Conceptual Model of Presentation of Fuzzy Knowledge

Volodymyr Polishchuk[a], Miroslav Kelemen[b], Iwona Włoch[c], Andriy Polishchuk[a], Marianna Sharkadi[a] and Yurii Mlavets[a]

[a] *Uzhhorod National University, Narodna Square, 3, Uzhhorod, 88000, Ukraine*
[b] *Technical University of Kosice, Rampova 7, Kosice, 04121, Slovak republic*
[c] *Rzeszow University of Technology, al. Powstańców Warszawy, 12, Rzeszow, 35959, Poland*

Abstract
Is conducted a study of the current problem of developing a conceptual model of fuzzy knowledge, using one-dimensional and multidimensional membership functions, taking into account any type of input data, as well as designing innovative software as a means of technical support to systems analysts and its implementation in the learning process. For the first time, the study proposed a conceptual model consisting of five stages. Approaches to the formalization of qualitative and hybrid data are demonstrated by examples. The most common types and kinds of membership function of one variable which can identify a set of criteria of internal and external factors of influence for various applied problems are resulted. Modeling of uncertainties of different types on the basis of multidimensional membership functions is also investigated. An experiment was performed on the example of awarding fuzzy knowledge for the problematic of complex evaluation of startup projects. Innovative software in the form of a web application - "Information modeling of fuzzy knowledge", which will find application in systems analysts and implemented in the learning process. The study will be a useful tool to support decision-making to formalize and gain knowledge from different data sets, in different areas of application.

Keywords [1]
Fuzzy knowledge, fuzzy set, membership function, expert evaluation, decision-making.

CEUR-WS.org/Vol-3018/Paper_1.pdf

1. Introduction

The use of information technology in various spheres of public life is associated with the use of clear and vague knowledge bases, including the construction of expert and intelligent systems. Data-driven innovations are already bringing benefits to citizens, improving their quality of life, for example, through improved personalized medicine, new mobility, and living in a Smart City. This is all aimed at the European Green Course. An effective modeling tool in many cybernetics and artificial intelligence problems dealing with fuzzy is the fuzzy knowledge base, which is a collection of facts, linguistic variables, and corresponding membership functions that are a valuable source for describing fuzzy concepts, extracting data, and making heterogeneous decisions various fields of science, business, and production. Today, decision support systems that use the knowledge gained from experts are widespread. The problem of multicriteria evaluation of objects lies in the plane of selection tasks, which are an integral part of the tasks of decision support systems.

During the design and development of an intelligent system, knowledge undergoes a similar transformation of data – from more generalized sets to narrower, specific to a given subject area. When developing intelligent systems, knowledge about the specific subject area for which the system is being developed is rarely complete and reliable. The use of accurate methods does not allow to take into

II International Scientific Symposium «Intelligent Solutions» IntSol-2021, September 28–30, 2021, Kyiv-Uzhhorod, Ukraine
EMAIL: volodymyr.polishchuk@uzhnu.edu.ua (V. Polishchuk); miroslav.kelemen@tuke.sk (M. Kelemen); iwloch@prz.edu.pl (I. Włoch); andrii.polishchuk@uzhnu.edu.ua (A. Polishchuk), marianna.sharkadi@uzhnu.edu.ua (M. Sharkadi); yurii.mlavets@uzhnu.edu.ua (Y. Mlavets)
ORCID: 0000-0003-4586-1333 (V. Polishchuk); 0000-0001-7459-927X (M. Kelemen); 0000-0002-9969-0827 (I. Włoch); 0000-0003-3093-8393 (A. Polishchuk); 0000–0002–2045–5956 (M. Sharkadi); 0000-0002-1480-9017 (Y. Mlavets)
© 2020 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

account the verbal inaccuracy and subjectivity of expert information, which in turn imposes restrictions on the quality of knowledge for decision-making.

Based on the above, there is an urgent problem of building a knowledge management system. Knowledge management is a set of processes that control the creation, dissemination, processing and use of knowledge within the object of study.

During developing knowledge management systems, can be identified a number of the following processes [1].

1. Collection of data on the object of study and its activities.
2. Excavation as a process of knowledge allocation for expert systems.
3. Structuring - the process of forming the structure of the received information.
4. Formalization - structured information to describe data and knowledge.
5. Maintenance - adjustment of formalized data and knowledge.

Thus, for various applications, given any type of data, there is a need and relevance to present fuzzy knowledge from the use of fuzzy set tools to the construction of information models. Information modeling of the reflection of fuzzy knowledge creates the possibility of adequate evaluation of alternative solutions while increasing the degree of its validity.

The study of this problem is divided into three stages:

1. Research and development of information models for the presentation of fuzzy knowledge on the example of various applied problems;
2. Development of a conceptual model for the presentation of fuzzy knowledge, taking into account any type of data, and innovative software as a means of technical support for systems analysts and its implementation in the educational process;
3. Approbation of the model of representation of fuzzy knowledge on applied problems of various spheres of application and testing of the software.

In this context, the authors have already conducted the first stage of the study, namely: the development of information models for the presentation of fuzzy knowledge for management decisions in the functioning of socio-economic systems in conditions of uncertainty according to input expert assessments [1].

The aim of the current study is to progress a conceptual model for the presentation of fuzzy knowledge, using one-dimensional and multidimensional membership functions, taking into account any type of input data, as well as designing innovative software as a means of technical support for systems analysts and its implementation in the learning process.

2. Overview of Domestic and Foreign Research Studies

Actual approaches to the development of information technology require the use of information and logical modeling at the stage of problem statement and design, which significantly increases the efficiency of communication between the customer and contractors. To solve the problems of designing information technology in the automation of control problems of complex systems, the use of fuzzy modeling and control methodology has become widespread [2].

