

Data Stream Mining & Processing

PROCEEDINGS of the
2018 IEEE Second International Conference on
Data Stream Mining & Processing (DSMP)



IEEE 
UKRAINE SECTION

IEEE Ukraine Section (Kharkiv)
SP/AP/C/EMC/COM
Societies Joint Chapter

IEEE Ukraine Section (West)
AP/ED/MTT/CPMT/SSC
Societies Joint Chapter

August 21–25, 2018

Lviv, Ukraine



Table of Contents

Topic #1. Big Data & Data Science Using Intelligent Approaches	1
Iryna Perova, Olena Litovchenko, Yevgeniy Bodyanskiy, Yelizaveta Brazhnykova, Igor Zavgorodnii and Pavlo Mulesa. MEDICAL DATA-STREAM MINING IN THE AREA OF ELECTROMAGNETIC RADIATION AND LOW TEMPERATURE INFLUENCE ON BIOLOGICAL OBJECTS	3
Polina Zhernova, Anastasiia Deineko, Yevgeniy Bodyanskiy and Vladyslav Riepin. ADAPTIVE KERNEL DATA STREAMS CLUSTERING BASED ON NEURAL NETWORKS ENSEMBLES IN CONDITIONS OF UNCERTAINTY ABOUT AMOUNT AND SHAPES OF CLUSTERS	7
Ganna Ponomaryova, Igor Nevlydov, Oleksandr Filipenko and Mariya Volkova. MEMS-BASED INERTIAL SENSOR SIGNALS AND MACHINE LEARNING METHODS FOR CLASSIFYING ROBOT MOTION.	13
Dmytro Lande, Valentyna Andrushchenko and Iryna Balagura. DATA SCIENCE IN OPEN-ACCESS RESEARCH ON-LINE RESOURCES	17
Nina Khairova, Svitlana Petrasova and Włodzimierz Lewoniewski. BUILDING THE SEMANTIC SIMILARITY MODEL FOR SOCIAL NETWORK DATA STREAMS	21
Gautam Pal, Gangmin Li and Katie Atkinson. BIG DATA REAL TIME INGESTION AND MACHINE LEARNING	25
Andrii Berko and Vladyslav Aliksieiev. A METHOD TO SOLVE UNCERTAINTY PROBLEM FOR BIG DATA SOURCES.	32
Yuriy Kondratenko and Nina Kondratenko. COMPUTATIONAL LIBRARY OF THE DIRECT ANALYTIC MODELS FOR REAL-TIME FUZZY INFORMATION PROCESSING	38
Oleksandr Gerasin, Yuriy Zaporozhets and Yuriy Kondratenko. MODELS OF MAGNETIC DRIVER INTERACTION WITH FERROMAGNETIC SURFACE AND GEOMETRIC DATA COMPUTING FOR CLAMPING FORCE LOCALIZATION PATCHES	44
Volodymyr Ostakhov, Viktor Morozov and Nadiia Artykulna. MODELS OF IT PROJECTS KPIS AND METRICS	50
Yuliya Kozina, Natalya Volkova and Daniil Horpenko. MOBILE APPLICATION FOR DECISION SUPPORT IN MULTI-CRITERIA PROBLEMS	56
Olena Basalkevych and Olexandr Basalkevych. FUZZY RECONSTRUCTIONS IN LINGUISTICS	60
Mykola Malyar, Oleksey Voloshyn, Volodymyr Polishchuk and Marianna Sharkadi. FUZZY MATHEMATICAL MODELING FINANCIAL RISKS	65
Peter Bidyuk, Aleksandr Gozhyj, Iryna Kalinina, Zdislaw Szymanski and Volodymyr Beglytsia. THE METHODS BAYESIAN ANALYSIS OF THE THRESHOLD STOCHASTIC VOLATILITY MODEL	70
Max Garkavtsev, Natalia Lamonova and Alexander Gostev. CHOSING A PROGRAMMING LANGUAGE FOR A NEW PROJECT FROM A CODE QUALITY PERSPECTIVE	75

Fuzzy Mathematical Modeling Financial Risks

Oleksey Voloshyn
Department of complex systems modelling
Taras Shevchenko National University
Kyiv, Ukraine
olvoloshyn@ukr.net

