based software platform that can be deployed and managed locally. All workflow processes are configured directly by the editors of each particular edition. OJS provides special tools for reading and viewing publications in PDF and HTML formats. A number of functions are available for working with bibliography and metadata. There are a number of other features.

OJS has a modular architecture, which allows you to develop your own classes and modules if necessary. The system has an MVC structure (Model-View-Controller), respectively, data storage, user interfaces and control functions are divided into different levels of interaction. This architecture provides fault tolerance, performance, flexibility and scaling of the entire system [7].

OJS uses a role model of users with various access rights and a multi-stage process of publishing resources. This process supports all stages of the life cycle of a scien-

tific article submitted for publication, from its initial download in the author's version to the publication of the final version on the Internet and the formation of relevant pointers and links. The functionality of the system allows the interaction of participants in the editorial process online. OJS interface modules are implemented as sets of Smarty templates (<http://smarty.php.net>), which allows you to change the user interfaces of the system.

### **3 Automation of the scientific peer review process**

An important distinguishing feature of the work of the editorial board of a scientific journal is the organization of independent peer-reviewing with the involvement of specialists in subject areas determined by the subject of the journal. The process of scientific review is, as a rule, the most time-consuming when considering scientific materials in the journal [1]. At the same time, the selection of reviewers for a qualified assessment of the work received in the journal, as well as the review itself, are critical in time. Automation of this process can reduce the time spent – this is also confirmed by practical experience (see, for example, [4]). Such automation involves the selection of a reviewer from the existing database of experts formed in the journal, taking into account their professional competencies, the amount of review and current workload.

The scientific review process support service allows you to select reviewers for each article submitted to the journal and automatically control the review process, including sending out notifications and reminders. One of the options for such a system of automation of the review process is implemented in the Lobachevskii Journal of Mathematics [4, 5].

## **4 Algorithm for Formation Recommendations for the Selection of Reviewers for Mathematical Papers**

### **4.1 Automated Selection of Reviewers for the Editorial Office of the Lobachevskii Journal of Mathematics**

Automated selection of reviewers of scientific articles was introduced into the practice of functioning of the Lobachevskii Journal of Mathematics (LJM) (<https://ljm.kpfu.ru/>) back in 2007 [4]. The implemented algorithm for selecting experts is based on an inaccurate comparison of the list of keywords given in the manuscript uploaded by the author into the journal's information system with the list of scientific interests contained in the created database of journal reviewers. In 2015, the editors of the journal switched to the publishing platform Open Journal Systems [5, 7], and this algorithm was modified taking into account the technical requirements of OJS [6]. Changes have also occurred in the editorial policy, requirements for the composition and formats of the materials presented, the scientific topic of the journal has expanded. The volume of the journal has significantly increased: from 4 issues a

year in the period 1998–2015 up to 12 issues currently. In addition, since 2008, LJM has been indexed in the scientometric databases Scopus, Web of Science and many bibliographic and abstracts databases (Zentralblatt Math, Google Scholar, Academic OneFile, Expanded Academic, Mathematical Reviews, OCLC, SCImago, STMA-Z, Summon by ProQuest, RSCI). As a result, the number of articles submitted to the journal has significantly increased, and the previously developed algorithm for selecting experts could not always sufficiently take into account the thematic diversity of articles. At the same time, the practice of issuing specialized thematic issues of the journal has significantly expanded. Therefore, there was an additional opportunity to connect reviewers with the editors-compilers of these thematic issues, and the existing automated system for selecting and appointing reviewers was transformed into a recommendation system.

The basis for the selection of recommendations, it was decided to use the mathematical subject classification Mathematics Subject Classification (MSC 2000, MSC 2010) [8], and other classification features and keywords extracted from the texts of articles, considered as clarifying [9].