Recent scientific studies [1-2] indicate the need for information modeling of fuzzy knowledge, which will allow on the basis of fuzzy, incomplete, and especially expert information to obtain adequate and objective knowledge about the object of study. Extracting large data sets without their proper, and most importantly high-quality processing is meaningless.

As of today, there are a large number of models of knowledge representation in complex systems. Conventionally, all models can be divided into the following groups [1]: logical, productive, semantically network, frame, mathematical. In general, the creation of mathematical models is based on adequate information about the object, but sometimes inaccurate, because the construction of the model uses expert information that shows the substantive features of the studied object, also formulated in natural language. Then the description of the object is vague.

In [5-6], general ideas and advantages are outlined, on which current opinions on the use of fuzzy logic for decision support systems are based. In [7-8] the use of fuzzy logic in different areas of application is given, which allows to determine the optimal variables in the conditions of uncertainty of the input data. Also in [9] introduces fuzzy decision support systems in maritime practice; [10]

represents an intelligent system for determining customer loyalty and decision-making based on fuzzy logic; [11] investigated product ratings by fuzzy dialing, using their online reviews and other.

To summarize, membership functions can be constructed by one of two methods; expert knowledge or generation-oriented [12]. Construction on the basis of expert knowledge decreases accuracy because there are subjective opinions of people or conflicts between experts [13]. It is also a time-consuming path and not always available [14]. For example, the construction of membership functions using the parameters obtained from expert opinions also has risks in terms of accuracy, due to the bias of experts, as human semantics and knowledge are limited by experience and knowledge [15]. Therefore, it is not suitable for solving fuzzy inference systems that require the presence of membership functions adaptive constructed with data definition or systems that give fast and dynamic constructions of membership functions.

Therefore, most research has been based on a data-driven and trend-based approach, as it allows you to automatically build membership functions. One such approach to constructing membership functions is the Fuzzy C-means (FCM) clustering algorithm. It has found application in various tasks, such as classification [16] and forecasting [17]. FCM clusters basic data sets and their results are used to generate membership functions. He considers only Gaussian membership functions, which is his disadvantage [18-19]. It is known that they are not always suitable for various applications. For example, common triangular and trapezoidal membership functions used in many studies, such as fuzzy antenna positioning controller [15], crime prevention analysis [20], gesture monitoring [21], fuzzy cognitive map [22-23] and others.

Based on the above, argues and confirms the relevance of our study, the development of a conceptual model for the presentation of fuzzy knowledge using one-dimensional and multidimensional membership functions, taking into account any type of input data. The relevance of this study proves the need to process data and obtain knowledge from them in various objects of research, for further decision-making. In addition, the relevance of this study is confirmed by the European Data Strategy [24] for its implementation until 2030, in order to establish a leading model in the EU to follow the society endowed with data to improve the solution – in business and the public sector.

3. Conceptual model of representation of fuzzy knowledge

3.1. Formal problem statement and input data

Fuzzy knowledge in the general case can be described through linguistic variables. For illustration [1]: If $(a_1 \Xi_1 X_1 \Lambda_1 a_2 \Xi_2 X_2 \Lambda_2 \ldots a_l \Xi_l X_l \Lambda_l \ldots)$ then $\Omega_p B \Xi_j L$. For example: $\Xi =$ affiliation relationship = { Belong, Most likely belong, …, Not belong }; $\Omega =$ Follow Ratio = {Follow, Most likely follow, …, Not follow}; $\Lambda =$ Communication relationship = {AND/OR, Likely AND/OR, …}; a_i– elements of the term set of a linguistic variable; X_i– values of linguistic variables; B – linguistic variable; L – the value of the linguistic variable.

The main element in the construction of the model of representation of fuzzy knowledge is the value of linguistic variables, which are obtained through the model of the corresponding function of belonging to the fuzzy set.

Let's mark some set of research objects $X = \{x_1, x_2, \ldots, x_n\}$. Under the objects of research, we will consider a broad idea, which includes elements of the system, system processes, management decisions, or even a complex whole system that needs to be evaluated. Objects are evaluated on many indicators, factors of external and internal influences, as well as, if necessary, to organize according to some rule, while using information modeling of fuzzy knowledge. Without reducing the generality, you can consider one object, and when there are many objects, then they can be sorted by the obtained initial estimates. In decision theory, this class of problems is called multicriteria evaluation problems. Then we consider the approach of modeling the problem of multicriteria selection using the tools of fuzzy sets and construction of their membership functions.

A system set-theoretic model of the problem of representing fuzzy knowledge can be represented as follows:

$$\{X, K_{FK}, M_{FK}, Z_{FK} | Y\} \quad (1)$$

Where:

- X – research object or set of research objects;
- K_{FK} – information model of criteria (groups of criteria) for evaluating the properties of the object of study on various factors of influence, based on data sets (expert knowledge, quantitative data);
- M_{FK} – a model for constructing membership functions to evaluate the object of study;
- Z_{FK} – model of aggregation of initial estimates.

As a result, we obtain an aggregate initial estimate Y of the object of study X.

Consider the case where the set X is finite and the valid alternatives can be listed. Let's mark $K = \{(K_i, \mu(K_i)), i = 1,2,\dots,m\}$ fuzzy set of performance criteria, which are used to evaluate research objects x_j. $\mu(K_i)$ – assessment of one-dimensional or multidimensional function of belonging to the relevant criterion, built by information modeling of the representation of fuzzy knowledge. If some evaluation criterion is a multifaceted content of the object of study, then it is proposed to use multidimensional membership functions. Estimates of the values of membership functions by alternatives can be offered in the form of a table 1.