Marianna Sharkadi
Department of cybernetics and applied mathematics
Uzhgorod National University
Uzhgorod, Ukraine
marianna.sharkadi@uzhnu.edu.ua

Mykola Malyar
Department of cybernetics and applied mathematics
Uzhgorod National University
Uzhgorod, Ukraine
malyarmm@gmail.com

Volodymyr Polishchuk
Department of Software Systems
Uzhgorod National University
Uzhgorod, Ukraine
v.polishchuk87@gmail.com

Abstract — The research of the urgent task of developing a fuzzy mathematical model of financial risks for evaluating projects regarding the level of security of their financing has been carried out. The development of such technology will provide an opportunity to adequately approach the consideration of projects, increase the degree of validity of investment decisions and increase economic security.

Keywords — projects, risks, linguistic evaluation, information technology, multicriteria, security financing, venture capital funds.

I. INTRODUCTION

Fuzzy mathematical modeling is one of the most active and promising areas of applied research in the field of management and decision making in weakly structured systems. The range of fuzzy methods is expanding every year, embracing various new areas. Fuzzy mathematical modeling is this when the elements of the study are not numbers, and some fuzzy sets or their combination. The basis of this approach lies not in traditional logic, but in logic with fuzzy truth, fuzzy ties and fuzzy rules of output. The main characteristics of this approach are the use of linguistic variables instead of numerical variables, the relationship between variables is described by fuzzy statements, and complex relationships are described by fuzzy algorithms.

When designing and managing a complex socio-economic system, a problem arises when a person is not able to give accurate and, at the same time, practical meanings of judgments about their behavior.

The paper proposes research of the actual problem of developing a mathematical model of information technology for risk assessment of projects regarding the level of security of their financing, using fuzzy mathematics, for various investment subjects. The development of such technology will provide an opportunity to adequately approach the consideration of projects, increase the degree of validity of investment decisions and, in general, increase economic and managerial security.

Financing projects of any nature (project startup or classical investment) is a risky activity. Depending on the origin of the project, there are various options for its financing, such as business angels, venture and investment funds, banks. Each of these institutions has its own risk management policy. But all of them combine one thing: to

find and finance a successful project with minimal risks.

Risk is closely linked to the concept of economic security of the project, both as the security of the entity representing the project, and the security of the investor. The subject's security is that a risky and unsuccessful project will lead to damage to the enterprise. The investor's security is directly dependent on an adequate assessment of the project and the entity presenting the project. Increasing the security of investment projects provides stability of the regional economy [1].

Recent scientific studies indicate the need to systematize risk minimization tools and develop a algorithm for selecting a model for evaluating projects of different origins. The issue of quantitative risk assessment and risk management during investing is disclosed in many papers [1-3], but a holistic concept for determining the level of risk, reducing it and taking into account the subjective aspects of the assessment has not yet been developed.

II. FORMULATION OF THE PROBLEM

Depending on the origin, commercial projects are dealt with in two types: classical - investment projects under which a well-formulated business plan emerges in a company operating on the market and requires a partial involvement of funds from the outside; startup projects - the "idea" that arises in companies whose business is based on innovative technologies, such companies have not entered the market or have just started to come out of it and need to attract external resources.

We formulate the task of evaluation as follows. Suppose we have some projects S_1, S_2, \dots, S_n , for which an evaluation of the risk with regard to the level of security of their financing should be done. Projects can have different perspectives, nature and security of implementation. Without diminishing the universality, we will continue to consider one project. In case of a plurality of projects, they can be ordered according to the initial estimates received. The model of the problem is represented in the following form:

$$SPF = O(O_S, O_G, O_R), \quad (1)$$

where O_S – evaluation of the project under consideration, depending on its origin (classic investment project [4] or

startup project [5]), O_G – assessment of the economy in which a commercial project will be implemented [6], O_R – aggregated risk assessment for project implementation. SPF – initial assessment and linguistic treatment of risk in relation to the level of project financing security. O – operator that matches the output variable SPF , with input estimates O_S, O_G, O_R .