## 4.2 Mathematical Subject Classification and their Applications

Mathematical Subject Classification (MSC) is created and maintained by Mathematical Reviews (MR, <http://www.ams.org/publications/math-reviews/math-reviews>) and Zentralblatt MATH (zbMATH, <https://zbmath.org>). This classification is used by all major mathematical journals, the archive of electronic publications of scientific articles and their preprints arXiv.org (<https://arxiv.org/>), and digital mathematical libraries [10]. The current version is MSC2010 (<http://msc2010.org/>; <https://mathscinet.ams.org/msc/msc2010.html>; <https://zbmath.org/classification>), the codes of the MSC2000 classifiers were used until 2010. The next version of the mathematical subject classification MSC2020 (<https://msc2020.org/>) has been announced.

MSC is a three-layer scheme using alphanumeric codes. The classifier code of the first layer consists of two digits, the second layer is determined by the capital letter of the Latin alphabet “A”, “B” or “C”, the third layer contains two digits. For example, the code “42C40 Wavelets and other special systems” refers to the third layer of the hierarchy, the parent code is “42C Nontrigonometric harmonic analysis”, and the code “42 Harmonic Analysis on Euclidean spaces” belongs to the top layer of the diagram.

MSC2010 contains 63 nodes of the first layer, starting with “00 General mathematics” and ending with “97 Mathematics education”. The numbering is not continuous, for example, codes “02”, “04”, “07”, “21” and some others are missing. The second layer contains 528 codes, and the third contains 5606.

In [11], the MSC2010 classifiers were converted to RDF Linked Data using the standardized Simple Knowledge Organization System (SKOS, <https://www.w3.org/TR/skos-reference/>). This makes it possible to assign classifiers not only to articles, but also to blog posts, figures, and formulas.

In [12], the results of machine learning MSC for the full text of articles in the mathematical digital libraries DML-CZ and NUMDAM are presented. These libraries have an article search service by MSC codes.

### 4.3 Main Steps of the Algorithm for Generating Recommendations for the Selection of Reviewers

The algorithm for generating recommendations for the selection of reviewers consists of two modules. The first module implements the creation and replenishment of the expert database in the journal information system, the second module selects experts for the specific article in question.

To form an expert profile, an automated search is conducted in the abstract and bibliographic databases of the articles of the author, included in the composition of experts. The MSC codes indicated in these articles (see Fig. 1) are included in the expert profile. In case codes MSC2000 are indicated, they are converted to codes MSC2010 [13].

The screenshot shows a search results page on MathSciNet. The search criteria are 'Items authored by Lipachev, E. K.'. The results list several articles with their MSC codes highlighted in red boxes:

- MR3830410: 35A02, 35C05, 35J25
- MR3525668: 94A14
- MR3408903: 94A08
- MR3284624: 68T30, 00A30
- MR2396104: 35A01, 35A02, 35A35, 35C05, 35J05, 35J25, 35P25, 35Q60, 41A15, 45B05, 45F05, 45G05, 65R20, 65Z05, 68T30, 78A45, 94A08, 94A14

Fig. 1. Extraction of codes of classifiers given in the expert profile in MathSciNet

Further, from the articles in which the expert participated in the review process (as a reviewer, compiler of the thematic issue or a member of the editorial board who submitted the article), the MSC codes are extracted and converted into MSC2010 codes. After that, the codes are added to the profile of the expert under consideration (see Fig. 2).

The screenshot shows the 'Managing Reviewer Profile' form in OJS. The fields are:

- ORCID ID: <http://orcid.org/0000-0001-7789-23>
- MathSciNet: <https://mathscinet.ams.org/mathscir>
- ZBMath: <https://zbmath.org/?q=ai%3Alipachev>
- URL: [Empty field]
- Phone: [Empty field]
- Fax: [Empty field]
- Reviewing interests: A grid of 24 MSC codes, each with a close button (x):
 

00A30	31B15	35A01	35A02	35A35	35C05	35J05
35J25	35P25	35Q60	41A15	45B05	45F05	45G05
65R20	65Z05	68T30	78A45	94A08	94A14	
- Private Notes: [Empty text area]
- Mailing Address: [Empty text area]

Fig. 2. Managing Reviewer Profile in OJS

If the classifier code is already present in the profile, the weight of this code increases and it is written to the database (see Fig. 3).

