

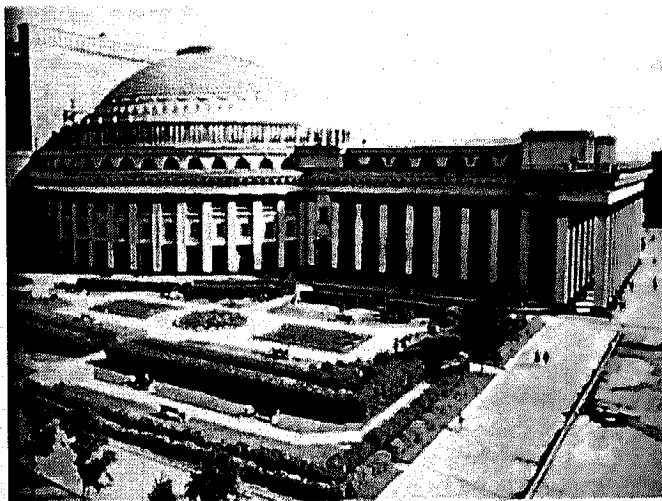


A Publication of the International Association of Science and Technology for  
Development – IASTED

---

Proceedings of the Second IASTED International Multi-Conference on

# **AUTOMATION, CONTROL, AND INFORMATION TECHNOLOGY**



---

## **SOFTWARE ENGINEERING**

Editors: Yu.I. Shokin, O.I. Potaturkin

June 20 – 24, 2005  
Novosibirsk, Russia

ISBN: 0-88986-504-3

ACTA Press

## SYSTEMATIZATION AND THE SEARCH OF MATHEMATICAL WEB-RESOURCES

V. Barakhnin, O. Klimenko  
 Institute of Computational Technologies of the SB RAS,  
 Lavrentiev pr.6, 630090, Novosibirsk, Russia  
 Email: bar@ict.nsc.ru, klimenko@ict.nsc.ru

### ABSTRACT

In the paper different mathematical classification schemes are examined. The methods of the systematization of mathematical web-resources are analyzed. The new method of the search for mathematical resources in treelike catalog is proposed.

### KEY WORDS

Scientific research, mathematical web-resources, classification schemes

### 1. Subject Classification Schemes

With the classification of mathematical web-resources it is necessary to have in the form that the different classification scheme must correspond to the different

types of resources. It is necessary to note that the compilers of classifiers considered below differently understand term "informatics". We include in survey the appropriate division when by "informatics" is understood "cybernetics", and we do not include, when by "informatics" is understood the "science about the scientific information".

First of all, the information systems, which contain different resources both scientific organizational (web page of scientific and training organizations and their subdivisions, dissertation councils, the personal pages of mathematicians, etc.) and scientific plan (first of all, mathematical web-publications), should be classified in accordance with their type [1].