III. MATERIALS AND METHODS

Let us offer the next set of start-up risk assessment criteria on which the platform can assess risks of start-ups. There are four groups of criteria: K_O – "operational risk"; K_I – "investment risks"; K_F – "financial risks"; K_S – "risks of innovation".

A person who decides (Decision Maker) from each group chooses risk criteria are criteria that can assess the proposed project. Let us represent each group of criteria in the form of a set of indicators. Then to the group of criteria K_O – "operational risk" the following indicators can be considered: K_{O1} – the risk of loss of the client base; K_{O2} – the risk of loss of the supplier; K_{O3} – the risk of losing market share; K_{O4} – the risk of lowering the level of management; K_{O5} – the risk of industrial conflict and ineffective motivation; K_{O6} – the risk of lowering the quality of the processes; K_{O7} – the risk of lowering labor productivity; K_{O8} – personnel risks; K_{O9} – the risk of unsecured resources.

The "investment risks" group – K_I we will express through the following indicators: K_{I1} – the risk of inefficiency of investments; K_{I2} – risk of disruption of the terms of creation of production assets; K_{I3} – the risk of failure to achieve the return on investment capital; K_{I4} – the risk of exceeding the amount of start-up investment; K_{I5} – the risk of a lack of investment capital.

The "financial risks" group – K_F we will express through the following indicators: K_{F1} – the risk of inefficient use of capital; K_{F2} – risk of loss (arises due to price changes, when sudden expenses cover revenue); K_{F3} – the risk of investor loss; K_{F4} – the risk of loss of solvency; K_{F5} – the risk of a suboptimal capital price.

The criteria of the "risks of innovation" group – K_S we will express with the help of such indicators: K_{S1} – the risk of ineffective innovation investments; K_{S2} – the risk of ineffective promotion of innovations; K_{S3} – risks of breaking the terms of innovation development; K_{S4} – risks of technology innovation; K_{S5} – risks of resource insufficiency when designing innovations.

This set of risk criteria can not reveal all aspects of any startup of the project in various areas of implementation, so

it is open and any expert can add some criteria depending on the scope of investment.

Each risk criteria is evaluated by experts with one of the terms of the following term-set of linguistic variables $L = \{H; HC; C; BC; B\}$, where: H – «low risk level»; HC – «risk level below average»; C – «average risk level»; BC – «risk level above average»; B – «high risk level». Also, an expert puts the number of «authenticity» for each assessment of risk level $\mu(L)$ of his consideration concerning the interval $[0; 1]$.

We describe a two-tier scheme of a project risk assessment model based on input linguistic variables. The inputs are presented in the form of linguistic variables and the reliability of the expert's consideration of their assignment. Therefore, at the first level, it is necessary to build membership rules and knowledge base in order to obtain the resulting term-assessment L^α for each group of risk criteria. On the basis of obtained resultant term evaluation L^α to determine the aggregated estimation of reliability $\mu(L^\alpha)$. At the second level, estimates are obtained L^α and $\mu(L^\alpha)$ we will design a "risk axis" to determine one project risk assessment for each group of criteria α .

Consider the first level - the construction of the rules of ownership of the resulting term evaluation by the groups of risk criteria.

Level H: «low risk level». The minimum amount of criteria with low risk level term should not be less than 60% and the remaining 40% of the criteria should not have terms lower than «risk level below average».

Level HC: «risk level below average». The project should have the minimal amount of criteria with the term «risk level below average» not less than 60%, and the other 40% of criteria should have terms not lower than the «average risk level».

Level C: «average risk level». The minimal amount of criteria with the term «average risk level» not less than 60%, and the other 40% of criteria should have terms not lower than the «risk level above average».

Level BC: «risk level above average». The minimal amount of criteria with the term «risk level above average» not less than 60%, and the other 40% of criteria should have terms not lower than the «high risk level».

Level B: «high risk level». The project gets the resulting term-evaluation «B» in case the amount of criteria with the term «high risk level» compiles 60% and more.

Then, based on the established rules of ownership of the resulting term evaluation for the groups of risk criteria, we can give a fragment of the knowledge base, for example, according to five criteria, Table 1.

