Forecasting the Reader's Demand Level Based on Factors of Interest in the Book

Vsevolod Senkivskyy[1], Sergii Babichev[2,3], Iryna Pikh[1,4], Alona Kudriashova[1], Nataliia Senkivska[1,4] and Iryna Kalynii[5]

[1]*Ukrainian Academy of Printing, 19, Pid Holoskom St., Lviv, 79020, Ukraine*
[2]*Jan Evangelista Purkyne University in Usti nad Labem, Ceske mladeze, 8, Usti nad Labem, 40096, Czech Republic*
[3]*Kherson State University, Universytetska st. 27, Kherson,73003, Ukraine*
[4]*Lviv Polytechnic National University, 12, Stepana Bandery St., Lviv, 79013, Ukraine*
[5]*Berezhany Agrotechnical Institute, 20, Academichna St., Berezhany, 47501, Ukraine*

Abstract

The selection and description of a set of factors related to the study of the process of interest in the book has been done, which became the initial precondition to assess the prognostic level of reader's demand. The growth of risks of the low reader's demand level among young people is indicated. A formalized version of the relationships between the factors of demand for the book is represented using a semantic network that provides a graphical and linguistic representation of reading intensity factors. Using the methodology of hierarchy modelling, the levels of factors preferences are established and the weight priorities of their influence on the studied process are calculated. An optimized variant of factors ranking according to the importance of forming the intensity of the reader's demand is obtained. Based on the methods of fuzzy set theory, the membership functions of linguistic variables (factors of interest in the book) are calculated, fuzzy knowledge bases are formed and fuzzy logical equations are derived, which became the basis for prognostic assessment of the reader's demand level.

Keywords

Factor, reader's demand, risks of low demand for the book, fuzzy knowledge base, fuzzy logical equations, fuzzification, defuzzification.

1. Introduction

There is a book in front of us – a complex and ingenious work of the human spirit. It provides the transfer and assimilation of knowledge, reproduces the life history of individuals and civilizations, enables the communicative connection between different nationalities. While reading the book, few people think about the difficult path of thought (ideas, considerations, views, life stories) before becoming a printed word in a usual format that ensures its dissemination, use and, consequently, impact on human community.

CITRisk'2021: 2nd International Workshop on Computational & Information Technologies for Risk-Informed Systems, September 16–17, 2021, Kherson, Ukraine
EMAIL: senk.vm@gmail.com (V.Senkivskyy); sergii.babichev@ujep.cz (S.Babichev); pikhirena@gmail.com (I.Pikh); kudriashovaaliona@gmail.com (A.Kudriashova); senkivskanata@gmail.com (N.Senkivska); iryna.kalynii@gmail.com (I.Kalynii)
ORCID: 0000-0002-4510-540X (V.Senkivskyy); 0000-0001-6797-1467 (S.Babichev); 0000-0002-9909-84444 (I.Pikh); 0000-0002-0496-1381 (A.Kudriashova); 0000-0002-6096-2282 (N.Senkivska); 0000-0003-1296-6454 (I.Kalynii)


A significant amount of research connected to the book relates to the processes and factors of formation and prognostic assessment of the edition quality, taking into account the technological stages of its production [1, 2]. Sociological and historical studies related to the book business, the issue of book distribution also attract the attention of researchers [3, 4]. In contrast, the problem of the demand for a book concerning the "life" of a book after it leaves the walls of a printing house remains somewhat unexplored. The active age of a book depends on the spiritual needs of its fans, which are determined by the reader's interest and many preconditions of an objective nature.

The reader's demand for a book, like any other process of human activity, is characterized by certain factors that affect its intensity. One of the important tasks will be the separation, description and structured representation of the relationship between factors, reasonable ranking by levels of importance, calculation of conditional weights and synthesis of structured, classified as multilevel, display of the priority of factors on the book demand based on the obtained data. The formulation and solution of problems of formalized assessment of the reader's demand summarizes the study of the process of interest in the book, which involves the formation of term-sets of values of linguistic variables that identify the factors of interest in the book, the calculation of membership functions of variables, the development of fuzzy knowledge bases of linguistic variables and the derivation and solution of fuzzy logical equations, which has become the basis for prognostic assessment of the reader's demand level based on the methods of fuzzy logic.

The solution of such weakly formalized problems should be carried out on the basis of informational approaches to systems analysis, which allow the use of universal tools of hierarchical systems theory, modelling theory, fuzzy sets and fuzzy logic methods to achieve numerical characteristics of the studied process.

2. Problem Statement

The formalization of the input database of the studied subject area should be performed using semantic networks [6, 7] and the method of hierarchy analysis [8] to establish the priority of the influence on the described process of a set of the related factors (the first stage of the problem), as well as fuzzy set theory means [9-11], which will provide a numerical expression of the reader's demand level, taking into account the term-set of values of linguistic variables (the second stage). The application of these approaches will not only rank the factors of the reader's interest in the book, which in itself will improve the thematic planning of the publishing activity and the book distribution processes, but also provide forecasting the reader's demand level taking into account possible combinations of values from a predetermined set.

The nodes of the semantic network will reflect the semantics of concepts, i.e. factors that in the second stage of the study are presented in the form of abstract linguistic variables. The arcs represent functional (semantic) relationships or connections between them [1, 6]. The combination of linguistics (semantics of linguistic variables) and mathematics (networks as a variant of the graph) provides, on the one hand, the use of ordinary language to describe the knowledge base of the studied process, and on the other hand, it allows the use of formal methods and fuzzy logic for the research, the ultimate goal of which is the prognostic assessment of the reader's demand level.

The model of the process of forecasting the reader's demand intensity can be presented by a set of appropriate steps, the implementation of which is shown in Fig. 1.

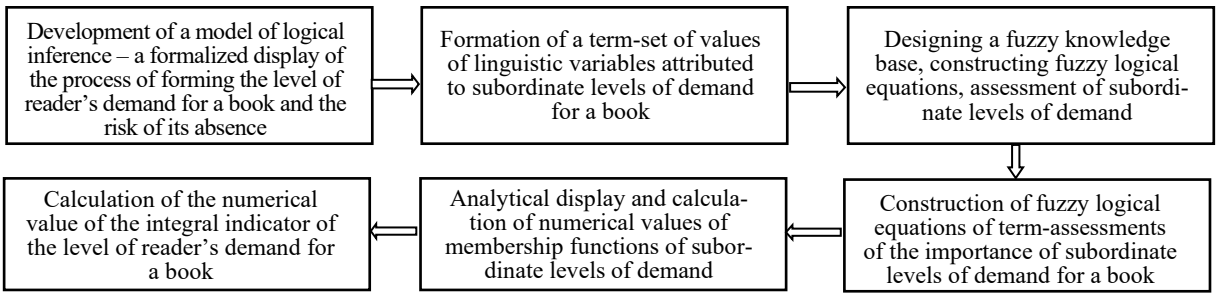

Figure 1: Structural and functional model of the process of calculating the level of reader's demand for a book

The essential element and advantage of fuzzy logic is the possibility of fuzzification, i.e. the replacement of components of a certain set with the corresponding concepts of a fuzzy set. It is known that its essence is to compare the term-set of values of the analysed factors corresponding to the fuzzy format of variables – membership functions. Fuzzification provides a fairly high level of conformity of the model to the real object and serves, as it will be shown later, as a basis for further modelling of the prognostic assessment of the reader's demand level.