user_id	controlled_vocab_entry_id	tf
5	440	0.3333333333333333
5	441	0.0833333333333333
5	442	0.0833333333333333
5	443	0.0833333333333333
5	444	0.1666666666666667
5	445	0.0833333333333333
5	446	2.75
5	447	0.3333333333333333
5	448	0.0833333333333333

  

controlled_vocab_entry_id	locale	setting_name	setting_value	setting_type
446		interest	35J05	string

**Fig. 3.** Calculation of weights of professional interests of the reviewer

If the expert, after viewing the article, selects the option “UNABLE TO DO THE REVIEW” (refusal to review) on the “Review” page of the OJS information system, then the codes in the expert profile that match the article codes have a reduced weight. Figure 4 shows a fragment of an expert profile with codes MSC2010.

When adding classifier codes to the expert profile, information retrieval methods are used. In particular, a model of the vector space of documents is created, by analogy with which the profiles of reviewers act as documents, and the classifier codes from the reviewer's profile act as the corresponding terms:

$$tf(t, d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t', d}}$$

where  $f_{t,d}$  is the number of occurrences of the MSC 2010 code in the general list of codes in the reviewer profile  $d$ , and the denominator contains the total number of codes in this profile.

As a rule, the first code is the main code that determines the topic of the article. For clarification, the rest of the codes given in the article are used. Therefore, in the algorithm, the first classifier in the list is assigned the highest weight (for example, twice as many as the next).

ojs.kpfu.ru/index.php/ljm/manager/userProfile/5

**Evgeny Konstantinovich Lipachev**

Profile

EDIT PROFILE

2010 Mathematics Subject Classification

elipachev

—

Evgeny

Konstantinovich

Lipachev

Kazan Federal University

—

EKL

M

elipachev@gmail.com

—

—

—

00A30, 31B15, 35A01, 35A02, 35A35, 35C05, 35J05, 35J25, 35P25, 35Q60, 41A15, 45B05, 45F05, 45G05, 65R20, 65Z05, 68T30, 78A45, 94A08, 94A14

—

—

Russian Federation

—

English; Русский

January 31, 2013 - 11:30 AM

February 24, 2016 - 10:34 AM

Enrollment

Lobachevskii Journal of Mathematics

USER

You are logged in as...

AdminOJS

- My Journals
- My Profile
- Log Out

NOTIFICATIONS

- View (13 new)
- Manage

FONT SIZE

INFORMATION

- For Authors
- For Librarians

**Fig. 4.** Inclusion of classifier codes MSC2010 in the profile of the reviewer in the information system

After calculating the weight of each extracted code, an entry is made in the expert profile. To perform this operation, you need to make changes to the structure of the database table. To do this, OJS provides the ability to modify the table using the database schema DTD adopted from ADOdb (see, <https://github.com/pkp/pkp-lib/blob/master/dtd/xmlSchema.dtd>, [http://adodb.org/dokuwiki/doku.php?id=v5:axmls:axmls\\_index](http://adodb.org/dokuwiki/doku.php?id=v5:axmls:axmls_index)). So, for example, the “tf” field was added to the “user\_interests” table (see Fig. 3). This field stores the value obtained by the above formula. Reviewers are selected based on a vector model. In accordance with the *tf-idf* method for this article, weights are calculated from three levels of MSC classifiers:

$$tf(t, d) \times idf(t, D) = \left( 0.5 + \left( 0.5 + \frac{f_{t,d}}{\max_{t' \in d} f_{t' \in d}} \right) \right) \times \log \frac{N}{df}$$

Here  $t$  is the MSC code,  $d$  is the expert profile,  $D$  is the set of expert profiles,  $N$  is the total number of expert profiles,  $df$  is the number of experts who have the MSC code  $t$ .

Next, the cosine measure is calculated between the query vector generated using the previous formula and each expert profile from database of the system:

$$\cos(q, d_i) = \frac{(\vec{v}(q), \vec{v}(d_i))}{\|\vec{v}(q)\| \|\vec{v}(d_i)\|}$$

where  $q$  is the vector of the analyzed article, and  $d_i$  is the vector obtained from the reviewer's profile. Elements of these vectors are MSC codes.

