

ANDRIIV Roman

Separate structural unit "Berezhany Professional College of NUBiP of Ukraine"

SENKIVSKYY Vsevolod

Lviv Polytechnic National University

INFORMATION TECHNOLOGY FOR PREDICTING THE RELIABILITY LEVEL OF TEXT MESSAGES

The research presents the results of the creation of an intellectualized information technology for predictive analysis of the reliability of text information messages, formed on the basis of concepts and tools of fuzzy logic. The use of the fuzzy set apparatus makes it possible to take into account the semantic ambiguity inherent in a natural language, as well as to formalize qualitative expert assessments by using linguistic variables, fuzzy term-sets and a rule base of the "if-then" type. This provides the possibility of creating adaptive decision-making models in conditions of incompleteness, inconsistency and subjectivity of input information.

The developed technology includes fuzzification of input characteristics of texts, aggregation of expert judgments, construction of a system of fuzzy rules for assessing the reliability level and defuzzification of the obtained results. A concept is implemented that allows for a predictive assessment of the veracity of data even before their potential appearance in the information space. Within the framework of the proposed approach, a structured information database is formed, which establishes a relationship between the input variables, their linguistic nature, permissible ranges of values of the universal term-set, as well as clearly defined linguistic terms used for qualitative interpretation of parameters. Based on the performed structuring of linguistic variables of the studied process, a method of logical inference is developed, which represents a multi-level hierarchy of relationships between database components and determines the algorithm for calculating the message reliability indicator. The method is based on a knowledge matrix, leading to the construction of fuzzy logical equations, which provide the calculation of normalized values of membership functions of linguistic variables at the division points of the universal set. The result is the defuzzification of the fuzzy set "the indicator of the reliability level of text information messages" and the calculation of its value using the centre of mass formula, taking into account the input data. As a result of the study, a structural model of the information technology component of assessing the veracity of news content is developed.

Keywords: reliability of text messages, linguistic variable, logical inference method, knowledge matrix, membership function, fuzzy logic equations, information technology.

АНДРІЙ Роман

Відокремлений структурний підрозділ «Бережанський фаховий коледж НУБіП України»

СЕНЬКІВСЬКИЙ Всеволод

Національний університет «Львівська політехніка»

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПРОГНОЗУВАННЯ РІВНЯ ДОСТОВІРНОСТІ ТЕКСТОВИХ ПОВІДОМЛЕНЬ

У межах дослідження представлено результати створення інформаційної технології прогнозного аналізу достовірності текстових інформаційних повідомлень, побудованої на основі концепцій і засобів нечіткої логіки. Застосування апарату нечітких множин дало змогу враховувати семантичну неоднозначність, притаманну природній мові, а також формалізувати якісні експертні оцінки шляхом використання лінгвістичних змінних, нечітких терм-множин і бази правил типу «якщо-то». Це забезпечує можливість побудови адаптивних моделей прийняття рішень в умовах неповноти, суперечливості та суб'єктивності вхідної інформації.

Розроблена технологія включає фазифікацію вхідних характеристик текстів, агрегування експертних суджень, побудову системи нечітких правил для оцінювання рівня достовірності та дефазифікацію отриманих результатів. Реалізовано концепцію, яка дозволяє здійснювати прогнозну оцінку правдивості даних ще до їх потенційної появи в інформаційному просторі. У межах запропонованого підходу сформовано структуровану інформаційну базу даних, яка встановлює зв'язок між вхідними змінними, їхньою лінгвістичною природою, допустимими діапазонами значень універсальної терм-множини, а також чітко визначеними лінгвістичними термами, що застосовуються для якісної інтерпретації параметрів. На основі виконаного структурування лінгвістичних змінних досліджуваного процесу розроблено метод логічного виведення, що відтворює багаторівневу ієрархію зв'язків між компонентами бази даних та визначає алгоритм розрахунку показника достовірності повідомлень. В основі методу лежить матриця знань, що обумовила побудову нечітких логічних рівнянь, які забезпечили розрахунок нормалізованих значень функцій належності лінгвістичних змінних у точках розподілу універсальної множини. Наслідком стала дефазифікація нечіткої множини «показник рівня достовірності текстових інформаційних повідомлень» та розрахунок за формулою центра мас його значення з урахуванням введених вхідних даних. У підсумку дослідження розроблено структурну модель компоненти інформаційної технології оцінювання правдивості новинного контенту.