In the works of the founder of fuzzy logic Zadeh [12, 13], the concept of a universal set D is introduced as such, which applies to the whole problem area. Then the fuzzy subset M of the set D is determined through the scale D and the membership function $\mu_M(d)$ [2], i.e.:

$$M = \left\{ \left(\mu_M(d), d \right), d \in D \right\}, \tag{1}$$

where $\left(0 \le \mu_M(d) \le 1\right)$.

The membership function establishes the degree to which each element of a fuzzy set belongs to a universal set, i.e. $M \in D$. Under the condition of the discreteness and finiteness of the basic scale (i.e. divided into quanta or intervals) the fuzzy set is:

$$M = \left(\mu_M(d_1)/d_1, \mu_M(d_2)/d_2, ..., \mu_M(d_n)/d_n \right) = \sum_{i=1}^{n} \mu_M(d_i)/d_i, \tag{2}$$

or simplified: $M = \sum_{i=1}^{n} \mu_i / d_i$. The record means "attachment" ΦН $\mu_M(d_i)$ to the element d_i.

Eventually the membership functions act as an identifier of the input values of linguistic variables in a fuzzy format, i.e. the set of values of the variable d is matched to the membership function $\mu(d)$.

3. Related Works

In the list of literary sources, much attention is paid to the study of factors that shape the habit of reading throughout a person's life and the state of the book and newspaper and magazine market related to it. Thus, the paper [5] justifies the need to study the reading audience, arguing that the proper implementation of all technological procedures for the production of books, newspapers or magazines without taking into account the information about the end reader can not fully ensure

the quality. In [14] it is about the influence of such factors as the purpose of reading, the level of education, gender, income, access to the Internet, etc. on the choice of the carrier of book products: e-books or printed on paper. It is also noted that e-books, despite their popularity, cannot replace printed ones because they have unique characteristics. However, the emergence of e-books contributes to the mass distribution and, consequently, the availability of books, on which the level of reading depends. Studies [15, 16] confirm the existence of a relationship between the level of reading and social status, cultural level, lifestyle. According to [17, 18], it is the family that instils the habit of reading in a person in the early life stages, shaping him as a person. In [18] it is noted that the profession and the level of education of parents significantly affect the interest in reading in children. [19, 20] confirm the influence of living environment, place of study and family on motivation to read. [21] presents a study of the level of reading and describes the results of a survey in which most respondents noted one of the main factors being the influence of the family. In contrast to the more definite influence of the above-mentioned factors, [16] deals with the obvious relationship between the age of the reader and the level of his reading, and in [22] – between the gender and the level of reading. It can be concluded that a person's socio-cultural environment, personal characteristics and access to literature determine the level of reading to a greater extent than the age or gender. The above considerations are taken into account in further research when choosing a set of factors influencing the level of reading, their interdependencies and term-sets.

It should be noted that the analysed works do not give a clear idea of the priority of factors and do not determine their impact on the level of a person's reading in the quantitative area. The essence of these publications is exclusively in the analytical and sociological description of the problem. That is why, in order to enable not only assessing, but also prognostic activity (and, as a consequence, corrective one), it is expedient to establish clear relationships between factors, their priorities and to carry out prognostic assessment of the reader's demand level.

The analysis of publications related to the above issues characterizes the lack of completeness of scientific research related to the formation of components of the information database, focused on the study of a poorly formalized social problem – the reader's demand for a book. In the vast majority of cases, researchers focus on processing sociological and statistical data, which, despite the relevance and demand as to the ways of intellectual development of society, does not provide a final prognostic assessment of the level of reader's interest in the book, which would operate in a separate expert set of factors of demand for printed products.

The above considerations determine the application of non-traditional, in our opinion, informational approach in this area, the result of which will be, on the one hand, the ranking of factors by importance and numerical weights based on the semantic network of relationships between them, which will synthesize multilevel structural models of priority influence of factors on reader's demand; on the other hand, prognostic numerical assessment of the reader's demand intensity using fuzzy set theory (fuzzy logic), including a hierarchical model of logical inference, membership functions of linguistic variables, fuzzy knowledge base and fuzzy logical equations, calculation of numerical value of generalized forecasting of interest in a book through the defuzzification of the linguistic term "level of reader's demand".

A separate problem today is a rather low reader's demand level among young people, and the risks of its increasing require separate research and coverage in scientific publications. This, however, does not detract from the general interest in the printed book.

4. Materials and Methods

According to the task, the research of the process of forecasting the intensity level of demand for a book, is performed in two stages, focusing first on obtaining a multilevel model of factors, structured by the importance of influencing readers' preferences, which will be the initial information component of determining and calculating a numerical indicator of the reader's demand level.

4.1. Semantic network of the reader's demand factors

For a formalized description and content filling of the subject area, a graphic representation is used in the form of semantic networks, which will highlight significant aspects of knowledge about the factors influencing the fundamental ability of forecasting the behaviour of readers as to the level of their interest in a book [1, 2, 5].

Let the list of factors related to the process of assessing the demand for a book contain a mathematical notation and its semantic interpretation. The nature of the factors relates more to the identity of the reader, and to a lesser extent to a specific book, which can be seen from the following description: x_1 – place of residence; x_2 – level of education; x_3 – profession (occupation); x_4 – content, subject of the book; x_5 – availability (accessibility) of literature; x_6 – family (role of the family); x_7 – reading traditions; x_8 – social status. The semantic network of connections between the above factors is shown in Fig. 2. The vertices of the network-graph identify the linguistic factors-arguments of the set $X = \{x_1, x_2, ..., x_8\}$, the arcs are pairs of vertices (x_i, x_j), for which the connection ($i, j = 1 \div 8$; $i \neq j$) is defined.