Table 1

Estimates of values of criteria by alternatives

Критерії	x_1	x_2	…	x_n
K_1	O_{11}	O_{12}	…	O_{1n}
K_2	O_{21}	O_{22}	…	O_{2n}
…	…	…	…	…
K_m	O_{m1}	O_{m2}	…	O_{mn}

Or decision matrices:

$$O = (O_{ij}), i = \overline{1,m}; j = \overline{1,n}. \quad (2)$$

Where O_{ij} – data sets, or in our interpretation of the estimate j alternatives by I criterion, $O \in [0;1]$.

We offer the following formal classification of data sets:

- quantitative – all data sets according to different criteria, the object of study are quantitative statistical or large data sets formed by technical means;
- quality (expert knowledge) – all data sets have a linguistic representation of some term set $T = \{T_1; T_2; \dots; T_l\}$. For example, they are assessed by an expert (group of experts) using a linguistic variable, based on experience and knowledge of the object of study X, analyze it, draw conclusions and make one linguistic assessment of each indicator K;
- hybrid – all datasets are evaluated in a hybrid way, based on the analysis of quantitative datasets, to which linguistic assessments are added, for example using the experience of experts.

For example, “Smart City Data Sources” analyzes and processes data sets from different sources, such as [25]:

- data on participation: from active citizens of the city, from actions of participation for residents and entrepreneurs in support of the concept of a smart city;
- data from private systems: from city companies, from the third sector (non-profit organizations, volunteers, public associations);
- survey data: data obtained by urban planners, scientists, students, etc.;
- additional data: data received from sensors, for example, about transport, air quality, water quality, soil quality, noise level in the city, etc.;
- open data: publicly available data for all citizens, data of the government and other governmental and non-governmental organizations.

Belonging functions for the presentation of fuzzy knowledge, according to the relevant criteria are selected taking into account the specific application problem. Examples of such tasks may be: assessing the solvency of enterprises; evaluation of investment projects; evaluation of startup projects; public health risk assessment; assessment of the quality of life of the region's residents; paratroopers landing safety assessment and others.

Figure 1 shows the steps of the method of presenting fuzzy knowledge for the formation of membership functions based on the data of the object of study, obtained experimentally.

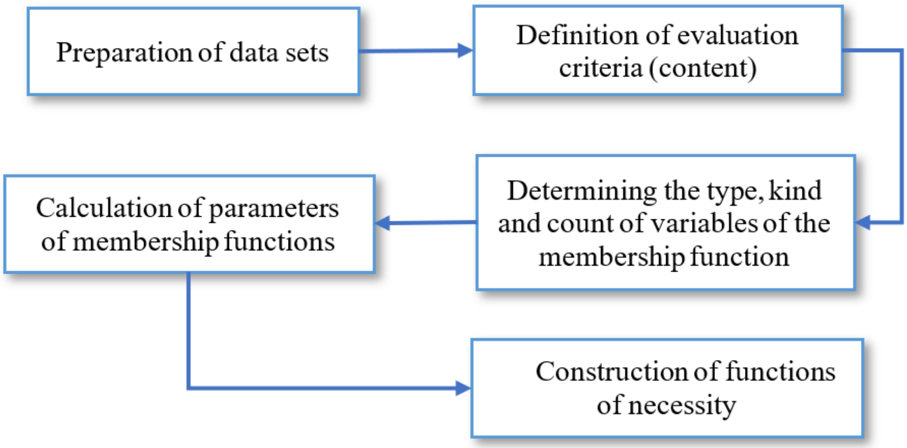

Figure 1: Stages of realization of constructions of functions of belonging

As can be seen from Figure 1, the first step is to prepare datasets specific to the universal set of membership functions: statistically large data sets (eg. collected from different sensors), expert (judgments about the relevant indicator for the object of study), or hybrid. The received information is analyzed, its content is evaluated and evaluation criteria are based on them. Next, the type, kind and count of variables of the membership function are determined. This is followed by the calculation of parameters and the construction of the membership function to represent fuzzy knowledge.

3.2. Conceptual model of presentation of fuzzy knowledge

The conceptual model of representation of fuzzy knowledge is given according to the following stages.

At the first stage, preparatory work is carried out on the sets, namely: the objects of research are identified, data are collected, processed, classified, etc.

As a result, we have a number of research objects $X = \{x_1, x_2, \ldots, x_n\}$, which need to be evaluated. We have a certain set of collected data on these objects O from which it is necessary to reveal the vagueness and gain knowledge, for further decision-making on the assessed objects.

In the second stage, the evaluation criteria (content) are determined.

Based on the collected data, information models of criteria (groups of criteria) for evaluating the properties of the object of study are built.

In the third stage it is necessary to determine the type, kind and count of variable membership functions.

For this purpose, first of all, data sets according to the offered classification are considered. If we have qualitative or hybrid datasets, then we first need to use some ratio to obtain quantitative estimates. To do this, can be specified certain correspondences between linguistic variables and quantitative, in the form of characteristic functions, or use the knowledge base of fuzzy inference, etc.

For example, if we have hybrid datasets then we can use the following approach. Let t_{ij} – variable from the term set T for i criterion j object of study; q_{ij} – quantitative assessment from the interval [0; 1], $i = \overline{1,m}$; $j = \overline{1,n}$.