Because the expert puts each variable L_i^α the reliability of their reasoning – $\mu(L_i^\alpha)$ from the interval $[0; 1]$, $\alpha = \{O; I; F; S\}$ then linguistic variables can be represented,

for example, in the form of triangular membership functions.

TABLE I. A FRAGMENT OF KNOWLEDGE BASE

N _o	K _{α1}	K _{α2}	K _{α3}	K _{α4}	K _{α5}	Resulting term evaluation
1	H	H	H	HC	HC	H
2	H	H	HC	HC	HC	HC
3	HC	HC	HC	C	C	
4	HC	HC	HC	H	C	
5	C	C	C	HC	HC	C
6	C	C	C	BC	BC	
7	C	C	C	HC	BC	
8	C	C	C	H	HC	
9	BC	BC	BC	C	C	BC
10	BC	BC	BC	C	B	
11	BC	BC	BC	B	B	
12	BC	BC	BC	HC	C	
13	B	B	B	BC	BC	B
14	B	B	B	BC	C	

The aggregated authenticity estimation $\mu(L^\alpha)$ $\alpha = \{O; I; F; S\}$ is calculated with the following formula:

$$\mu(L^\alpha) = \frac{1}{n} \sum_{i=1}^m \mu(L_i^\alpha), \quad \alpha = \{O; I; F; S\}, \quad (2)$$

Where $\mu(L_i^\alpha)$ – is the authenticity estimation of the linguistic variables which match the resulting term-evaluation for i -criterion of α risk criteria group.

At the second level, we will design the data on the risk criteria groups into a "axis of risk" to determine a generalized risk assessment of the project for each group of criteria α and obtaining an aggregated risk assessment, as well as its linguistic interpretation.

Next for each group of criteria α , let's express it x^α :

$$x^\alpha = \begin{cases} \sqrt{\frac{\mu(L^\alpha)}{2}}(b-a) + a, & 0 \leq \mu(L^\alpha) \leq 0,5; \\ b - \sqrt{\frac{1-\mu(L^\alpha)}{2}}(b-a), & 0,5 < \mu(L^\alpha) \leq 1. \end{cases} \quad (3)$$

To obtain a generalized project risk assessment for groups of criteria α , use the following formula:

$$z^\alpha = \frac{x^\alpha}{100}. \quad (4)$$

Three variables $(x(L^\alpha); \mu(L^\alpha); z^\alpha)$ we interpret the three-dimensional coordinate system, where $x = x(L^\alpha)$ – the value of a function equal to the numerical interpretation of the resulting term-estimates $L = \{H; HC; C; BC; B\}$ $y = \mu(L^\alpha)$ – aggregated assessment of the reliability of the expert's thoughts, and the axis $z = z^\alpha$ – project risk assessment for each group of criteria α , which project is on the "axis of risk".

Aggregated risk assessment for all groups of criteria α we calculate as follows:

$$O_R = \frac{1}{4} \sum_{\alpha} (1 - z^\alpha). \quad (5)$$

Since the evaluation is received O_R normalized, then to compare it with the output variable R the following scale is proposed: $r_1 = \langle \text{negligible risk level of project} \rangle$; $r_2 = \langle \text{low risk of project} \rangle$; $r_3 = \langle \text{average risk of a project} \rangle$; $r_4 = \langle \text{high risk of project} \rangle$; $r_5 = \langle \text{extreme risk level of project} \rangle$. Linguistic treatment of aggregated risk assessment $R = \{r_1, r_2, r_3, r_4, r_5\}$ define, for example, the following scale: $O_R \in (0,87; 1] - r_1$; $O_R \in (0,67; 0,87] - r_2$; $O_R \in (0,36; 0,67] - r_3$; $O_R \in (0,21; 0,36] - r_4$; $O_R \in [0; 0,21] - r_5$.

IV. RESULTS AND DISCUSSION

Formulate a generic algorithm to obtain aggregated safety assessment project financing.

1 step. Determine the resulting term-evaluation.

Based on expert input imposed on the project and built the knowledge base determines the resulting term-evaluation criteria for groups: $K_O; K_I; K_F; K_S$.

2 step. Determination of the aggregated estimation of reliability of expert considerations.

Aggregate validation calculates $\mu(L^\alpha)$, $\alpha = \{O; I; F; S\}$ according to the formula (2).