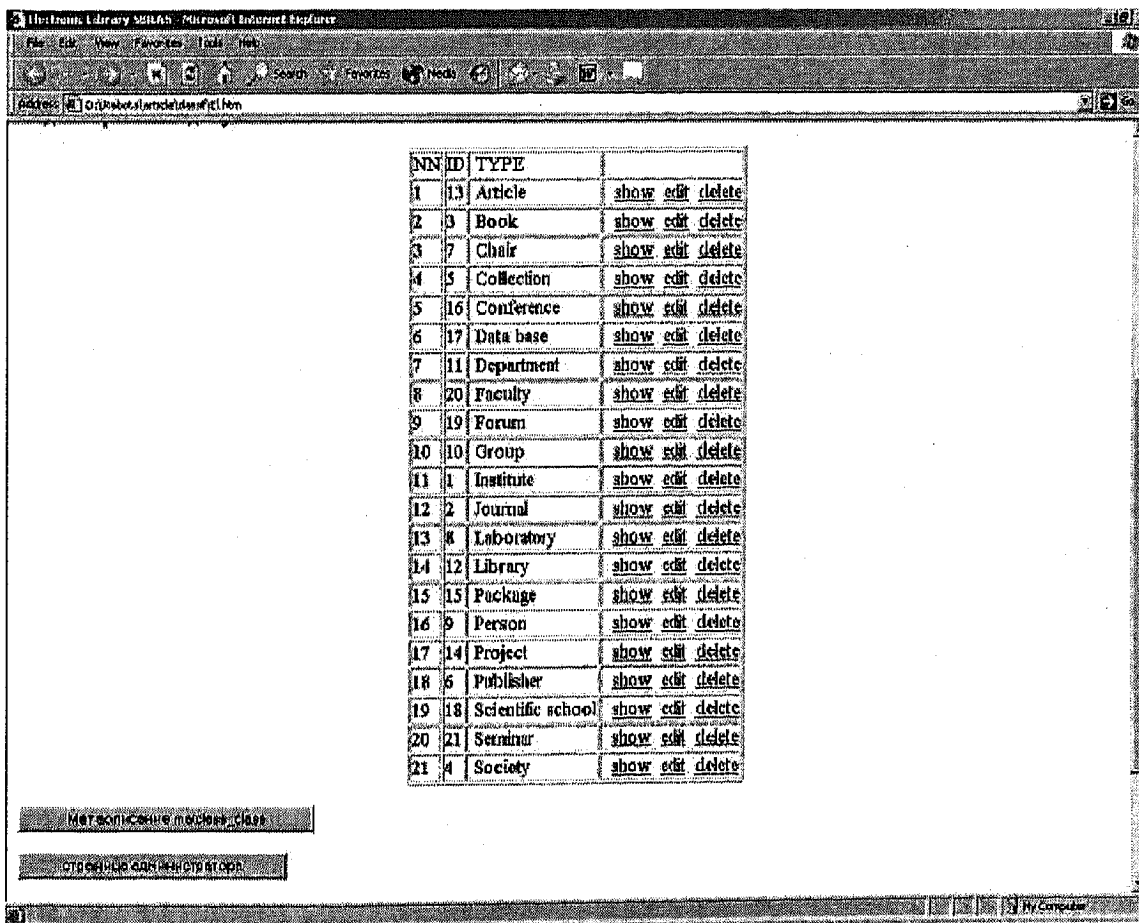


Fig. 1. Types of mathematical resources into information system "Mathematical Web-resources"

The resources of scientific organizational plan are classified in accordance with the rules, established by the appropriate scientifically administrative structure. Such classifiers consist of the small number of divisions, which correspond to the most general divisions of mathematical science.

In Russia scientist must prepare many documents for the protection of theses. These documents are classified with the use of Nomenclature of the specialties of scientific workers, affirmed by the Ministry of science and technologies of Russian Federation [2]. Nomenclature contains 8 designations of mathematical specialties and more than 10 designations of specialties, connected with the information theory.

The mathematical studies conducted in the Russian Academy of Sciences (RAN) are classified in accordance with "Basic directions of basic research" affirmed by Presidium RAN. They contain 13 divisions of mathematics and 6 divisions of information theory, which are concerned cybernetic questions.

Russian scientists must use the classifier, which contains 14 divisions of mathematics and 9 for information theory for the classification of the projects carried out with the support of the Russian Foundation for Basic Research (RFBR). However, absence of detailed information about the content of projects on RFBR site leads to the conclusion this classifier is not very useful.

The state rubricist of scientific and technical information (GRNTI) [3] occupies intermediate position by the volume between the brief scientific organizational classifiers and the detailed classifiers, intended for the systematization of the materials of scientific type (electronic publications of articles, the books, et cetera), accepted for the systematization of scientific and technical information in Russia. GRNTI is developed and is supported by Interdepartmental commission for classification of scientific and technical center.

GRNTI contains more than 140 divisions in the field of mathematics and more than 70 in the field of cybernetics (hierarchy in each of the regions two-level). This classifier is connected in its code part with the Nomenclature of the specialties of scientific workers. GRNTI is widely used in the distributed information systems, which work with the bases of data (for example, by server ZooPARK, created on the basis of protocol Z39.50). But mathematicians don't use GRNTI for the direct classification of mathematical web-publications.

Universal Decimal Classification (UDC) [4] and Mathematics Subject Classification (MSC) [5] are the

most suitable for the classification of scientific mathematical resources. UDC supported by International federation on the information and the documentation by consortium UDC. Russian version UDC is supported by VINITI (All-union Institute of Scientific and Technical Information). MSC utilized by the leading world abstract publications: "Mathematics Review" and "Zentralblatt MATH". We stopped our selection at Mathematics Subject Classification, which, in our opinion, is more fitted out for the solution of the problem presented than UDC by the following reasons:

1) At present the world community of mathematicians much more frequently uses MSC than UDC. For example, with the aid of MSC is conducted the data base "Zentralblatt MATH" that contains abstracts of more than 2 million articles. The presence of this base is extremely important for the solution of the problem of the automatic classification of the web- resources of the mathematical content.