Ключові слова: достовірність текстових повідомлень, лінгвістична змінна, метод логічного виведення, матриця знань, функція належності, нечіткі логічні рівняння, інформаційна технологія.

Introduction

In the modern information environment, the problem of assessing the reliability of text messages is becoming particularly relevant due to the high level of information noise, the spread of manipulative content and fake messages. The process of determining the reliability level is characterized by a significant number of subjective

factors that are difficult to formalize by traditional methods. In this context, there is a need to develop flexible and adaptive solutions that can take into account the uncertainty and ambiguity inherent in a natural language.

For this purpose, the study proposes an information technology for predictive assessment of the reliability of text information messages, based on the use of fuzzy logic. This approach allows one to model linguistic variables, formalize expert judgments and implement the decision-making process in conditions of vaguely defined input parameters. The proposed technology is universal in the sense that it can be adapted to various thematic domains, types of information and sources, which provides wide possibilities for practical application – from automated monitoring of news streams to assessing the reliability in social networks.

An important component of the study is a review of publications related to the declared topic, which will determine the relevance and reliability of the results obtained.

The development of automated fake news detection methods is a key area of modern research in the field of information security. The works [1–4] present a wide range of approaches, covering probabilistic models, structural text analysis, graph convolutional networks and intelligent frameworks for data analysis in social networks. These solutions are aimed at identifying unreliable information using formalized characteristics that indicate potential distortion of facts or manipulateness. Algorithmic methods, in particular machine learning, are increasingly used as an alternative to traditional manual fact-checking [5], but remain dependent on the quality and completeness of training samples. In this context, there is growing interest in models that take into account not only the content, but also the behavioural aspects of users of the information environment [6], and are also aimed at developing the digital literacy of future media professionals [7]. Of particular scientific value are approaches that integrate fuzzy logic tools into the process of automated message reliability analysis [8–9]. Such solutions allow creating adaptive systems that can work effectively in conditions of incomplete or fuzzy data, in particular in library information environments. The use of hybrid models that combine fuzzy logic with deep learning methods significantly improves the accuracy and reliability of fake news detection [9], especially in conditions of increased sensitivity to data security, for example, in IoT environments [10]. In this case, decision-making algorithms based on extended fuzzy logical inference are implemented, which takes into account the complexity and dynamics of information flows. The works [11–12] substantiate the advantages of using intelligent control systems using fuzzy logic to increase the reliability of decision-making in complex information environments. By using membership functions and fuzzy rules, such systems are able to effectively respond to multifactorial influences and ensure stability in situations of information uncertainty. Fake news is increasingly being considered as a tool for targeted disinformation activities [13], which requires a comprehensive analysis taking into account both technical and social factors of content distribution. For a wide range of users, it is advisable to apply reliability assessment criteria that do not require special training, but allow one to recognize signs of unreliability, in particular emotionality, sensationalism and lack of accuracy [14]. In the media analytical context, a balance between facts and value judgments is important, which ensures the completeness and objectivity of the information presented [15].

The analysis of scientific sources allows one to conclude that despite the significant number of modern approaches to detecting fake news, the application of the reliability fuzzy logic apparatus remains limited. In this regard, it is urgent to create unified information models and information technologies based on the fuzzy set theory, which will provide a predictive assessment of the reliability of information messages.

Formation of an indicator for assessing the reliability of text information messages

Taking into consideration the complexity and multi-factor nature of the process of assessing the reliability of text information messages, this study substantiates, as noted above, the feasibility of using the fuzzy logic apparatus. Within the framework of the study, a structured information database is formed that establishes the relationship between input variables, their linguistic nature, permissible ranges of values of the universal term-set, as well as clearly defined linguistic terms used for qualitative interpretation of parameters.

The linguistic variable of the reliability of text messages is presented as a function:

$$Q = F_Q(B, T, L), \quad (1)$$

the arguments of which are linguistic variables (LV) of the second level B, T, L . At this level, the linguistic variable B is oriented towards factors of organizational orientation (b_1 – a source of information, b_2 – fact checking, b_3 – multiple publication); T defines a procedure-function related to the author and the context of the message (t_1 – professionalism of the author, t_2 – objectivity of the author, t_3 – informativeness of the message context); L characterizes the share of the indicator assessed by the attitude of users towards the received news (l_1 – refutation and criticism, l_2 – social trust)

Let one form a table of term-sets of linguistic variables, including the description of LV and a database of values [17].