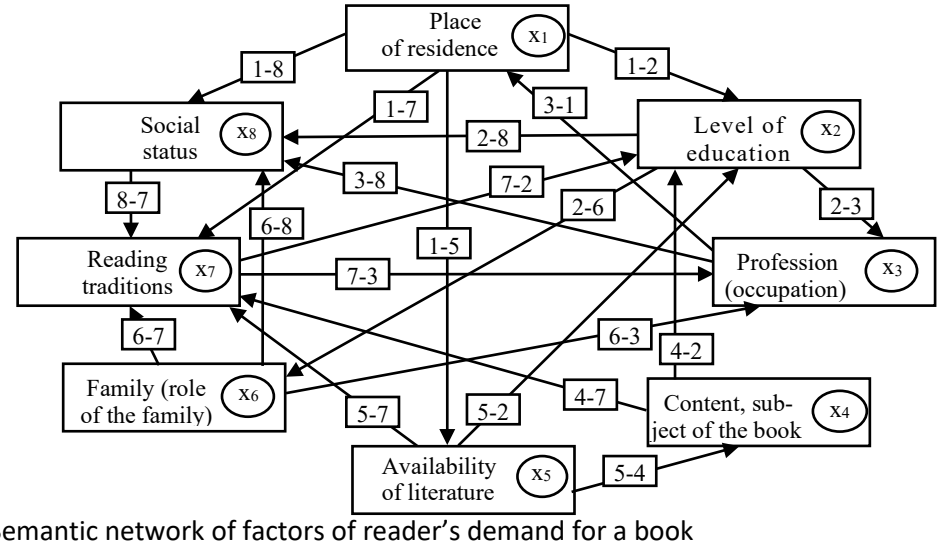

Figure 2: Semantic network of factors of reader's demand for a book

To obtain a model of the priority influence of factors on reader's demand, the method of hierarchy analysis has been used [8], the implementation of which involves: the construction of a pairwise comparison matrix of factors using the scale of relative importance of objects; the calculation and normalization of the values of the components of the main eigenvector of the matrix, which determine the weight advantages of the factors; the verification of the obtained results according to the

criteria of the maximum value of the eigenvector, normative values of the consistency index and the consistency ratio; the establishment of levels of priority influence of factors on reader's demand.

4.2. Model of fuzzy logical inference

Let the level of reader's demand be the function $Q = F\left(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\right)$, the arguments of which are the factors described above. Then the value of the function will determine the prognostic integrated indicator of the reader's demand level Q, which should be divided into partial indicators, according to the semantic load: $Q = F_Q\left(A, B, C\right)$.

The argument A determines the total indicator, which shows the level of influence of the socio-cultural environment and contains the linguistic variables a_1 – "place of residence" (belonging of the settlement to a certain category by population), a_2 – "family" (the role of the family), a_3 – "reading traditions": $A = F_A\left(a_1, a_2, a_3\right)$.

The argument B identifies the total indicator that accumulates the level of personal indicators and includes the linguistic variables b_1 – "level of education", b_2 – "profession" (occupation), b_3 – "social status" (social class): $B = F_B\left(b_1, b_2\right)$.

The argument C is related to the total indicator of the level of the book market and contains the linguistic variables c_1 –"content" (subject); c_2 – "availability and accessibility of literature": $C = F_C\left(c_1, c_2\right)$.

The above structuring makes it possible to represent the process of forming an integrated indicator of the reader's demand level on the basis of the model of logical inference, taking into account the values of linguistic terms of factors.

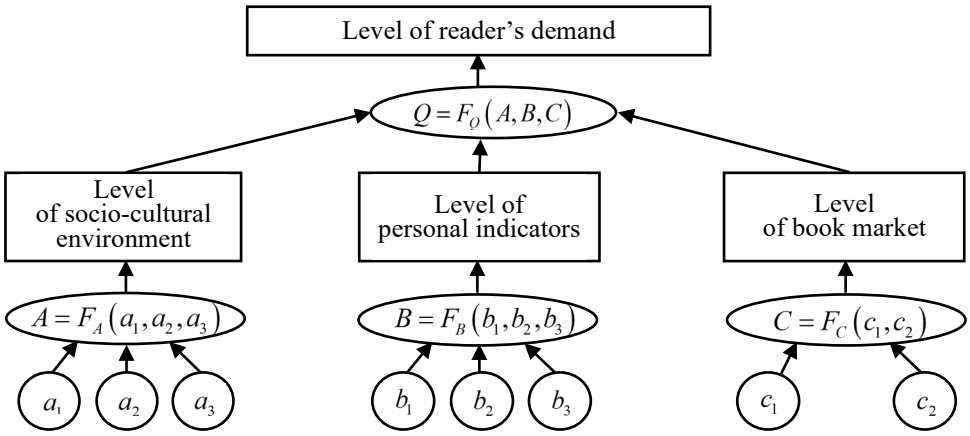

Figure 3: Multilevel model of fuzzy logical inference: the formation of an integrated indicator of the reader's demand level

The next step in this stage of the study will be the formation of a term-set of values of linguistic variables that determine the reader's demand level.

5. Experiment

Returning to the previous stage, a square inversely symmetric pairwise comparison matrix is constructed, taking into account the logic of connections in the semantic network, the order of which is determined by a number of factors. The known scale of relative importance of objects is used to establish the results of the expert comparison [2, 8].

Taking into consideration the above conditions, the elements of the matrix are represented in Table 1. From the previous table it is clear that the elements of the main diagonal of the matrix will be equal to one. The rest of the elements are obtained by comparing the factors of the first information column with the factors similar to the purpose of the first additional line.

Table 1

Pairwise comparison matrix of factors of demand for a book

Factors	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	1	6	7	5	3	4	8	9
x_2	1/6	1	2	1/3	1/5	1/5	5	6
x_3	1/7	1/2	1	1/5	1/6	1/6	3	5
x_4	1/5	3	5	1	1/4	1/3	5	7
x_5	1/3	5	6	4	1	3	7	8
x_6	1/4	5	6	3	1/3	1	7	8
x_7	1/8	1/5	1/3	1/5	1/7	1/7	1	2
x_8	1/9	1/6	1/5	1/7	1/8	1/8	1/2	1

To obtain the priority vector of the pairwise comparison matrix, the method described in [2, 8] is used. First, the main eigenvector $W(w_1, w_2, ..., w_n)$ of the matrix is defined, the components of which are obtained from the expression:

$$w_i = \sqrt[n]{a_{i1} \cdot a_{i2} \cdot ... \cdot a_{in}} \quad i = \overline{1,n}, \tag{3}$$

where n is a number of factors.

Normalized components of the vector W_{norm}

$$w_{i\,norm} = \frac{\sqrt[n]{a_{i1} \cdot a_{i2} \cdot ... \cdot a_{in}} \quad i = \overline{1,n}}{\sum_{i=1}^{n} \sqrt[n]{a_{i1} \cdot a_{i2} \cdot ... \cdot a_{in}}} \tag{4}$$

determine the previous numerical priorities of the factors.