The screenshot shows the 'Select Reviewer' interface in the OJS system. The submission details are as follows:

- Submission #843:** Poincaré group (IGNING submissions with totally untrivial fibers)
- Section:** General
- Editor:** None assigned
- Review Version:** 09.04.2019

The 'Peer Review' section is set to 'Round 1'. The 'Editor Decision' section shows options for 'Select decision', 'Decision', 'Notify Author', 'Review Version', 'Author Version', and 'Editor Version'.

The 'Reviewers' section is titled 'Select Reviewer' and includes a search bar for 'Reviewing interests' and a 'Search' button. Below the search bar is an alphabetical index (A-Z) and the subject classification '53020; 53C20; 53C50'.

The main table lists potential reviewers with their reviewing interests, recommendation percentages, and a 'DONE' status:

NAME	REVIEWING INTERESTS	RECOMMENDATION	DONE
[REDACTED]	General topology (Spaces with richer structures); Differential geometry (Global differential geometry)	67.0%	5
[REDACTED]	Global analysis, analysis on manifolds (Pseudogroups, differentiable groups and general structures on manifolds); Partial differential equations (General topics)	38.8%	1
[REDACTED]	Global analysis, analysis on manifolds (General theory of differentiable manifolds); Differential geometry (Global differential geometry)	35.2%	7
[REDACTED]	Differential geometry (Global differential geometry); Global analysis, analysis on manifolds (General theory of differentiable manifolds)	33.5%	7
[REDACTED]	Functions of a complex variable (Miscellaneous topics of analysis in the complex domain); Partial differential equations (Elliptic equations and systems)	29.2%	0

**Fig. 5.** Formed recommendation containing a list of reviewers most preferred for the selected article

In the next step, the cosine measure for each classifier layer is calculated and then summation is performed using a decreasing coefficient for each layer (0.5; 0.3 and 0.2, respectively):

$$score(q, d_i) = \sum_k \cos(q, d_i) \times \delta_k,$$

where  $\delta$  is the decreasing coefficient for each layer.

The assignment of reviewers in the OJS system is done by a user who has the roles of “Editor” or “Section Editor”. Reviewers are selected from the list of system users who have the role of “Reviewer”. This choice can be made only in manual mode – based on the professional experience of the editor.

The OJS system provides opportunities for changing its functionality by adding modules (plug-ins) compiled according to certain rules. Since the system is written in PHP and is open, it is possible to make changes to the functionality of the system without using special OJS tools. Finally, the result is provided at Figure 5.



## 5 Conclusion

A method for the automated selection of possible reviewers for scientific papers submitted for publication in the information system of a scientific journal is proposed, based on the use of Mathematics Subject Classification 2010 to determine the subject area to which the research results submitted for publication belong. The method is implemented as a service using the OJS plug-in technology and can be used in any mathematical journal that runs on the OJS platform. The method was tested in the information system of the Lobachevskii Journal of Mathematics.

This work was funded by the subsidy allocated to Kazan Federal University for the state assignment in the sphere of scientific activities, grant agreement No. 1.2368.2017). The work partially contains the results of the project “Monitoring and standardization of the development and use of technologies for storing and analyzing big data in the digital economy of the Russian Federation”, carried out as part of the program of competence of the National Technological Initiative “Center for storing and analyzing big data”, supported by the Ministry of Science and Higher Education Of the Russian Federation under the Treaty of Moscow State University named after M.V. Lomonosov with the Project Support Fund of the National Technological Initiative dated 08/15/2019 No. 7/1251/2019. The work was also carried out with the partial support of the Russian Fund for Basic Researches and the Government of the Republic of Tatarstan within the framework of scientific projects 18-29-03086, 18-47-160012.