To do this, each input value $(t_{ij}; q_{ij})$ it is necessary to match the value O_{ij}. Let be a term set of linguistic variables T showed on some numerical interval, to distinguish terms $[a_0; a_l]$, where $T_1 \in [a_0; a_1]$, $T_2 \in [a_1; a_2]$, …, $T_l \in [a_{l-1}; a_l]$. Interval values can be adjusted and changed by experts or other approaches. Next, we calculate the criteria O_{ij}, using linguistic variables T, quantitative estimates q and the value of the partitioning $[a_0; a_l]$, for example, using the following characteristic function:

$$O_{ij} = \begin{cases} a_1 \cdot q_{ij}, & if \quad t_{ij} \in T_1; \\ a_2 \cdot q_{ij}, & if \quad t_{ij} \in T_2; \\ \ldots & \ldots \quad \ldots; \\ a_l \cdot q_{ij}, & if \quad t_{ij} \in T_l. \end{cases} \tag{3}$$

Thus, we moved from hybrid estimates to quantitative ones. After that we can carry out modeling of fuzzy knowledge using membership functions.

Then it is necessary for each criterion K_i build membership functions $\mu(K_i), i = \overline{1, m}$ to evaluate the object of study. To do this, must be determined the type, kind and count of variables of the membership function.

The most common types and kinds of membership functions that can identify many criteria of internal and external factors of influence for different applications are below.

- Triangular membership functions are used for uncertainty problems of the type: "approximately equal to", "average value", "located in the interval", "similar to the object", "like the object", etc. These include triangular and trapezoidal membership functions.
- Z-shaped membership functions are used for uncertainty problems such as: "small count", "small value", "low level", etc. This type includes quadratic and harmonic Z-splines, Z-sigmoidal and Z-linear membership functions.
- S-shaped membership functions used for type uncertainty problems: "large count", "large value", "significant value", "high level", etc. There are quadratic and harmonic S-splines, S-sigmoidal and S-linear membership functions.
- π -shaped membership functions used for uncertainty problems such as "approximately in the range from and to", "approximately equal to", "approximately", etc. This type includes bell-shaped and Gaussian membership functions.

There are also other membership functions of fuzzy sets given as compositions of the above basic functions (double Gaussian, double sigmoidal, etc.), or as combinations by areas of growth and decrease (sigmoidal-Gaussian, spline-triangular, etc.).

There are problems when membership functions are used in reverse order. For example, if we have hybrid datasets for the criterion of security threat to network systems [26]. This indicator is assessed by a security expert on the basis of input expert data $(T_{ij}; q_{ij}), (i = \overline{1, m}; j = \overline{1, n})$. T – the consequences of the implementation of security threats to network systems, $T = \{M; \quad A; \quad H; \quad C\}$, where M – minimal consequences of the threat, A – average threat consequences, H – maximum consequences of the threat, C – critical consequences of the threat. q – the degree of possibility of realization of the threat, which is quantified from the interval [0; 1], putting the following content: 0 – it is impossible to realize the threat, 0.4 – minimal threat realization, 0.6 – average threat realization, 0.8 – high threat realization, 1 – critical realization of the threat. Terms $T = \{M; \quad A; \quad H; \quad C\}$ can be adequately determined on a percentage scale [0; 100], each of which is assigned a value from the interval $[a; b]$. For example, M – [0; 20], A – [20; 50], H – [50; 80], C – [80; 100]. This presentation has the following meaning. If the threat is realized, for example, by 90% then it is treated as a "critical consequence of the threat."

Dependence of the consequences of the threat T_{ij} and the degree of possibility of such implementation q_{ij} naturally to consider in the form S-shaped membership function. This is due to the fact that the higher the possibility of a threat and the more critical the consequence, the more dangerous the incident of network systems, which entails a high threat to the security of the system:

$$q_{ij} = \begin{cases} 0, & O_{ij} \leq a; \\ 2\left(\frac{O_{ij}-a}{b-a}\right)^2, & a < O_{ij} \leq \frac{a+b}{2}; \\ 1 - 2\left(\frac{b-O_{ij}}{b-a}\right)^2, & \frac{a+b}{2} < O_{ij} < b; \\ 1, & O_{ij} \geq b. \end{cases} \tag{4}$$

Thus, from formula (4) we express the dependence O_{ij}, since, the value of the degree of possibility of realization of the threat q_{ij} known and known intervals of numerical values for T_{ij}.

In many cases, it is impossible to do without using membership functions of only one variable. To model uncertainties of the form "approximately equal to" or "average value" in multidimensional space, it is proposed to use a cone-shaped or pyramidal membership function. These functions are a certain analogue of triangular membership functions [27-28]. The ambiguity of the type "located in the interval" is translated in the multidimensional case into the uncertainty "located in the region". This type of uncertainty is quite common when describing limitations in management tasks. Uncertainties of this type are revealed by means of a trapezoidal pyramidal membership function. An analog of Z-shaped and S-shaped functions in the multidimensional case can be a comprehensive sigmoidal function that can be used to model uncertainties at the extreme points of the area. Bell and Gaussian membership functions can also be generalized for use in n-dimensional space. These functions specify a smoother and nonlinear transition of membership values, which can be useful in many fuzzy control problems [28].

For example, the conical membership function is defined as follows [28]:

$$\mu(O_{ij}) = \begin{cases} 1 - \sqrt{\sum_{i=1}^{m} \frac{\left(O_{ij}-x_i^0\right)^2}{(h_i)^2}}, if \sqrt{\sum_{i=1}^{m} \frac{\left(O_{ij}-x_i^0\right)^2}{(h_i)^2}} < 1. \\ 0, otherwise. \end{cases} \quad (5)$$

where O_{ij} – rating *j* alternatives by *i* criterion (value *i* components of the vector of variables); x_i^0 – *i* the value of the center of the base of the cone; h_i – non-zero numerical parameters that specify scaling by vector coordinates $\bar{x}$; *m* – count of variables in the vector.