One of the determining criteria for the reliability of the constructed matrix is the maximum eigenvalue $\lambda_{\max}$, which is used to calculate the subordinate criteria and is obtained as a result of such actions. The normalized vector W_{norm1} is calculated by multiplying the matrix on the right by the vector W_{norm}. Dividing the components of the vector W_{norm1} into the corresponding components of the vector W_{norm}, the vector W_{norm2} is obtained, whose components are correlated with the final levels of factors. The maximum eigenvalue of a positive inversely symmetric matrix $\lambda_{\max}$ is the arithmetic mean of components of the vector W_{norm2}. The assessment of the obtained solution is determined by the consistency index IU, which is calculated by the formula:

$$IU = (\lambda_{\max} - n)/(n-1). \tag{5}$$

The value of the consistency index is compared with a random index WI, which is considered a standard and depends on the number of objects. In this case, performing the inequality $IU < 0{,}1 \times WI$ determines the appropriate level of results.

The continuation of experimental research on the use of fuzzy logic is to further identify linguistic variables that correspond to the selected expert factors of demand for a book. Linguistic variables are accompanied by mathematical notations presented in the model in Fig. 2, and the additional semantic (linguistic) nature of the variable. A universal set of values and the corresponding linguistic terms are introduced, defined by a fuzzy scale that express the qualitative property of a variable. The elements of the universal set of values for variables, the limits of which are indefinite, are denoted by conventional units. The representation of the above characteristics will be presented in a table for convenience.

Table 2

Term-sets of values of linguistic variables

Variable	Linguistic nature of a variable	Universal set of values *D*	Linguistic terms (set *L*)
a_1	Place of residence (belonging of the settlement to a certain category by population)	(0–1500) thousand people	Small, average, large, more significant, the most significant
a_2	Family (role of the family)	(1–5) c.u.	Low (poorly developed values), average (formed basic values), high (formed social and spiritual values)
a_3	Reading traditions	(1–5) c.u.	Weak, average, strong
b_1	Level of education	(1–5) c.u.	Low (general secondary education: preschool, primary, basic, complete), below secondary (vocational education), secondary (professional pre-higher education), above secondary (higher education: short, first, second cycle), high (high education: third cycle, PhD)
b_2	Profession (occupation)	(1–5) c.u.	Gnostic, transforming, research
b_3	Social status (social class)	(1–5) c.u.	Lower, average, higher
c_1	Content (subject)	(1–5) c.u.	Reference and scientific editions, popular science and educational editions, literary and artistic editions
c_2	Availability and accessibility of literature	(1–5) c.u.	Custom editions, limited editions, mass distribution

The level of the reader's demand formation is denoted by a linguistic term Q. In this case, the universal set D is divided into parts (quanta). At the points of division, the linguistic variables and ranks $r_q(d_i)$ which identify linguistic terms are specified. Therefore, the output database will

be the set $D=\{d_1,d_2,...d_n\}$ and ranks $r_q(d_i)$ that set the priority of linguistic terms in the ranges d_i $(i=1,...,n)$. Taking into account the above, the linguistic term "level of reader's demand" Q is presented in the form of some fuzzy set, the elements of which form a set of pairs [2, 9, 10]:

$$Q=\left\{\frac{\mu_q(d_1)}{d_1},\frac{\mu_q(d_2)}{d_2},...,\frac{\mu_q(d_n)}{d_n}\right\}, \tag{6}$$

where: $Q\subset D$; $\mu_q(d_i)$ is a membership degree of the element $d_i\in D$ to the set Q.

Degrees or membership functions $\mu_q(d_i)$, are basic components of logical equations, the solution of which provides the numerical value of the membership function of the linguistic term Q. For membership functions, the rationing condition is satisfied: $\mu_1+\mu_2+...+\mu_n=1$.

The distribution of membership degrees (functions) meets the following conditions:

$$\frac{\mu_1}{r_1}=\frac{\mu_2}{r_2}=...=\frac{\mu_n}{r_n}, \tag{7}$$

where: $\mu_i=\mu_q(d_i)$; $r_i=r_q(d_i)$ for all $i=1,...,n$.

To graphically represent linguistic terms, the range of values of linguistic variables is divided into four parts, resulting in five points (d_1,d_2,d_3,d_4,d_5).

With known, or obtained on the basis of pairwise comparison matrices, ranks for each of the linguistic terms, the membership functions μ_i are calculated as a result of processing the matrix:

$$A=\begin{bmatrix} 1 & \frac{r_2}{r_1} & \frac{r_3}{r_1} & \frac{r_4}{r_1} & \frac{r_5}{r_1} \\ \frac{r_1}{r_2} & 1 & \frac{r_3}{r_2} & \frac{r_4}{r_2} & \frac{r_5}{r_2} \\ \dots & \dots & \dots & \dots & \dots \\ \frac{r_1}{r_5} & \frac{r_2}{r_5} & \frac{r_3}{r_5} & \frac{r_4}{r_5} & 1 \end{bmatrix}. \tag{8}$$

Obtaining the final result is to achieve the maximum value of the function, which characterizes the level of reader's demand at the maximum values of the membership functions of the assessment terms of factors – linguistic variables.

Let one move on to an important component of fuzzy logic – the formation of a fuzzy knowledge base, which according to the model of logical inference in Table 2 will look like:

IF (*A* = low) AND (*A* = average) AND (*A* = high)
AND (*B* = low) AND (*B* = average) AND (*B* = high)
AND (*C* = low) AND (*C* = average) AND (*C* = high),
THEN (*Q* = low) AND (*Q* = average) AND (*Q* = high).

On the basis of the formed conditions, a knowledge matrix is constructed:

Table 3

Knowledge matrix for linguistic variable *Q*

Level of socio-cultural environment *A*	Level of personal indicators *B*	Level of book market *C*	Level of reader's demand *Q*
low	low	low	low
low	average	low	

average	average	average	average
low	high	average	
high	average	high	high
high	high	high	

The knowledge matrices for the linguistic variable Q will correspond to fuzzy logical equations, which will define the procedures for obtaining the values of membership functions for the set of terms of the integral indicator of the level of reader's demand. For the terms "low", "average", "high", fuzzy logical equations are presented below:

$$\mu_{низький}(Q) = \mu_{low}(A) \wedge \mu_{low}(B) \wedge \mu_{low}(C) \vee \mu_{low}(A) \wedge \mu_{average}(B) \wedge \mu_{low}(C)$$

$$\mu_{average}(Q) = \mu_{average}(A) \wedge \mu_{average}(B) \wedge \mu_{average}(C) \vee \mu_{low}(A) \wedge \mu_{high}(B) \wedge \mu_{average}(C)$$

$$\mu_{high}(Q) = \mu_{high}(A) \wedge \mu_{average}(B) \wedge \mu_{high}(C) \vee \mu_{high}(A) \wedge \mu_{high}(B) \wedge \mu_{high}(C)$$

Based on expert statements about the sets $L(a_1, a_2, a_3)$, $L(b_1, b_2, b_3)$, $L(c_1, c_2)$, fuzzy knowledge bases, knowledge matrices and fuzzy logical equations for linguistic variables of the level of reader's demand are designed.