2) MSC considerably more accurately reflects the contemporary state of mathematical science than UDC - the last processing MSC took place in 2000, and the division "Mathematics" of UDC in 1975 (into 1988 it underwent sufficiently insignificant changes).

It suffices to say that in UDC there is no term "Pseudodifferential operators". It is dedicated the class 35Sxx of second level for pseudodifferential operators in MSC. There is one point 512.736 for K- theory in UDC. There is isolated class of the first level 19-XX MSC for K-theory. It is including 11 classes of the second level, divided more than into 50 classes of the third level. In division 519.6 - "Computational mathematics, numerical analysis" - are indicated only the field of application of computational methods and indications of the type of these methods, in contrast to division MSC 65-XX, be absent. Such examples can be continued.

3) There are special classes in MSC, which cover history and philosophy of mathematics, sciences adjacent with mathematics (information theory and theoretical physics, including different branches of mechanics), and also dedicated to the application of mathematical methods in the natural and social sciences. In order to attain with the aid of UDC of this scope of concepts, it would be necessary to use not only divisions 51 - "Mathematics", 52 - "Astronomy. Astrophysics. Space research. Geodesy", 53 - "Physics", 004 - "Information technologies. Computer technology", but also many other divisions, which are concerned natural and social sciences, which would sharply increase the volume of redundant information. The alternative solution - to include in the created thesaurus only the separate, most connected with

mathematics, subsections of these divisions UDC - would destroy the integrity of classifier.

4) From the point of view of the number of those permitted for the mapping in the structure of the members of classification division at each of its steps UDC it relates to the classifications with the distinction constraints, since in it a quantity of classes of the sequential level is limited by 10. At the same time MSC, in which a quantity of classes of the first and third level is limited to 100, and the second - 26, which completely satisfies practical requirements, it can be related to classifications with the natural constraints.

5) The three-level structure MSC is considerably more fitted out for the automation of search and classification of information resources, than the structure UDC, which sometimes counts to 7 levels (number of levels it is calculated inside the division UDC "Mathematics").

However, classifier MSC has its deficiencies. Some divisions of mathematics are classified in several free places. For example, to inverse problems correspond classes 35R30, 49N45, 65M32, 8Γ22. This hampers understanding the development of the theory of inverse problems and retrieval for information on this thematics. However, a similar deficiency is inherent in UDC.

In MSC by certain large divisions of mathematics and special division correspond the classes of one level. For example, for the nonassociative rings and the algebras and for the numerical analysis, very extensive division which consists of many subsections, are selected the classes of one level: 16-XX and 65-XX respectively.

The frequently salient mathematical results are obtained at the joint of the different of the fields of mathematics. This bears out the fact that interpenetration of the different branches of mathematics gives impetus for the development of entire mathematics. There is a need for tracing trends in the development of mathematics and seeing the most important points of increase, predicting the appearance of new scientific disciplines both inside mathematics itself and in connection with other field of knowledges. To make this, using only classifier MSC, is extremely difficult.

## 2. Systematization and the search

In the Siberian Branch of RAN is conducted project on the creation of the treelike catalog of mathematical resources [6]. Drafters proposed the classification, based on MSC was decided to divide entire systemizer into seven main divisions (branches): algebra and logic; geometry and topology; analysis; differential, integral

and difference equations; the probability theory, mathematical statistics; mathematical programming and Computer Science; application and other.

Division into the basic branches is conditional and reflects the ideas of participants in the project about the development of mathematics. However, it makes it possible to facilitate the search for the resources, connected with the specific mathematical problem. It is possible to search for necessary information, moving on the branches of tree from the more general divisions to the quotients. Thus, it is possible to obtain information about the laboratories, the departments, personas, publications, and conferences, connected with united thematics.

It is possible to achieve a search on the keywords. Search can be not effective because in the keywords is indicated the not conventional term or because the author of the desired publication uses another terminology. The new terms, which not all know, appear. For example, one and the same axiom is called terms the "axiom of Archimedes" and "axiom of Eudoxus", widely known quadrature formula is called both the "formula of Simpson" and "formula of parabolas", there are tow terms for "polynomial row" in the Russian mathematical literature, etc.

In order to solve this problem it is proposed the method, when the knots of tree are connected with the word group from the thesaurus, which is composed by specialists in the concrete division. Then it is supplemented by the keywords, which are introduced with the description of the new of the service life of system. The keywords of all previous resources apply to the following. Requirements for the description of resource are reduced. Thesaurus is filled naturally.