The center of the cone corresponds to a vector of values, which is approximately equal to the value being modeled. The cone-shaped function determines the uniform change of the membership index at a distance from the center of the cone according to the scaling parameters.

In the fourth stage, the parameters of the membership function are calculated.

This is a very important stage and is a difficult task, the solution of which depends on the adequacy of information modeling of fuzzy knowledge. Different mathematical or expert approaches can be used in different situations. Example:

- search for parameters using the least squares method or the Monte Carlo method;
- parameters can be built expertly on the basis of the available initial data set, with their future adjustment;
- parameters can be considered in the form of reference values;
- parameters can be built depending on the goals of the object of study;
- use of various measures of proximity, etc

The parameters of membership functions should be chosen by analyzing the available data sets by research objects. In addition, the parameters of the membership function make it possible to weed out unnecessary data classes.

In the last *fifth stage*, membership functions are built by formalizing fuzzy inputs. The result is a matrix of formalized data according to the constructed membership functions:

$$\mu(O) = \mu(O_{ij}), i = \overline{1,m}; j = \overline{1,n}. \quad (6)$$

After disclosure of uncertainty, processing and formalization of data sets using membership functions, it is necessary to obtain an aggregate initial estimate Y_j by researched objects x_j, $j = \overline{1,n}$. For this purpose, a model of aggregation of initial estimates using different convolutions is proposed [29].

The most common type of convolution is a linear convolution, which is defined as follows:

$$Y_j = \sum_{i=1}^{m} w_i \cdot \mu(O_{ij}),\ j = \overline{1,n}. \quad (7)$$

Coefficients $\{w_i\}$ are called scales and can serve to bring partial criteria to the same scale. Rationing of weights is sometimes required $\sum_{i=1}^{n} w_i = 1$.

The main advantages are the ease of convolution and maintaining the usual order for any system of weights $\{w_i\}$. The disadvantage is the possibility of compensation for small (unsatisfactory) values of some criteria due to good values of others. If the characteristics of objects are measured on a logarithmic scale (different scales), then it is advisable to use multiplicative convolutions [29]:

$$Y_j = \prod_{i=1}^{m} \left(\mu(O_{ij})\right)^{w_i}. \quad (8)$$

$$Y_j = \prod_{i=1}^{m} w_i \cdot \mu(O_{ij}). \quad (9)$$

Where $w_i > 0, j = \overline{1,n}$.

The advantages and disadvantages are similar to linear convolution.

The next common convolution is the Hermeier convolution proposed by Yu. Hermeyer, which is to assess the quality of the object at the worst value [30]:

$$Y_j = \min_i \left\{\frac{Y_j}{w_i}\right\}. \quad (10)$$

The advantage of this convolution is that it does not compensate for unsatisfactory values of some parameters at the expense of others. A significant disadvantage of this convolution is that it does not maintain the usual order.

The nonlinear scheme of compromises proposed by A. Voronin [31] can be written in the form:

$$Y_j = \sum_{i=1}^{n} \frac{w_i}{1-\mu(O_{ij})}, \ j = \overline{1,n}. \quad (11)$$

The main advantages of this type of convolution are that it behaves differently in different situations. For example, the evaluation of the object of study on some criteria is close to its limit value (unit), then this convolution expresses the minimax model. If it is far from its limit values, then this convolution is equivalent to the model of integral optimality. In the intermediate stages of the situation receive different degrees of partial alignment.

4. Results

Consider an example of presenting fuzzy knowledge for the task of comprehensive evaluation of startup projects. Let the set of startup projects be considered $S = \{s_1, s_2, \ldots, s_5\}$, which are evaluated by the following models [7-8]: M_1 – model for evaluating startup projects in conditions of information uncertainty; M_2 – model of information technology risk assessment of project financing; M_3 – information model for evaluation and ranking of startup project development teams.

Projects were evaluated using models, and as a result received generalized evaluations, table 2. The process of obtaining generalized evaluations is described in detail in Model of evaluation and selection of start-up projects by investor Goals.

Table 2
Input data for startup projects

Evaluation models	s_1	s_2	s_3	s_4	s_5
M_1	0,87	0,82	0,6	0,77	0,69
M_2	0,66	0,83	0,71	0,98	0,91
M_3	0,78	0,4	0,54	0,85	0,82

The task is to aggregate the knowledge about the object of study to build a ranking of startup projects and select the most promising for funding. To do this, we use the model to represent fuzzy knowledge based on multidimensional membership functions. For example, it is appropriate to use the cone-like membership function given by formula (5), where the value of the center of the base of the cone is given by a unit vector $(x_1^0; x_2^0; x_3^0) = (1; 1; 1)$, because our data are normalized; non-zero numerical parameters that specify scaling by vector coordinates $(h_1; h_2; h_3) = (1; 1; 1)$; count of variables 3; j – count of startup projects:

$$\mu(\overline{x_j}) = \begin{cases} 1 - a_j, & if\ a_j < 1, \\ 0, otherwise. \end{cases}$$

where: $a_j = \sqrt{\left(x_{1j} - 1\right)^2 + \left(x_{2j} - 1\right)^2 + \left(x_{3j} - 1\right)^2}, j = \overline{1,5}.$

Substituting the input data, we obtain the following result, table. 3.