The generalized version of the logical statement for the linguistic variable "level of socio-cultural environment" and the knowledge matrix (Table 4) will look like this:

IF $(a_1) =$ (small, average, large, more significant, the most significant)

AND $(a_2) =$ (low, average, high)

AND $(a_3) =$ (weak, average, strong),

THEN $(A) =$ (low, average, high).

Table 4

Knowledge matrix for linguistic variable A (level of socio-cultural environment)

Place of residence a_1	Family (role of the family) a_2	Reading traditions a_3	Level of socio-cultural environment A
small	low	weak	low
small	average	weak	
average	average	average	average
large	average	average	
more significant	high	strong	high
the most significant	high	strong	

Fuzzy logical equations for the terms "low", "average", "high" are:

$$\mu_{low}(A) = \mu_{small}(a_1) \wedge \mu_{low}(a_2) \wedge \mu_{weak}(a_3) \vee \mu_{small}(a_1) \wedge \mu_{average}(a_2) \wedge \mu_{weak}(a_3)$$

$$\mu_{average}(A) = \mu_{average}(a_1) \wedge \mu_{average}(a_2) \wedge \mu_{average}(a_3) \vee \mu_{large}(a_1) \wedge \mu_{average}(a_2) \wedge \mu_{average}(a_3)$$

$$\mu_{high}(A) = \mu_{more\ significant}(a_1) \wedge \mu_{high}(a_2) \wedge \mu_{strong}(a_3) \vee \mu_{the\ most\ significant}(a_1) \wedge \mu_{high}(a_2) \wedge \mu_{strong}(a_3)$$

The logical statement for the linguistic variable "level of personal indicators" and the knowledge matrix (Table 5) will be presented the following way:

IF $(b_1) =$ (low, below average, average, above average, high)

AND $(b_2)=$ (gnostic, transforming, research)

AND $(b_3)=$ (weak, average, strong),

THEN $(B)=$ (lower, average, higher).

Table 5

Knowledge matrix for linguistic variable *B* (level of personal indicators)

Level of education b_1	Profession (occupation) b_2	Social status (social class) b_3	Level of personal indicators *B*
low	gnostic	lower	low
below average	gnostic	lower	
average	gnostic	average	average
average	transforming	average	
above average	transforming	average	high
high	research	higher	

Fuzzy logical equations for the terms “low”, “average”, “high” are:

$$\mu_{low}(B)=\mu_{low}(b_1)\wedge\mu_{gnostic}(b_2)\wedge\mu_{lower}(b_3)\vee\mu_{below\ average}(b_1)\wedge\mu_{gnostic}(b_2)\wedge\mu_{lower}(b_3)$$

$$\mu_{average}(B)=\mu_{average}(b_1)\wedge\mu_{gnostic}(b_2)\wedge\mu_{average}(b_3)\vee\mu_{average}(b_1)\wedge\mu_{transforming}(b_2)\wedge\mu_{average}(b_3)$$

$$\mu_{high}(B)=\mu_{above\ average}(b_1)\wedge\mu_{transforming}(b_2)\wedge\mu_{average}(b_3)\vee\mu_{high}(b_1)\wedge\mu_{research}(b_2)\wedge\mu_{higher}(b_3)$$

The logical statement and the knowledge matrix (Table 6) for the linguistic variable "level of book market " will look like this:

IF $(c_1)=$ (reference and scientific editions, popular science and educational editions, literary and artistic editions)

AND $(c_2)=$ (custom editions, limited editions, mass distribution),

THEN $(C)=$ (lower, average, higher).

Table 6

Knowledge matrix for linguistic variable *C* (level of book market)

Content (subject) c_1	Availability and accessibility of literature c_2	Level of book market *C*
reference and scientific editions	custom editions	low
popular science and educational editions	custom editions	
reference and scientific editions	limited editions	average
popular science and educational editions	limited editions	
popular science and educational editions	mass distribution	high
literary and artistic editions	mass distribution	

Fuzzy logical equations for the terms “low”, “average”, “high” are:

$$\mu_{low}(C)=\mu_{reference\ and\ scientific\ editions}(c_1)\wedge\mu_{custom\ editions}(c_2)\vee$$

$$\vee\mu_{popular\ science\ and\ educational\ editions}(c_1)\wedge\mu_{custom\ editions}(c_2)$$

$$\mu_{average}(C) = \mu_{reference\ and\ scientific\ editions}(c_1) \wedge \mu_{limited\ editions}(c_2) \vee$$
$$\vee \mu_{popular\ science\ and\ educational\ editions}(c_1) \wedge \mu_{limited\ editions}(c_2)$$
$$\mu_{high}(C) = \mu_{popular\ science\ and\ educational\ editions}(c_1) \wedge \mu_{mass\ distribution}(c_2) \vee$$
$$\vee \mu_{literary\ and\ artistic\ editions}(c_1) \wedge \mu_{mass\ distribution}(c_2)$$

The general fuzzy set of the linguistic variable Q for the analysed membership functions in relation to the fuzzy terms "low", "average", "high" and the corresponding values of the variable Q will look like:

$$Q(A,B,C) = \left\{ \frac{\mu_{low}(Q)}{k_1}, \frac{\mu_{average}(Q)}{k_2}, \frac{\mu_{high}(Q)}{k_3} \right\},$$

where k_1, k_2, k_3 are quantitative values of the variable Q in relation to the analysed terms.

6. Results

The calculation of numerical results can be performed according to the suggested and implemented above structuring of the research process in accordance with the theoretical and experimental principles of the selected stages.

As a result of processing the pairwise comparison matrix of factors and performing the calculations, a normalized vector is obtained:

$$W_{norm} = (0{,}354;\ 0{,}062;\ 0{,}042;\ 0{,}102;\ 0{,}238;\ 0{,}160;\ 0{,}023;\ 0{,}016),$$

components of which are transformed into integers for convenience of perception by multiplication by some scaling coefficient, for example, $k = 1000$. One will get:

$$W_{norm} \times k = (354;\ 62;\ 42;\ 102;\ 238;\ 160;\ 23;\ 16).$$

To assess the consistency of the weight priorities of the factors, the elements of the pairwise comparison matrix on the right are multiplied by the vector W_{norm}. The vector is obtained:

$$W_{norm1} = (3{,}225;\ 0{,}541;\ 0{,}367;\ 0{,}914;\ 2{,}121;\ 1{,}405;\ 0{,}204;\ 0{,}145).$$

Next, by dividing the components of the vector W_{norm1} into the corresponding components of the vector W_{norm}, the components of the eigenvector W_{norm2} .are found:

$$W_{norm2} = (9{,}092;\ 8{,}691;\ 8{,}725;\ 8{,}930\ 8{,}897;\ 8{,}769;\ 8{,}670;\ 9{,}008).$$

The arithmetic mean of the components of the vector W_{norm2} determines the maximum eigenvalue of the pairwise comparison matrix $\lambda_{\max} = 8{,}85$. The assessment of the obtained solution is determined by the consistency index $IU = 0{,}12$ calculated by formula (5).

