Fuzzy Relations Based Intelligent Information Retrieval for Digital Library Users

Marat Rakhmatullaev
Library information systems Department Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
Tashkent, Uzbekistan
marat56@mail.ru

Sherbek Normatov
Library information systems department Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
Tashkent, Uzbekistan
shb.normatov@gmail.com

Fayzi Bekkamov
Library information systems department Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
Tashkent, Uzbekistan
bfayzi93@gmail.com

Abstract. **It is known that today information and library systems are one of the main sources of information needs of the population and have many users. Information and library systems have a large amount of valuable information resources and information retrieval services have been established to allow users to find the necessary literature. We know that search engines take requests and return results they think are relevant. As a result, the user again faces the problem of finding what he needs among the many sources of information provided.**

Today, a number of information systems effectively use recommendation systems based on artificial intelligence to recommend objects. In information and library systems, high efficiency can be achieved by identifying the information needs of users and recommending relevant literature.

To do this, it is necessary to determine the information needs of library users by analyzing information about their age, interests, level of knowledge in a particular area, previous requests, professions, etc. By introducing recommender systems into library information systems, it is possible to facilitate the work of librarians, increase speed and accuracy finding the necessary source of information, increase the efficiency of management and the level of satisfaction of the information needs of the population.

The article proposes a fuzzy model for solving the problem of assessing the needs of users of information and library systems and recommending relevant literature to them.

Keywords: Digital library, user needs, fuzzy relations, recommender systems, scientific and technical information.

I. INTRODUCTION

The human need for information has been studied as research in the field of librarianship, psychology and computer science since the 1950-1960s of the last century [1]. Charles Naumer and Karen Fisher in their work explained the process of studying the need for information through the idea of "human-computer symbiosis" [2]. According to him, people and computers should work together to solve problems related to information needs. People's information needs are also unique, such as areas of interest, knowledge levels in a particular area, goals, and other characteristics. Today, despite the availability of search engines and information databases with a large volume of information sources, it can be noted that there is a lack of work on creating information resource recommendation systems that meet the information needs of the user.

The increase in the amount of information is normal, and many digital libraries usually do not take into account parameters such as the level of knowledge, skills and characteristics of users. The available digital library search engines can be used in both simple and advanced modes, and they return sources of information that match the keywords entered, reflecting user queries. As a result, the user faces several more information sources.

Digital library systems can also use artificial intelligence and machine learning to help users find the information they need. The conducted studies emphasize that significant efficiency of library processes can be achieved through the use of artificial intelligence-based systems in digital libraries [3], [4]. The purpose of these systems is to determine the information needs of the user by analyzing information about him and recommending suitable content or services.

In the process of determining the information needs of digital library users, it is necessary to process ambiguous and incorrect information provided by

linguistic variables, and the solution of this problem by deterministic or stochastic methods may not be effective enough. The purpose of this article is to study a fuzzy model for solving the problem of assessing the information needs of digital library users and recommending appropriate sources of information.

II. STAGES OF ASSESSING THE INFORMATION NEEDS OF DIGITAL LIBRARY USERS

In recent years, as artificial intelligence has developed, search engines have given way to recommender systems. Research shows that digital libraries can achieve greater efficiency by using recommender systems in finding information sources. The process of classifying users of digital libraries and sources of information according to a number of parameters plays an important role in creating a system of recommendatory sources of information in digital libraries. There are basically two types of recommender systems: recommender systems that aim to find similar sources of information or users with similar information needs [5]. For both of them the information needs of users can be identified in four steps.

a) Formation of types of information needs of users: At this stage, a set of parameters is formed, such as user interests, profession, age, demographic information, research areas, the level of knowledge in their field of science, and on their basis, possible information needs are classified. These actions are performed using either expert judgment or machine learning.

b) Collection of information about a specific user: Information about the registered user of the electronic library system is formed on the basis of previously known parameters. Naturally, not all parameters can be involved in this process.

c) Analysis of user data: At this stage, information about users is analyzed and it is determined what type of information needs they belong to. As a result, classes of users with the same information needs are formed. This allows you to recommend sources of information to users with similar information needs without expert judgment.

d) Create a user profile: At this stage, initial information about the user of the electronic library and data sets are created that characterize the dynamics of information needs.