Based on the obtained estimates, we can build a ranking of startup projects (s_4; s_5; s_1; s_3; s_2).

Table 3
The resulting assessment

s_1	s_2	s_3	s_4	s_5
0,575	0,29	0,325	0,725	0,63

In this study, the practical result is the development of innovative software in the form of a web application, which we called "Information modeling of fuzzy knowledge" [32]. The software is designed using the proposed conceptual model. This web application has the functionality to process fuzzy data using one-dimensional and multidimensional membership functions. Also, it is possible to calculate membership functions and plot using operations on one-dimensional membership functions, such as intersection, union, difference, symmetric difference, and disjunctive sum.

The developed software, as a tool for modeling fuzzy knowledge, is a useful tool for systems analysts, who not only has the opportunity to get rid of routine calculations, but also has the opportunity to visualize the processed data. In addition, the web application is planned to be implemented in the educational process of Uzhhorod National University as a means of supporting data processing for master's and PhD students. The main window is divided into two categories regarding the dimensionality of membership functions: "One variable" and "Multiple variables", fig. 2.

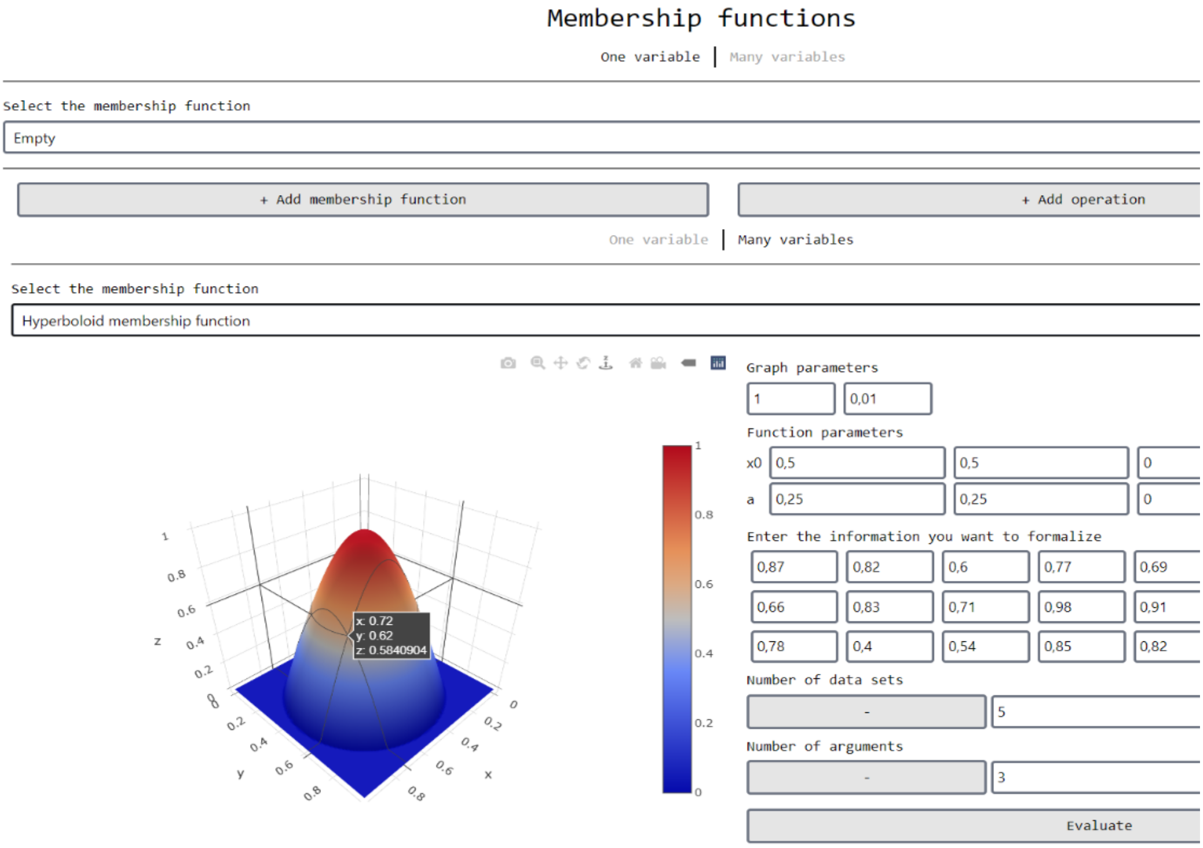

Figure 2: Web-application "Information modeling of fuzzy knowledge"

5. Discussion

It is undisclosed sufficiently to implement the tools of fuzzy sets to build models for mapping fuzzy knowledge. The apparatus of fuzzy sets requires the person making the decision to provide not point estimates, but with the help of an interval that determines the range of values of the relevant parameters. The advantage of such methods is the increase in the degree of validity of decisions. Because all possible scenarios are taken into account here, representing a continuous spectrum, on the contrary, for example on a discrete set of scenarios, is calculated by the Hurwitz method.

The parameters of the membership functions should be chosen based on the available data sets for a certain indicator of the object of study. Confirmation of this is an example of solving the problem of expert assessment of public health risks and quality of life of the Roma community. Based on data from the “Atlas of the Roma Community of the Slovak Republic 2019” [33], we have a situation where data, both qualitative and quantitative, indicators are in different evaluation scales (distance, number of people, percentage, presence or absence of the object, etc.). In addition, the notion of the quality of life of Roma families in the camp is very relative. To present it, it is necessary to compare data and normalize them against the performance of other camps in the evaluated region. We cannot use any "absolute indicator" because it will not give us any practical result of evaluation and comparison, because achieving an "absolute indicator" in the evaluated conditions is an abstract concept.