3 step. Obtaining a single generalized project risk assessment for groups of criteria α .

For each group of criteria we calculate the relative percentage scale $[a; b]$ and resultant term evaluation L^α , (which has the level of risk content) by the formula (3). A generalized project risk assessment for each group of criteria α gets for (4).

4 step. Calculation of aggregated risk assessment for all groups of criteria.

Aggregated risk assessment is determined by (5).

5 step. Output level of project financing security.

Match the evaluation O_R with output variable R to obtain a linguistic interpretation of the level of security of project financing.

For an example, consider the following problem. You need to build an initial estimate SPF and a linguistic interpretation of risk regarding the level of security of project financing. At this stage, we have an assessment of the project, depending on its origin – O_S , an assessment of the economy in which a commercial project will be implemented – O_G , aggregated risk assessment for project implementation – O_R .

Let the Decision Maker for each assessment may specify weight ratios $\{p_S, p_G, p_R\}$ from some interval. Then we will determine the normalized weight coefficients accordingly:

$$\alpha_\delta = \frac{p_\delta}{\sum_\delta p_\delta}, \delta = \{S, G, R\}, \sum_\delta p_\delta = 1. \quad (6)$$

Since all the estimates obtained are normalized from the interval $[0; 1]$, we use the following approach to obtain a final assessment of the security level of the project financing. Depending on the psychological perception of the situation Decision Maker can choose one of the convolutions [6]:

$$\text{Pessimistic } M_1(S) = \frac{1}{\sum_\delta \frac{\alpha_\delta}{O_\delta}}; \quad (7)$$

$$\text{Cautious } M_2(S) = \prod_\delta (O_\delta)^{\alpha_\delta}; \quad (8)$$

$$\text{Average } M_3(S) = \sum_\delta \alpha_\delta O_\delta; \quad (9)$$

$$\text{Optimistic } M_4(S) = \sqrt{\sum_\delta \alpha_\delta (O_\delta)^2}. \quad (10)$$

Thus, we obtain an initial estimate from the interval $[0; 1]$. For the linguistic interpretation of risk, the value obtained by formulas (7) - (10) is comparable to one of the term sets $SPF = \{SPF_1, SPF_2, \dots, SPF_5\}$. The scale of estimates can be determined as follows: $M(S) \in (0,77; 1] - SPF_5$ («high level of security of project financing»); $M(S) \in (0,57; 0,77] - SPF_4$ («the level of security of project financing is above average»); $M(S) \in (0,36; 0,57] - SPF_3$ («average level of project financing security»); $M(S) \in (0,21; 0,36] - SPF_2$ («low level of project financing security»); $M(S) \in [0; 0,21] - SPF_1$ («very low level of

project financing security»).

Depending on the different periods of the project implementation, we can review the initial assessment and aggregate risk assessment of the project implementation.

V. EXPERIMENTS

Let some investment project undergo an expert evaluation. The values of the linguistic variables and the authenticity of their assignment are as follows:

1. K^O – "operational risk": K_1^O (H; 0,8); K_2^O (H; 0,7); K_3^O (HC; 0,9); K_4^O (H; 0,6); K_5^O (HC; 0,7); K_6^O (C; 0,5); K_7^O (H; 0,7); K_8^O (H; 0,8); K_9^O (H; 0,9).

2. K^I – "investment risks": K_1^I (HC; 0,7); K_2^I (H; 0,5); K_3^I (C; 0,6); K_4^I (HC; 0,8); K_5^I (HC; 0,9).

3. K^F – "financial risks": K_1^F (HC; 0,3); K_2^F (HC; 0,6); K_3^F (HC; 0,2); K_4^F (H; 0,7); K_5^F (H; 0,6).

4. K^S – "risks of innovation": K_1^S (H; 0,8); K_2^S (H; 0,9); K_3^S (HC; 0,1); K_4^S (HC; 0,7); K_5^S (HC; 0,6).

To obtain an aggregated safety assessment of a project financing, use the following algorithm:

1 step. Determination of the resultant term evaluation.

Based on the knowledge base for each group of risk criteria, we define the resulting term evaluation: "operational risk" – H; "investment risks" – HC; "financial risks" – HC; "risks of innovation" – HC.