Table 1

Term-sets of values of linguistic variable			
LV	Linguistic description of the variable	Universal base of values U	Linguistic terms (term-set H)
b_1	A source of information (reliability)	(1-5) c. u.	Low, medium, high
b_2	Fact checking	(1-5) c. u.	Infrequent, frequent, constant
b_3	Multiple publication (quantity)	(1-5) c. u.	Small, medium, large
t_1	Professionalism of the author	(1-5) c. u.	Low, medium, high
t_2	Objectivity of the author	(1-5) c. u.	Low, medium, high
t_3	Informativeness of the message context	(1-5) c. u.	Low, medium, high
l_1	Refutation and criticism (frequency)	(1-5) c. u.	Small, medium, significant
l_2	Social trust	(1-5) c. u.	Low, medium, high

Based on the structuring of linguistic variables in the predictive assessment process of the reliability of text information messages, a method of logical inference is developed that reproduces a multi-level hierarchy of relationships between database components and determines the algorithm for calculating the RTIM indicator.

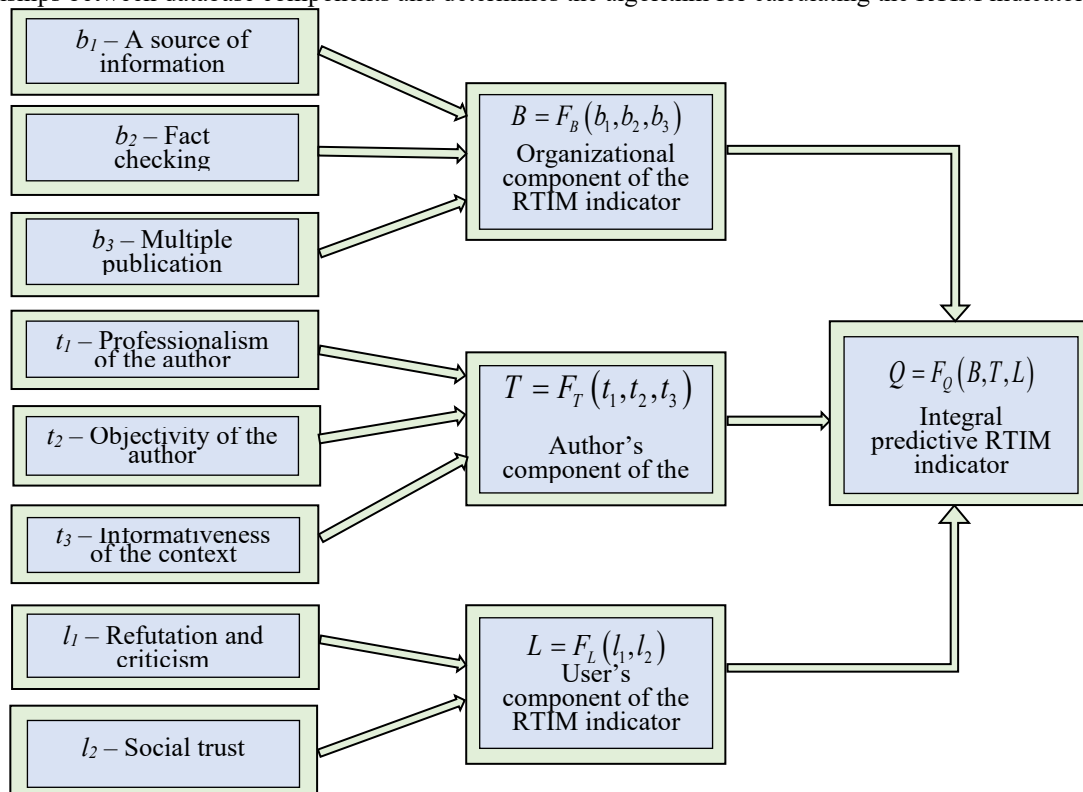


Fig. 1. Structural model of the logical inference method – formation of the indicator of the reliability of text information messages

Taking into account the above, the linguistic term “the indicator of reliability of text messages” is displayed as a fuzzy set [16]:

$$Q = \left\{ \frac{\mu_q(u_1)}{u_1}, \frac{\mu_q(u_2)}{u_2}, \dots, \frac{\mu_q(u_n)}{u_n} \right\}, \quad (2)$$

where: $Q \subset U$; $\mu_q(u_i)$ is a membership function (MF) of the element $u_i \in U$ to the set Q .