The adequacy of the solution of the problem is confirmed under the condition of inequality $IU < 0{,}1 \times WI$ (where $WI = 1{,}41$ is the standard value of the random index for eight objects). The solution is acceptable because $0{,}12 < 0{,}1 \times 1{,}41$. Finally, the results are assessed by the consistency ratio, the value $WU = IU/WI$ of which must satisfy the ratio $WU \le 0{,}1$. For our variant $WU = 0{,}08$, which allows asserting the reliability of the results of pairwise comparisons according to the above criteria.

As a result, the factors declared at the beginning of the study form the following levels of influence priority on the process of the reader's demand formation: x_1 – place of residence; x_5 – availability (accessibility) of literature; x_6 – family (role of the family); x_4 – content, subject of

the book; x_2 – level of education; x_3 – profession (occupation); x_7 – reading traditions; x_8 – social status. The interpretation of the priorities of the factors of interest in the book is quite close (especially for the first two levels) to those obtained as a result of analytical and sociological surveys [15, 20, 21], performed using fundamentally different methods, which indicates the reliability of the first stage.

The study completes the calculation of the predicted numerical expression of the level of reader's demand using the mechanisms of fuzzy logic.

At the beginning, fuzzy sets for the terms of linguistic variables are formed according to the characteristics given in Table 2. Pairwise comparison matrices W are constructed for the linguistic variable "place of residence" (belonging of a settlement to a certain category by population) with a universal set of values $D(a_1)=[50;250;500;1000;1500]$ thousand people and the term-set of values $L(a_1)=$ <small, average, large, more significant, the most significant>. For the terms "small", "average", "large", "more significant" and "the most significant" matrices will look like this:

$$W_{small}(a_1)=\begin{bmatrix} 1 & 5/7 & 4/7 & 3/7 & 1/7 \\ 7/5 & 1 & 4/5 & 3/5 & 1/5 \\ 7/4 & 5/4 & 1 & 3/4 & 1/4 \\ 7/3 & 5/3 & 4/3 & 1 & 1/3 \\ 7 & 5 & 4 & 3 & 1 \end{bmatrix} \quad W_{average}(a_1)=\begin{bmatrix} 1 & 7/9 & 5/9 & 3/9 & 1/9 \\ 9/7 & 1 & 5/7 & 3/7 & 1/7 \\ 9/5 & 7/5 & 1 & 3/5 & 1/5 \\ 9/3 & 7/3 & 5/3 & 1 & 1/3 \\ 9 & 7 & 5 & 3 & 1 \end{bmatrix}$$

$$W_{large}(a_1)=\begin{bmatrix} 1 & 6 & 9 & 3 & 1 \\ 1/6 & 1 & 9/6 & 3/6 & 1/6 \\ 1/9 & 6/9 & 1 & 3/9 & 1/9 \\ 1/3 & 6/3 & 9/3 & 1 & 1/3 \\ 1 & 6 & 9 & 3 & 1 \end{bmatrix} \quad W_{more\ significant}(a_1)=\begin{bmatrix} 1 & 2 & 4 & 6 & 8 \\ 1/2 & 1 & 4/2 & 6/2 & 8/2 \\ 1/4 & 2/4 & 1 & 6/4 & 8/4 \\ 1/6 & 2/6 & 4/6 & 1 & 8/6 \\ 1/8 & 2/8 & 4/8 & 6/8 & 1 \end{bmatrix}$$

$$W_{the\ most\ significant}(a_1)=\begin{bmatrix} 1 & 4 & 6 & 8 & 9 \\ 1/4 & 1 & 6/4 & 8/4 & 9/4 \\ 1/6 & 4/6 & 1 & 8/6 & 9/6 \\ 1/8 & 4/8 & 6/8 & 1 & 9/8 \\ 1/9 & 4/9 & 6/9 & 8/9 & 1 \end{bmatrix}$$

After calculating the matrices, the values of the membership functions are obtained for the terms "small", "average", "large", "more significant" and "the most significant":

$$\mu_{small}(d_1)=0,35;\ \mu_{small}(d_2)=0,25;\ \mu_{small}(d_3)=0,2;\ \mu_{small}(d_4)=0,15;\ \mu_{small}(d_5)=0,05.$$

$$\mu_{average}(d_1)=0,36;\ \mu_{average}(d_2)=0,28;\ \mu_{average}(d_3)=0,2;\ \mu_{average}(d_4)=0,12;$$

$$\mu_{average}(d_5)=0,04.$$

$$\mu_{large}(d_1)=0,05;\ \mu_{large}(d_2)=0,3;\ \mu_{large}(d_3)=0,45;\ \mu_{large}(d_4)=0,15;\ \mu_{large}(d_5)=0,05.$$

$$\mu_{more\ significant}(d_1)=0,047;\ \mu_{more\ significant}(d_2)=0,095;\ \mu_{more\ significant}(d_3)=0,19;$$

$$\mu_{more\ significant}(d_4)=0,285;\ \mu_{more\ significant}(d_5)=0,38.$$

$$\mu_{the\ most\ significant}(d_1)=0,035;\ \mu_{the\ most\ significant}(d_2)=0,142;\ \mu_{the\ most\ significant}(d_3)=0,214;$$

$$\mu_{the\ most\ significant}(d_4)=0,285;\ \mu_{the\ most\ significant}(d_5)=0,321.$$

The values of membership functions are normalized with respect to one and the normalization coefficients are determined for linguistic terms:

$$k_e=1/\max\mu_e(d_i),\ \ (i=1,2,3),$$

where: e are the terms of the analyzed linguistic variable $\mu_{e_n}(d_i)=k_e\times\mu_e(d_i)$.