III. ACCOUNTING FOR THE LEVEL OF COMPETENCE OF USERS IN THE AREA UNDER STUDY

When searching for information, it is necessary to take into account the level of competence of the user. Even the same information request can give a result that suits users in different ways. A survey was conducted of 148 users of 6 different levels of competence and education: Researcher, Lecturer, Engineer, Applicants for academic degrees, undergraduates and doctoral students, Student, Businessman. The respondents were people who work or study in the field of information technology.

Everyone was given the same question: "What types of information sources are more important to you? Please rate on a scale of 100. The following types of information sources were indicated: Scientific research (articles and books); Articles on technologies; Dissertations; Materials on professional technologies (scientific and applied reports); encyclopaedias; Standards; Textbooks, teaching materials; Articles on business.

The survey gave the following results. (Table 1).

TABLE 1 SURVEY RESULTS OF USERS OF DIFFERENT LEVELS COMPETENCE AND EDUCATION

#	Interest Category and materials	Users					
		Researcher	*Teacher*	*Engineer*	*Undergraduates and PhD students*	*Student*	*Businessman*
1	Scientific research (articles and books)	98.3	82.5	40.2	99.4	20.1	12.6
2	Articles on technology	22.4	32.1	99.2	70.3	56.6	42.2
3	Dissertations	99.6	80	56.7	98.7	15.6	10.4
4	Materials for prof. Technologies	51.4	50	99.6	50.4	47.3	21.3
5	Encyclopedias	92.4	63.9	51.3	54.3	42.4	11.2
6	Standards	10.6	32.5	90.4	41.5	21.5	52.4
7	Textbooks, teaching materials	71.4	100	43.1	72.8	99.7	12.5
8	Articles on business	22.5	20,5	47.1	31.6	13.1	99.8

As can be seen from Table 1, different categories of users have different needs by types of information sources. For users, they are close. For example, researchers and teachers may have similar requests. For others, they may differ dramatically. Accordingly, when forming search images (queries) and developing search algorithms in databases, it is necessary to take into account the categories of users. These search algorithms should allow "cutting off" the unnecessary and unimportant part of the search results, as well as providing more information that corresponds to users of this particular category.

Accounting for the category and competence of users should allow:

• Reduce the time of searching for information,

• Increase the share of the information that interests the user the most and reduce the share of non-important information.

To further reduce the amount of data that includes redundant information that may not be useful to the user, fuzzy correspondence models can be used [6]. These models of the type "Situation-Cause" and "Situation-Cause-Action" are used with situational advising systems

and have great prospects for solving problems that are difficult to formalize.

The use of fuzzy logic methods in the search for scientific and educational information is advisable if the following conditions are present:

a)   Lack of an unambiguous correspondence between the request and the source of information;

b)   The impossibility of arbitrarily accurate measurement of the amount and volume of data on search images;

c)   The impossibility of a complete and clear description of the search image;

d)   The inaccuracy of the functional actions performed by the system, which often do not achieve the goals set by the system when searching for data in the database;

e)   Insufficient dimension of the model, which does not allow to reflect all the significant properties of the search object.

f)   Not all query elements can be expressed in one or more words, even using Boolean algebra in its formation.

g)   Insufficient "literacy" of the reader in the formulation of their requests.

IV.   ALGORITHM FOR CONSTRUCTING A FUZZY COMPOSITION

It is known that the information needs of a particular user are characterized by such parameters as his interests, level of knowledge in a particular area, goals, and these parameters are characterized by uncertainty. Therefore, the composition of fuzzy correspondence can be used to model the task of determining the need of a digital library user for sources of scientific and technical information and recommending suitable sources of information. In studies conducted on the construction of recommender systems using fuzzy correspondences and their composition, the high efficiency of this model was noted in solving problems of this type [7] , [8].