For each LV and its corresponding three linguistic term-sets H , square inversely symmetric matrices $A = a_{ij}$ ($a_{ij} = r_i/r_j$; $i, j = 1, \dots, 5$) are constructed, the elements of which are formed by comparing ranks with each other. Fuzzy logical inference for the highest level has the following formalized form [17]:

IF (B = low) AND (B = medium) AND (B = high)
 AND (T = low) AND (T = medium) AND (T = high)
 AND (L = low) AND (L = medium) AND (L = high),
 THEN (Q = low) AND (Q = medium) AND (Q = high).

The formulated logical inference determines the development of the corresponding knowledge matrix in the form of Table 2.

Table 2

Knowledge matrix for the linguistic variable Q

Organizational component of RTIM B	Author's component of RTIM T	User's component of RTIM L	Integral RTIM indicator Q
low	low	low	low
low	medium	low	
medium	low	medium	medium
medium	medium	high	
high	high	medium	high
high	high	high	

Based on the knowledge matrix in Table 2, fuzzy logical equations are constructed that provide the calculation of the values of membership functions.

$$\mu_{low}(Q) = \mu_{low}(B) \wedge \mu_{low}(T) \wedge \mu_{low}(L) \vee \mu_{low}(B) \wedge \mu_{low}(T) \wedge \mu_{low}(L).$$

$$\mu_{medium}(Q) = \mu_{medium}(B) \wedge \mu_{low}(T) \wedge \mu_{medium}(L) \vee \mu_{medium}(B) \wedge \mu_{medium}(T) \wedge \mu_{high}(L).$$

$$\mu_{high}(Q) = \mu_{high}(B) \wedge \mu_{high}(T) \wedge \mu_{середня}(L) \vee \mu_{high}(B) \wedge \mu_{high}(T) \wedge \mu_{high}(L).$$

The specified approach is applied to LV of the following levels B , T and L .

The linguistic term Q "the indicator of reliability of text information messages" is presented in the form of a fuzzy set [16]:

$$Q(B, T, L) = \left\{ \frac{\mu_{low}(Q)}{q_1}, \frac{\mu_{medium}(Q)}{q_2}, \frac{\mu_{high}(Q)}{q_3} \right\}, \quad (3)$$

where q_1, q_2, q_3 determine the quantitative values of the linguistic variable Q in relation to the above-mentioned terms at the division points of the universal set.

Experiments

At the final stage of theoretical research, after determining the main linguistic variables and constructing the corresponding membership functions, these variables are formalized in the form of table structures. For this purpose, tables with normalized values of membership functions at five characteristic division points of the universal set are formed for each of the selected linguistic variables. This approach allows ensuring the consistency and correctness of further calculations within the fuzzy logic system.

An example of constructing the corresponding table for a linguistic variable is presented below, which at the same time serves as the basis for moving to the experimental stage of research, where the constructed model is tested on the basis of real or simulated input data.

Table 3

Membership functions of the term-set $H(b_i)$ (a source of information – reliability)

u_i , c. U.	1	2	3	4	5
$\mu_{low}(u_i)$	1	0,86	0,54	0,34	0,11
$\mu_{medium}(u_i)$	0,12	0,66	1	0,56	0,12
$\mu_{high}(u_i)$	0,11	0,33	0,56	0,78	1

The next step is to select the points of the universal set U and their corresponding values of the membership functions. Let the following input data be selected for the linguistic variables:

$$b_1 = 5; b_2 = 3; b_3 = 4; t_1 = 4; t_2 = 3; t_3 = 2; l_1 = 3; l_2 = 4$$

and their corresponding values of the membership functions.

By substituting the given values into fuzzy logical equations, the values of the membership functions for linguistic variables B, T, L are obtained:

$$\begin{aligned}\mu_{low}(B) &= 0,11; \mu_{medium}(B) = 0,12; \mu_{high}(B) = 0,40; \\ \mu_{low}(T) &= 0,30; \mu_{medium}(T) = 0,50; \mu_{high}(T) = 0,12; \\ \mu_{low}(L) &= 0,22; \mu_{medium}(L) = 0,30; \mu_{high}(L) = 0,50.\end{aligned}$$

The membership functions of the highest-level linguistic variable are calculated:

$$\mu_{low}(Q) = 0,11; \mu_{medium}(Q) = 0,12; \mu_{high}(Q) = 0,12.$$

The process of determining the quantitative indicator of the reliability level of text information messages is implemented by using the centre of mass method, which is one of the most common defuzzification methods within fuzzy logic. This method allows one to move from linguistic assessments, presented in the form of fuzzy sets, to a specific numerical value that characterizes the predictive level of information reliability.