2 step. Calculation of the aggregated estimation, of the reliability, of the experts reasoning. Aggregate validation $\mu(L^\alpha)$, $\alpha = \{O; I; F; S\}$ calculate according to the formula (2):

$$\mu(L^O) = 1/6 * (0,8 + 0,7 + 0,6 + 0,7 + 0,8 + 0,9) = 0,8;$$

$$\mu(L^I) = 1/3 * (0,7 + 0,8 + 0,9) = 0,8;$$

$$\mu(L^F) = 1/3 * (0,3 + 0,6 + 0,2) = 0,6;$$

$$\mu(L^S) = 1/3 * (0,1 + 0,7 + 0,6) = 0,5.$$

3 step. Obtaining a single generalized project risk assessment for groups of criteria α .

For each group of criteria α we calculate by the formula (3). A generalized project risk assessment for each groups of criteria α get for (4).

$$x^0 = 20 - \sqrt{\frac{1-0,8}{2}}(20-0) = 13;$$

$$z^0 = \frac{13}{100} \approx 0,13;$$

$$x^I = 40 - \sqrt{\frac{1-0,8}{2}}(40-20) = 33,7; \quad z^I = \frac{33,7}{100} \approx 0,34;$$

$$x^F = 40 - \sqrt{\frac{1-0,6}{2}}(40-20) = 29; \quad z^F = \frac{29}{100} \approx 0,29;$$

$$x^S = \sqrt{\frac{1-0,5}{2}} \cdot 20 + 20 = 30; \quad z^S = \frac{30}{100} \approx 0,3.$$

4 step. Calculation of aggregated risk assessment for all groups of criteria α . Aggregate risk assessment is determined by (5):

$$O_R = \frac{1}{4}((1-0,13) + (1-0,34) + (1-0,29) + (1-0,3)) = 0,74.$$

5 step. Determine the level of security of project financing.

Match the evaluation O_R with output variable R to obtain a linguistic interpretation of the level of security of project financing.

Because $O_R \in (0,67; 0,87] - r_2$, then the project under consideration will receive "a low level of project risk or a level of security of project financing above average".

Depending on the different periods of the project implementation, we can review the initial assessment and aggregate risk assessment of the project implementation.

Built in such a way two-level fuzzy mathematical model,

obtaining an aggregated risk assessment of the project, has a number of advantages, namely: uses the expert's reasoning for assessing the various risk criteria; the reliability of his reasoning and, based on this; the aggregation of opinions according to the groups of criteria in the final assessment. The disadvantages of this approach include the use of different models of membership functions, which can lead to ambiguity of end results.

REFERENCES

- [1] M. Kelemen, "Problems of Protected Interests in the Security Sectors," Warsaw: Wydawnictwo Wyższej szkoły menedżerskiej w Warszawie im. Prof. Leszka J. Krzyżanowskiego, 2015.
- [2] I. I. Verbitska, "Ryzik-menedzhment yak suchasna systema upravlinnya ryzykamy pidpryemnytskykh struktur," Staly rozvytok ekonomiky, Vol. 5, pp. 282-291, 2013.
- [3] M. Kelemen, S. Krizovsky, S. Kocan, "Vyuzitie technologie LVA (vrstvena analiza hlasu) v bezpecnostnej praxi, na prevenciu proti podvodom u poisovacich a financnych spolocnosti", 1. vyd. Kosice: Vysoka skola bezpecnostneho manazerstva v Kosiciach. 100, 2012.
- [4] M. Malyar, V. Polishchuk, "Choice and evaluation methodics of investment projects," Kosicka bezpecnostna revue, Kosice, 1, pp. 117-126, 2013.
- [5] M. Malyar, V. Polishchuk, M. Sharkadi, and I. Liakh, "Model of start-ups assessment under conditions of information uncertainty," Eastern European Journal of Enterprise Technologies, Mathematics and cybernetics – applied aspects, vol. 3/4(81), pp.43-49, 2016.
- [6] M. Malyar, V. Polishchuk. Ranking method of alternative options of inhomogeneous nature. Kosicka bezpecnostna revue, Kosice, 1, 60-67, 2016