Normalized values of membership functions of the linguistic variable "place of residence" are:

$$\mu_{small_n}(d_1)=1;\ \mu_{small_n}(d_2)=0,714;\ \mu_{small_n}(d_3)=0,571;\ \mu_{small_n}(d_4)=0,429;\ \mu_{small_n}(d_5)=0,143.$$

$$\mu_{average_n}(d_1)=1;\ \mu_{average_n}(d_2)=0,779;\ \mu_{average_n}(d_3)=0,556;\ \mu_{average_n}(d_4)=0,333;$$

$$\mu_{average_n}(d_5)=0,111.$$

$$\mu_{large_n}(d_1)=0,111;\ \mu_{large_n}(d_2)=0,667;\ \mu_{large_n}(d_3)=1;\ \mu_{large_n}(d_4)=0,333;\ \mu_{large_n}(d_5)=0,111.$$

$$\mu_{more\ significant_n}(d_1)=0,124;\ \mu_{more\ significant_n}(d_2)=0,25;\ \mu_{more\ significant_n}(d_3)=0,5;$$

$$\mu_{more\ significant_n}(d_4)=0,75;\ \mu_{more\ significant_n}(d_5)=1.$$

$$\mu_{the\ most\ significant_n}(d_1)=0,109;\ \mu_{the\ most\ significant_n}(d_2)=0,442;\ \mu_{the\ most\ significant_n}(d_3)=0,667;$$

$$\mu_{the\ most\ significant_n}(d_4)=0,888;\ \mu_{the\ most\ significant_n}(d_5)=1.$$

The terms of the linguistic variable "place of residence" are recorded in fuzzy sets according to the expression (6) with a visual graphical representation.

$$small\ settlement=\left\{\frac{1}{50};\frac{0,714}{250};\frac{0,571}{500};\frac{0,429}{1000};\frac{0,143}{1500}\right\}\ \text{thousand people;}$$

$$average\ settlement=\left\{\frac{1}{50};\frac{0,779}{250};\frac{0,556}{500};\frac{0,333}{1000};\frac{0,111}{1500}\right\}\ \text{thousand people;}$$

$$large\ settlement=\left\{\frac{0,111}{50};\frac{0,667}{250};\frac{1}{500};\frac{0,333}{1000};\frac{0,111}{1500}\right\}\ \text{thousand people;}$$

$$more\ significant\ settlement=\left\{\frac{0,124}{50};\frac{0,25}{250};\frac{0,5}{500};\frac{0,75}{1000};\frac{1}{1500}\right\}\ \text{thousand people;}$$

$$the\ most\ significant\ settlement=\left\{\frac{0,109}{50};\frac{0,442}{250};\frac{0,667}{500};\frac{0,888}{1000};\frac{1}{1500}\right\}\ \text{thousand people}$$

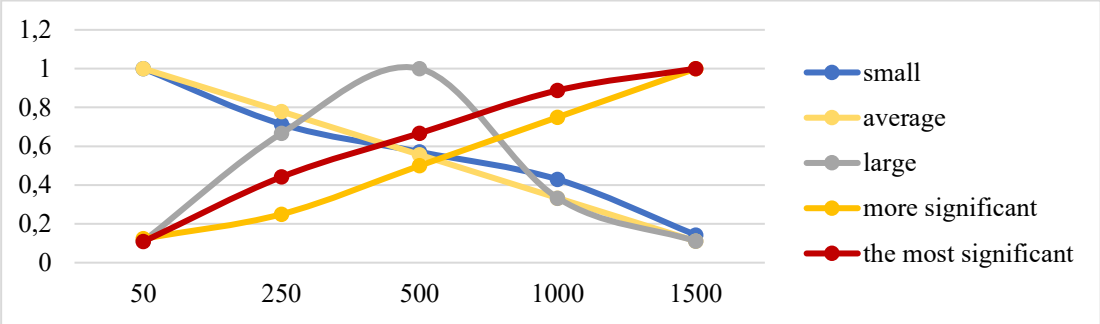

Figure 4: Membership functions of the linguistic variable "place of residence" (belonging of a settlement to a certain category by population)

Omitting similar calculations for the rest of the linguistic variables, one can move on to an important component of fuzzy logic – the process of defuzzification, which will provide reasonable receiving of the predicted numerical value of the reader's demand level.

The implementation of the process is performed by analogy with the option described above, taking into account the linguistic variable "place of residence". Preliminarily, a table with the normalised values of membership functions at the points of division of the universal set D is constructed.

Table 7

Membership functions of the term-set (place of residence – by population)

d_i, thousand people	50	250	500	1000	1500
$\mu_{small}(d_i)$	1	0,714	0,571	0,429	0,143
$\mu_{average}(d_i)$	1	0,779	0,556	0,333	0,111
$\mu_{large}(d_i)$	0,111	0,667	1	0,333	0,111
$\mu_{more\ significant}(d_i)$	0,124	0,25	0,5	0,75	1
$\mu_{the\ most\ significant}(d_i)$	0,109	0,442	0,667	0,888	1

Tables for other linguistic variables are formed similarly to the above.

The values of membership functions for the terms "low", "average", "high" are substituted into fuzzy logical equations for linguistic variables *A, B, C*.

$$\mu_{low}(A) = 0,571 \wedge 0,625 \wedge 0,665 \vee 0,571 \wedge 1 \wedge 0,665 = 0,571$$

$$\mu_{average}(A) = 0,556 \wedge 1 \wedge 1 \vee 1 \wedge 1 \wedge 1 = 1$$

$$\mu_{high}(A) = 0,5 \wedge 0,555 \wedge 0,556 \vee 0,667 \wedge 0,555 \wedge 0,556 = 0,555$$

$$\mu_{low}(B) = 0,428 \wedge 0,499 \wedge 0,428 \vee 0,625 \wedge 0,499 \wedge 0,428 = 0,428$$

$$\mu_{average}(B) = 1 \wedge 0,499 \wedge 1 \vee 1 \wedge 1 \wedge 1 = 1$$

$$\mu_{high}(B) = 0,556 \wedge 1 \wedge 1 \vee 0,555 \wedge 0,555 \wedge 0,713 = 0,556$$

$$\mu_{low}(C) = 0,555 \wedge 0,625 \vee 1 \wedge 0,625 = 0,625$$

$$\mu_{average}(C) = 0,555 \wedge 1 \vee 1 \wedge 1 = 1$$

$$\mu_{high}(C) = 1 \wedge 0,667 \vee 0,555 \wedge 0,667 = 0,667$$

For the highest level Q, the membership functions will receive the following numerical values:

$$\mu_{low}(Q) = 0,571 \wedge 0,428 \wedge 0,625 \vee 0,571 \wedge 1 \wedge 0,625 = 0,571$$

$$\mu_{average}(Q) = 1 \wedge 1 \wedge 1 \vee 0,571 \wedge 0,556 \wedge 1 = 1$$

$$\mu_{high}(Q) = 0,555 \wedge 1 \wedge 0,667 \vee 0,555 \wedge 0,556 \wedge 0,667 = 0,555$$