Let us construct a fuzzy model based on binary fuzzy relations $\tilde{S}$ and $\tilde{T}$. To construct the first relation, we will use two basic sets X and Y. For the second, we will use Y and Z. In this example, the set X describes the finite set of digital library users:$X = \{x_1, x_2, \dots x_n\}$. Y is the set of types of user information needs (TUIN) of electronic libraries: $Y = \{y_1, y, \dots y_m\}$. Z is a set of types of information sources in digital libraries: $Z = \{z_1, z_2, \dots z_l\}$.

In the context of this task, the relation $\tilde{S}$ will describe the characteristics of users of electronic libraries, and the relation $\tilde{T}$ will describe the qualitative indicators of the correspondence of information sources with types of users.

To find the correspondence $\tilde{S}$, the membership function $\mu_S(x, y)$ to the Cartesian product $X \times Y$ is constructed. Here $(x, y) \in X \times Y$. It should be noted that the $X \times Y$ correspondence table is filled in by the user himself.

In the Cartesian multiplication $Y \times Z$, the fuzzy correspondence $\tilde{T}$ is determined by the correspondence function $\mu_T(y, z)$. Here $(x, y) \in X \times Y$. Let us assume that the values of the membership functions were obtained by expert means or as a result of machine learning.

The fuzzy correspondence between $\tilde{S}$ and $\tilde{T}$ makes it possible to determine what type of user a digital library user belongs to and what sources of information are suitable for each type of user. Having built a composition of fuzzy matches $\tilde{S}$ and $\tilde{T}$, the stage of recommending which source of information is passed to each user:

$$\tilde{S} \circ \tilde{T} = \{(x, z), max_{y \subseteq Y}(\min (\mu_{\tilde{S}}(x, y),$$

$$\min (\mu_{\tilde{T}}(y, z)))\} \qquad (1)$$

or

$$\mu_{\tilde{S} \circ \tilde{T}}(x, z) = \bigvee_{y \subseteq Y}\{\mu_{\tilde{S}}(x, y) \wedge \mu_{\tilde{T}}(y, z)\} \quad (2)$$

here $x \in X, y \in Y, z \in Z$.

$\mu_{\tilde{S}}$ and $\mu_{\tilde{T}}$ are the membership function of fuzzy relations on fuzzy sets.

V.   EXAMPLE

Let the sets X, Y and Z be given in the following form:

$X = \{user_1, user_2, user_3\};$

$Y = \{TUIN_1, TUIN_2, TUIN_3, TUIN_4, \};$

$Y = \{source_1, source_2, source_3, source_4,$

$source_4, \}$

TABLE 2 FUZZY RELATION $\tilde{S}$: "USER – TYPE OF USER INFORMATION NEEDS (TUIN)"

	TUIN1	TUIN2	TUIN3	TUIN4
User1	0.4	0.9	0.2	0.1
User2	0.3	0.5	1	0.2
User3	0.2	0.5	0.3	0.9

TABLE 3 FUZZY RELATION $\tilde{T}$: "TYPE OF USER INFORMATION NEEDS – SOURCE"

	Source1	Source2	Source3	Source4	Source5
TUIN1	0.5	0	0.3	0.7	1
TUIN2	0.9	0.6	0	0.2	0.6
TUIN3	0.7	1	0	0.6	0.3
TUIN4	0	0.2	0	0.5	1

At first we compute $\tilde{S} \circ \tilde{T}$ by using max-min composition.