Limitations to the knots of tree appear. The part of the thesaurus, connected with each knot, must not be too wide and too narrow. In the first case will have to carry out repeated search in the abundance of the resources, which will be connected with the data by knot. In the second case, will have to investigate set main for obtaining the necessary information.

There is a visualization of the tree of mathematics in the system (see Figure 2), which makes it possible to facilitate search. Subsequently it is intended with the search to show the knots of tree, in which are found the resources.

The correspondence between the divisions of catalog and the classes MSC is established in the represented treelike catalog. A question of the establishment of the correspondence between different classifiers is very important. If correspondence is established, it is possible to obtain information about the mathematical

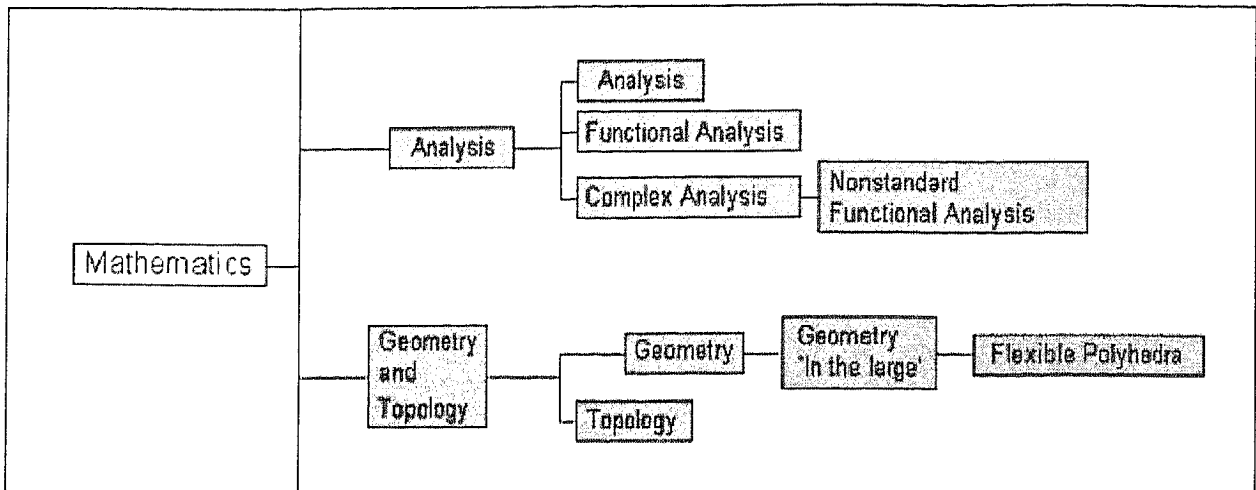


Fig. 2. Visualization of divisions “Analysis” and “Geometry and Topology” in the tree

resources, which are connected with one of the classifiers. However, the establishment of the correspondence between different classifiers is complex and ambiguous procedure. A similar work is carried, for example, in [7].

### 3. Conclusions

Thus, for the classification of the mathematical resources of scientific organizational plan is expedient to use nomenclature of the specialties of scientific workers, basic directions of basic research, and also the state rubricist of scientific and technical information. For the classification of mathematical web-publications is expedient to use Mathematics Subject Classification (MSC), modified in accordance with the scientific interests of mathematicians, who support concrete project. Search can be more effective, if it is conducted according to the treelike catalog, in which to each knot corresponds the group of the keywords.

### 4. Acknowledgement

This research was supported by Russian Foundation for Basic Research (project N 03-07-90423).

### References

- [1] Information system “Mathematical Web-resources”  
[[http://www-sbras.nsc.ru/win/elbib/data/show\\_page.dhtml?2+184](http://www-sbras.nsc.ru/win/elbib/data/show_page.dhtml?2+184)]
- [2] The nomenclature of the specialties of the scientific workers  
[<http://www.aspirantura.spb.ru/pasp/index.html>]
- [3] The state rubricist of scientific and technical information  
[<http://www.extech.ru/php/grnti/fl.php?.kod1=27>]
- [4] Universal Decimal Classification  
[<http://www.udcc.org/>]
- [5] Mathematics Subject Classification  
[<http://www.ams.org/msc/>]
- [6] Mathematical tree-catalog  
[<http://www.mathtree.ru>]
- [7] A. De Robbio, D. Maguolo, A. Marini, Mathematical subject classification and related schemes in the OAI framework. *Lecture Notes in Computer Science*, 2730, Springer, 100-111.