Based on the obtained data, the defuzzification of the fuzzy set is performed by the formula:

$$Q = \frac{\sum_{i=1}^{m}\left[\underline{Q} + (i-1)\frac{\overline{Q} - \underline{Q}}{m-1}\right]\mu_i(Q)}{\sum_{i=1}^{m}\mu_i(Q)}, \tag{9}$$

where: $\underline{Q}$ and $\overline{Q}$ – are the maximum and the minimum values of the quality indicator; m – is a number of fuzzy terms. The conditional limits for the variable Q are: $\underline{Q} = 1\%$, $\overline{Q} = 100\%$. The calculations are performed at three points of division: 1%, 50%, 100%. As a result of calculation, the numerical value of an integral indicator of the reader's demand level is obtained:

$$Q_{forecast} = \frac{1 \cdot 0,571 + 50 \cdot 1 + 100 \cdot 0,555}{0,571 + 1 + 0,555} = 49,89\ \%$$

The results of the study allow for more sound planning of publishing and book trade organizations due to the possibility of taking into account the priorities of factors of interest in books and final forecasting of reader's demand – the important components, the assessment of the importance of which is obtained on the basis of the application of methods of system analysis and fuzzy sets, taking into account the expert output data.

7. Conclusions

The analysis of the literary sources concerning the subject of the recommended paper is carried out. Despite the considerable interest in the issues raised in the work, the amount of research on the theoretical orientation of the processes of the reader's demand formation is clearly insufficient taking into account the dominance of modern media space with electronic products of "light consumption". In view of the above, it is considered appropriate to use methodologies of information orientation, significant developments of which in other areas of human activity in the form of theoretically balanced and successfully applied models, methods, software, which form the basis of modern information technology, would lead to significant progress in the interest in the book and the intensification of the reader's demand.

It is clear that the attraction to the book can be considered a subjective category of human nature, at the same time it is very important as to its social nature and the enormous impact on the intellectual level of society. It is characterized by the presence of certain factors that affect its intensity. Therefore, an important task is to identify, describe and structure the relationship between the factors, their ranking substantiated by levels of importance in relation to the impact on the book demand. The adequacy of this stage of the study is confirmed by the acceptable values of the criteria of the methods used. An important stage in the study of the process of interest in a book is the formulation and solution of problems of formalized assessment of the reader's demand, which involves the formation of term-sets of values of linguistic variables that identify factors of

interest in a book, the calculation of membership functions of variables, designing fuzzy knowledge bases and deriving and solving fuzzy logical equations. This has become the basis of the defuzzification process, the implementation of which provides a prognostic assessment of the reader's demand level.

Based on the obtained results, it is logical to state that the numerical value of the integrated indicator of the reader's demand level is inversely proportional to the risk of low demand for a book.

References

[1] V.M.Senkivskyy, I.V.Pikh, A.V.Kudriashova, N.M.Lytovchenko, Theoretical foundations of quality assurance of publishing and printing processes (Part 2: Synthesis of models of priority of factors action), Printing and Publishing, 2016, 1(71), pp 20-29

[2] I.V.Pikh, B.V.Durnyak, V.M.Senkivskyy, T.S.Holubnyk, Information technologies of quality formation of book editions, Monograph, Lviv, Ukrainian Academy of Printing, 2017, 308 p.

[3] H.Chen, Y.J.Hu, M.D.Smith, The impact of E-book distribution on print sales: analysis of a natural experiment, Management Science, 2019, 65(1), pp. 19-31

[4] K.L.Anderson,T. S.Atkinson, E.A.Swaggerty, K.O'Brien, Exploring the short-term impacts of a community-based book distribution program, Literacy Research and Instruction, 2019, 58(2), pp. 84-104

[5] G.Mytton, P.Diem, P.H.Van Dam, Media Audience Research: a guide for professionals, SAGE Publications India, 2016

[6] V.Matveiev, Semantic networks. URL: matveev.kiev/exprt/t5.pdf

[7] A.V.Kudriashova, Semantic network of factors of compositional design of the edition, Scientific Papers, Ukrainian Academy of Printing, 2016, 2(53), pp. 112-119

[8] E.Matrosova, A.Tikhomirova, N.Matrosov, K.Dmitriy, Visualization of T. Saati Hierarchy Analysis Method, In Biologically Inspired Cognitive Architectures Meeting, Springer, Cham, 2020, pp. 253-264

[9] N.Raychev, Hybrid system with fuzzy hierarchical evaluation model, Nature of Artificial Life, 2020, 1(1), pp. 10-37686

[10] P.Cintula, C.G.Fermüller, C.Noguera, Fuzzy logic, 2016

[11] P.Gupta, Applications of Fuzzy Logic in Daily life, International Journal of Advanced Research in computer science, 2017, 8(5).

[12] L.A.Zadeh, R.A.Aliev, Fuzzy logic theory and applications: part I and part II, World Scientific Publishing, 2018

[13] L.A.Zadeh, Fuzzy logic – a personal perspective, Fuzzy sets and systems, 2015, 281, pp. 4-20

[14] Y.Zhang, S.Kudva, E-books versus print books: Readers' choices and preferences across contexts, Journal of the Association for Information Science and Technology, 2014, 65(8), pp. 1695-1706

[15] B.Anderson, Influence of education, income and age on newspaper use and platform preference, Elon Journal of Undergraduate Research in Communications, 2018, 9(1), pp. 108-114

[16] G.Layefa, W.A. ohnson, A.Taiwo, Newspaper Readership Pattern in Ekiti State, Nigeria, IOSR Journal of Humanities and Social Science (IOSR-JHSS), 2016, 21(5), pp. 121-13

[17] M.B.Harji, K.Balakrishnan, K.Letchumanan, SPIRE Project: Parental Involvement in Young Children's ESL Reading Development, English Language Teaching, 2016, 9(12), pp. 1-15

[18] G.Ozturk, S.Hill, G.Yates, Family context and five-year-old children's attitudes toward literacy when they are learning to read, Reading Psychology, 2016, 37(3), pp. 487-509
[19] N.O.R.A.I.E.N.Mansor, Exploring perceptions on ESL students' reading habits, Journal of Business and Social Development, 2017, 5(2), pp. 19-24
[20] Q.Chen, Y.Kong, W.Gao, L.Mo, Effects of socioeconomic status, parent–child relationship, and learning motivation on reading ability, Frontiers in psychology, 2018, 9, 1297 p.
[21] E.Herrmann, Note from the Editor, How do we choose what we read? TXT, 2016(1), URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4804024
[22] M.Becker, N.McElvany, The interplay of gender and social background: A longitudinal study of interaction effects in reading attitudes and behavior, British Journal of Educational Psychology, 2018, 88(4), pp. 529-549