Building a conceptual model of fuzzy knowledge has a number of advantages, namely: accuracy, work with abstractions, transfer of information in a logically uniform way and increase the objectivity of expert and quantitative assessments, reveals the subjectivity of experts, quantifies informal tasks. The use of different models of membership functions, as well as the ability to determine the parameters, sometimes lead to ambiguity of the final results. This is a disadvantage of this approach.

6. Conclusions

Conducted a study of the current problem of developing a conceptual model of fuzzy knowledge, using one-dimensional and multidimensional membership functions, taking into account any type of input data, as well as was designed innovative software as a means of technical support for systems analysts and implementation in the learning process. In the first time the following results have been obtained:

- the conceptual model of representation of fuzzy knowledge on the basis of membership function of estimations on criteria, one and many variables, taking into account any type of input data, and their possibility of application for various applied problems is offered. The conceptual model consists of five stages. In the first stage, preparatory work is carried out on data sets, namely: the objects of research are identified, data are collected, processed, classified, etc. Next, the evaluation criteria (content) are determined. In the third stage, the type, kind and count of variable membership functions are determined. The approaches of formalization of qualitative and hybrid data on examples are demonstrated. The most common types and kinds of membership function of one variable which can identify set of criteria of internal and external factors of influence for various applied problems are resulted. Modeling of uncertainties of different types on the basis of multidimensional membership functions is also investigated. In the fourth stage, the parameters of membership functions are calculated and known approaches to their definition are described. On the last stage the one-dimensional or multidimensional membership function is constructed, fuzzy input data are formalized;
- the formal classification of data sets is given taking into account the type of input data, namely: quantitative, qualitative and hybrid. The known approaches of convolutions of receiving the aggregate initial estimation on objects of research are resulted;
- an experiment was performed on the example of displaying fuzzy knowledge for the task of evaluating startup projects in the complex. In this case, use the data obtained using: models for evaluating startup projects in conditions of information uncertainty; models of information technology risk assessment of project financing; information model for evaluating and rating teams of startup project developers;
- Innovative software in the form of a web application called "Information modeling of fuzzy knowledge" is designed on the basis of the developed conceptual model. This tool for modeling fuzzy knowledge, as a practical result of the study, will be used by systems analysts and implemented in the educational process of Uzhhorod National University for master's and PhD students.

The rationality of the conceptual model of representation of fuzzy knowledge proves the advantages of the developed model. The reliability of the obtained results is ensured by the correct use of fuzzy set theory and data formalization using one-dimensional and multidimensional membership functions, which is confirmed by the research results.

We see further research of the problem in the study of the third stage of this problem: testing of the model of representation of fuzzy knowledge on applied problems of different areas of application and testing of software.

7. Acknowledgements

This work was supported by the Slovak Research and Development Agency (APVV) under the contract no. PP-COVID-20-0002.

8. References

[1] O. Voloshyn, M. Malyar, V. Polishchuk, and M. Sharkadi. “Informatsiyne modelyuvannya nechitkykh znan”. [Information modeling of fuzzy knowledge]. Radioelektronika, informatyka, upravlinnya, 2018/4 (2018). (In Ukranian). doi: 10.15588/1607-3274-2018-4-8

[2] V. Polishchuk, M. Malyar, M. Kelemen and A. Polishchuk, The Technology for Determining the Level of Process Control in Complex Systems, in: Proceedings of VII International conference Information Technology and Interactions (IT&I), CEUR Workshop Proceedings, Kyiv, Ukraine, 2020, pp. 138-148. doi: http://ceur-ws.org/Vol-2845/Paper_14.pdf

[3] G.J. Klir, and B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, NJ, USA, Prentice Hall, Englewood Cliffs, 1995.

[4] M. Sugeno, G. T. Kang. “Structure identification of fuzzy model”. Fuzzy sets and systems, Vol. 28, no. 1 (1988): 15-33.

[5] C. Nugent. “Decision-making”. The Southwest Respiratory and Critical Care Chronicles, Vol. 7(30) (2019).

[6] O. Buchelnikov, A. Grishkov, F. Malchinskiy, L. Medyanikova, and S. Nekrasov. “Modelirovaniye sistemy otbora letchikov-istruktorov” [Modeling of the system of selection of pilots-instructors], Yuzhno-Rossiyskiy zhurnal sotsialnykh nauk, Vol. 19, №4, (2018). (In Russian).

[7] V. Polishchuk, M. Kelemen, B. Gavurová, C. Varotsos, R. Andoga, M. Gera, J. Christodoulakis, R. Soušek, J. Kozuba, P. Blišťan, S. and Szabo Jr. “A Fuzzy Model of Risk Assessment for Environmental Start-up Projects in the Air Transport Sector”. Int. J. Environ. Res. Public Health, Vol. 16, 3573 (2019). doi: 10.3390/ijerph16193573.

[8] M. Kelemen, V. Polishchuk, B. Gavurová, S. Szabo, R. Rozenberg, M. Gera, J. Kozuba, J. Hospodka, R. Andoga, A. Divoková, P. Blišt’an. “Fuzzy Model for Quantitative Assessment of Environmental Start-up Projects in Air Transport”. Int. J. Environ. Res. Public Health, Vol. 16, 3585 (2019). doi: 10.3390/ijerph16193585

[9] V. Suong Nguyen. “Investigation on a Novel Support System for Automatic Ship Berthing in Marine Practice”. Journal of Marine Science and Engineering, Vol. 7(4), 114 (2019). doi: 10.3390/jmse7040114.