$$\mu_{\tilde{S} \circ \tilde{T}}(x_1, z_1) = \max\{\min(0.4, 0.5), \min(0.9, 0.9), min(0.2, 0.7), \min(0.1, 0)\} = max\{0.4, 0.9, 0.2, 0\} = 0.9$$
$$\mu_{\tilde{S} \circ \tilde{T}}(x_2, z_1) = \max\{\min(0.3, 0.5), \min(0.5, 0.9), min(1, 0.7), \min(0.2, 0)\} = max\{0.3, 0.5, 0.7, 0\} = 0.7$$
$$\mu_{\tilde{S} \circ \tilde{T}}(x_3, z_1) = \max\{\min(0.2, 0.5), \min(0.5, 0.9), min(0.3, 0.7), \min(0.9, 0)\} = max\{0.2, 0.5, 0.3, 0\} = 0.5$$
$$\mu_{\tilde{S} \circ \tilde{T}}(x_1, z_2) = \max\{\min(0.4, 0), \min(0.9, 0.6), min(0.2, 1), \min(0.1, 0.2)\} = max\{0, 0.6, 0.2, 0.2\} = 0.6$$
$$\mu_{\tilde{S} \circ \tilde{T}}(x_2, z_2) = \max\{\min(0.3, 0), \min(0.5, 0.6), min(1, 1), \min(0.2, 0.2)\} = max\{0, 0.5, 1, 0.2\} = 1$$
$$\mu_{\tilde{S} \circ \tilde{T}}(x_3, z_2) = \max\{\min(0.2, 0), \min(0.5, 0.6), min(0.3, 1), \min(0.9, 0.2)\} = max\{0, 0.5, 0.3, 0.2\} = 0.5$$

Using max-min composition, we get the following result:

TABLE 3 FUZZY RELATION $\tilde{S} \circ \tilde{T}$: "USER – SOURCE"

	Source1	Source2	Source3	Source4	Source5
User1	0.9	0.6	0.3	0.4	0.6
User2	0.7	1	0.3	0.6	0.5
User3	0.5	0.5	0.2	0.5	0.9

Table 3 shows that *Source1* can be recommended for *User1*. *Source2* for *User2* is fully compatible, it is recommended to use *Source5* for *User2*. This algorithm can be used when the types of information needs of users are known.

VI. CONCLUSION AND FURTHER WORK

The proposed composition of fuzzy correspondences "user - user type" and "user type - information source" can be considered as the basis for building an expert system that allows users of digital libraries to recommend information sources suitable for their information needs. When the information needs of users need to be determined by several quantitative or qualitative parameters, each parameter can be specified in the form of fuzzy variables. Fuzzy modelling is more convenient in the process of describing and processing expert knowledge that is specified in linguistic variables.

REFERENCES

[1] Bates, Marcia J. Information Behavior In Encyclopedia of Library and Information Sciences, 3rd Ed. 2010.

[2] Charles Naumer, Karen E. Fisher, Information needs, Encyclopedia of Library and Information Sciences, 2017, doi: 10.1081/E-ELIS3-120043243

[3] 3.Zhang, Q., Lu, J. & Jin, Y. Artificial intelligence in recommender systems. Complex Intell. Syst. 7, 439–457 (2021). doi: 10.1007/s40747-020-00212-w

[4] Cox, Andrew. (2022). How artificial intelligence (AI) might change academic library work: applying the competencies literature and the theory of the professions. Journal of the Association for Information Science and Technology. doi: 74.10.1002/asi.24635

[5] F.O.Isinkaye, Y.O.Folajimi, B.A.Ojokoh. Recommendation systems: Principles, methods and evolution, Egyption Informatics Journal 2015.

[6] Rakhmatullaev M. Situational Management of Complex Information Systems Based on Fuzzy Correspondence Models. ICISCT-2019, TUIT named after Muhammad Al-Khwarizmi, Tashkent, Uzbekistan_4-6.11.2019. doi: 10.1109/ICISCT47635.2019.9012035

[7] N. Cao, R. Valášek and S. Ožana, "Composition Models of Fuzzy Relations Considering Importance Levels of Features," 2022 14th International Conference on Knowledge and Systems Engineering (KSE), Nha Trang, Vietnam, 2022, pp. 1-6, doi: 10.1109/KSE56063.2022.9953776

[8] Morawski, Jason & Stepan, Torin & Dick, Scott & Miller, James. (2017). A Fuzzy Recommender System for Public Library Catalogs. International Journal of Intelligent Systems. doi: 32.10.1002/int.2188