## References

1. Binswanger, M.: Excellence by Nonsense: The Competition for Publications in Modern Science. In: Bartling S., Friesike S. (Eds). *Opening Science. The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*, pp. 49–72. Springer International Publishing (2014), [https://doi.org/10.1007/978-3-319-00026-8\\_3](https://doi.org/10.1007/978-3-319-00026-8_3).
2. Binfield, P.: Novel Scholarly Journal Concepts. In: Bartling S., Friesike S. (Eds) *Opening Science. The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*, pp. 155–163. Springer International Publishing (2014), [https://doi.org/10.1007/978-3-319-00026-8\\_10](https://doi.org/10.1007/978-3-319-00026-8_10).
3. Elizarov, A.M., Lipachev, E.K., and Khaydarov, S.M.: Automated system of services for processing of large collections of scientific documents. In: *CEUR Workshop Proceedings*, vol. 1752, pp. 58–64 (2016).
4. Elizarov, A.M., Lipachev, E.K., Malahal'cev, M.A.: Web-tehnologii v rabote jelektronogo matematicheskogo zhurnala Lobachevskii Journal of Mathematics. In: *NAUCHNYJ SERVIS V SETI INTERNET: MNOGOJADERNYJ KOMP'JUTERNYJ MIR. 15 LET RFFI. Trudy Vserossijskoj nauchnoj konferencii. Moskovskij Universitet im. M.V. Lomonosova, Juzhnyj Federal'nyj Universitet, Institut vychislitel'noj matematiki RAN*, pp. 355–356 (2007).
5. Elizarov, A.M., Zuev, D.S., Lipachev, E.K.: Lifecycle Management of Electronic Publications in Information Systems Scientific Journal. *Proceedings of Voronezh State University. Series: Systems analysis and information technologies* 4, 81–88 (2014).

6. Elizarov, A. M., Lipachev, E. K., Zuev, D. S.: Electronic Scientific Journal Management Systems. *Sci. Tech. Inf. Proc.* 41 (1), 66–72 (2014), <https://doi.org/10.3103/S0147688214010109>.
7. MacGregor, J., Stranack, K., Willinsky, J.: The Public Knowledge Project: Open Source Tools for Open Access to Scholarly Communication. In: Bartling S., Friesike S. (Eds). *Opening Science. The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*, pp. 165–175. Springer International Publishing (2014), [https://doi.org/10.1007/978-3-319-00026-8\\_3](https://doi.org/10.1007/978-3-319-00026-8_3).
8. Mathematics Subject Classification MSC2010, <http://msc2010.org/Default.html>, last accessed 2019/11/21.
9. Khaydarov, S.M, Yamalutdinova, G.S.: Recommender System of Physical and Mathematical Documents Classification. In: *CEUR Workshop Proceedings*, vol. 2260, pp. 480–486 (2018).
10. Elizarov, A.M., Lipachev, E.K., Zuev, D.S.: Digital Mathematical Libraries: Overview of Implementations and Content Management Services. In: *CEUR Workshop Proceedings*, vol. 2022, pp. 317–325 (2017).
11. Lange, C., Ion, P., Dimou, A., Bratsas, B., Sperber, W., Kohlhase, M., and Antoniou, I.: Bringing Mathematics to the Web of Data: The Case of the Mathematics Subject Classification In: Simperl, E., Cimiano, P., Polleres, A., Corcho, O., Presutti, V. (eds) *The Semantic Web: Research and Applications, ESWC 2012, Lecture Notes in Computer Science*, vol. 7295, pp. 763–777. Springer, Berlin, Heidelberg (2012), [https://doi.org/10.1007/978-3-642-30284-8\\_58](https://doi.org/10.1007/978-3-642-30284-8_58).
12. Řehůřek, R., Sojka, P.: Automated Classification and Categorization of Mathematical Knowledge. In: *INTELLIGENT COMPUTER MATHEMATICS, 9th International Conference, AISC 2008, 15th Symposium, CALCULEMUS 2008, 7th International Conference, MKM 2008, Birmingham, UK, July 28 - August 1, 2008. Proceedings*, pp. 1–15 (2008), [https://doi.org/10.1007/978-3-540-85110-3\\_44](https://doi.org/10.1007/978-3-540-85110-3_44).
13. Table of conversions between the 2000 and 2010 versions of the Mathematics Subject Classification (MSC), <http://msc2010.org/2000to2010.html>, last accessed 2019/11/21.