[10] U. Ghani, IS. Bajwa, and A. Ashfaq. “A Fuzzy Logic Based Intelligent System for Measuring Customer Loyalty and Decision Making”. Symmetry. Vol. 10(12), (2018): 761.

[11] Y Liu, J.W. Bi, and Z.P. Fan. “Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory”. Inf. Fusion, Vol. 36, (2017): 149–161. doi: https://doi.org/10.1016/j.inffus.2016.11.012

[12] M.H. Hasan, J. Jaafar and M.F. Hassan, Fuzzy C-Means and two clusters’ centers method for generating interval type-2 membership function, in: Proceedings of the 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 2016, pp. 627–632.

[13] J.C. Dunn. “A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters”. J. Cybern., 3, (1973): 32–57.

[14] J.-R. Jang. “ANFIS: Adaptive-network-based fuzzy inference system”. IEEE Trans. Syst. Man, Cybern. 23, (1993): 665–685,
[15] V. Kalist, P. Ganesan, B.S. Sathish, J.M.M. Jenitha, K. Bashashaik. “Possiblistic-Fuzzy C-Means Clustering Approach for the Segmentation of Satellite Images in HSL Color Space”. Procedia Comput. Sci. 57, (2015): 49–56.
[16] B. Durnyak, M. Lutskiv, P. Shepita, R. Karpyn, N. Savina. Determination of the optical density of two-parameter tone transfer for a short printing system of the sixth dimension, in: Proceedings of the 2nd International Workshop on Intelligent Information Technologies & Systems of Information Security, CEUR Workshop Proceedings, Khmelnytskyi, Ukraine, 2021, pp. 134-140. doi: http://ceur-ws.org/Vol-2853/short11.pdf.
[17] S. Babichev, O. Khamula, B. Durnyak, J. Škvor. Technique of gene expression profiles selection based on SOTA clustering algorithm using statistical criteria and shannon entropy, in: Advances in intelligent systems and computing, Springer International Publishing, 2020, pp. 23-38. doi: 10.1007/978-3-030-54215-3_2.
[18] O. Castillo and P. Melin, Design of Intelligent Systems with Interval Type-2 Fuzzy Logic. In Handbook of Granular Computing. Berlin, Germany, Springer, 2008, pp. 575–601.
[19] R. Krishnapuram and J.M. Keller. “A possibilistic approach to clustering”. IEEE Trans. Fuzzy Syst. 1, (1993): 98–110.
[20] L. Wang and J. Wang. “Feature Weighting Fuzzy Clustering Integrating Rough Sets and Shadowed Sets”. Int. J. Pattern Recognit. Artif. Intell. 26, (2012): 26.
[21] S. Saha, M. Pal and A. Konar. “Triangular membership function based real-time gesture monitoring system for physical disorder detection”. Comput. Vis. Sci. 22, (2019): 1–14.
[22] G. Felix, G. Nápoles, R. Falcon, W. Froelich, K. Vanhoof, R. Bello. “A review on methods and software for fuzzy cognitive maps”. Artif. Intell. Rev. 52, (2019): 1707–1737.
[23] F. Liu, Y. Peng, Z. Chen, Y. Shi. “Modeling of Characteristics on Artificial Intelligence IQ Test: A Fuzzy Cognitive Map-Based Dynamic Scenario Analysis”. Int. J. Comput. Commun. Control. 14, (2019): 653–669.
[24] A European strategy for data, 2020. URL: https://ec.europa.eu/info/sites/default/files/communication-european-strategy-data-19feb2020_en.pdf
[25] FUTURISTIQ: Places, data, perspectives: The future of city planning in the age of big data, 2019. URL: https://www.apums.sk/e_download.php?file=data/editor/211sk_1.pdf&original=FUTURISTIQ_Budúcnosť%20plánovania%20mesta.pdf
[26] M. Kelemen, V. Polishchuk, B. Gavurová, R. Andoga, S. Szabo, W. Yang, J. Christodoulakis, M. Gera, J. Kozuba, P. Kaľavský, and M. Antoško. “Educational Model for Evaluation of Airport NIS Security for Safe and Sustainable Air Transport”. Sustainability, Vol. 12, (2020): 6352. doi: 10.3390/su12166352.
[27] A. Kumar, S. Yoon. “Development of Fast and Soft Landing System for Quadcopter Drone using Fuzzy Logic Technology”. International Journal of Advanced Trends in Computer Science and Engineering, Vol. 9, No.1 (2020).
[28] M. Daňková. “Functional Partial Fuzzy Relations”. Mathematics, Vol. 9, (2021): 2191. doi: 10.3390/math9182191.
[29] S. Thukral, J. S. Bal. “Design of an intelligent fuzzy logic predictive system for celiac disease assessment”. Acta Technica Corviniensis-Bulletin of Engineering, Vol. 13(4), (2020): 137-141.
[30] K.-S. Chen, T.-H. Huang. “A Fuzzy Evaluation Model Aimed at Smaller-the-Better-Type Quality Characteristics”. Mathematics, Vol. 9, (2021): 2513. doi: 10.3390/math9192513.
[31] B. Savun-Hekimoğlu, B. Erbay, M. Hekimoğlu, S. Burak. “Evaluation of water supply alternatives for Istanbul using forecasting and multi-criteria decision making methods”. Journal of Cleaner Production, Vol. 287, (2021): 125080.
[32] Information modeling of fuzzy knowledge, 2021. URL: https://z4cc76024-gtw.qovery.io/?test=1
[33] Atlas of Roma Communities, 2019. URL: https://www.minv.sk/?atlas-romskych-komunit-2019