The indicator of predictive reliability of information messages P is calculated using the following formula [17]:

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

where: \underline{Q}, \bar{Q} mean the minimum and maximum values of the message reliability indicator; m is a number of fuzzy terms of the linguistic variable Q . For calculations, the following initial values are set: $m=3$; $\mu_1(Q) = \mu_{low}(Q)$, $\mu_2(Q) = \mu_{medium}(Q)$, $\mu_3(Q) = \mu_{high}(Q)$. Lower and upper limits of values for the linguistic variable Q are presented as a percentage: $\underline{Q}=1\%$; $\bar{Q}=100\%$.

The calculations are performed at three points of the specified interval: $q_1 = 1$; $q_2 = 50$; $q_3 = 100$. Finally, according to formula (9), the following value of the message reliability level indicator is obtained: $P = 51,74\%$. The indicator calculated at the given input values of linguistic variables indicates the *average* reliability level of the information message.

Taking into account the theoretical and experimental components of the study, a key component of information technology will be designed for predictive assessment of text message reliability.

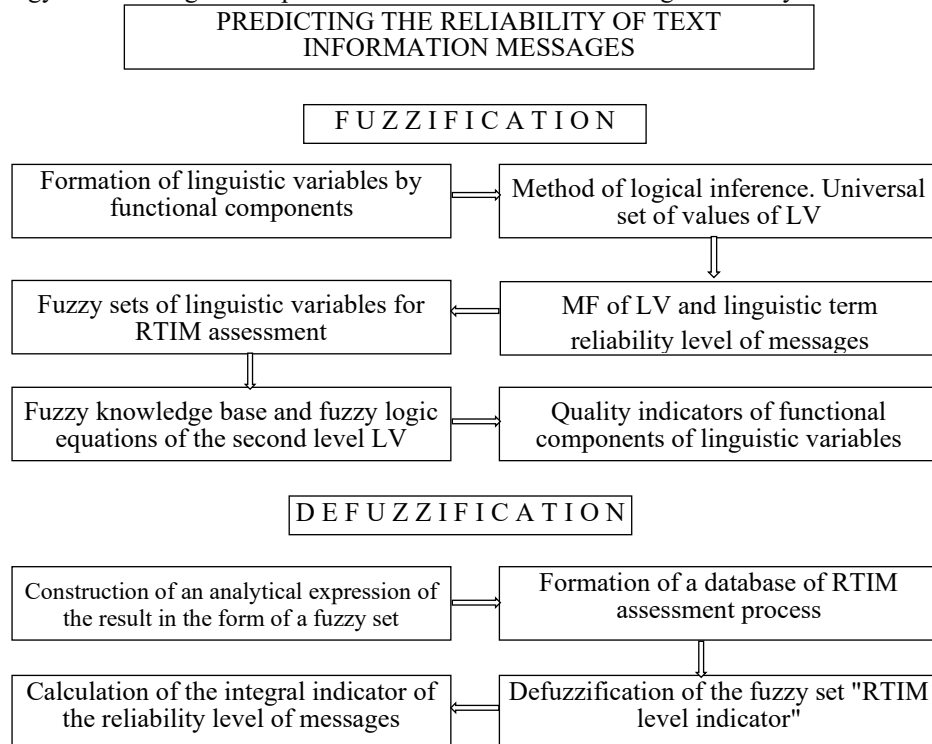


Fig. 2. Structural model of the information technology component of predictive assessment of the text message reliability

The developed structural model of the key component of the information technology for predictive assessment of the text message reliability provides a holistic view of the interaction of the main subsystems focused on processing, analysis and interpretation of text data, taking into account multi-factor influence. The central element of the model is the method of fuzzy logical inference, which integrates the results of preliminary linguistic and semantic processing, reliability parameters, as well as contextual characteristics of messages. The architecture of the model provides for adaptive tuning based on expert knowledge, weight coefficients of influence factors and membership functions, which allows forming a generalized indicator of the reliability level, taking into account incompleteness, inconsistency or unstructuredness of information. The model can be effectively integrated into application systems for monitoring the information environment, increasing the accuracy of detecting fake news and contributing to making informed decisions in the field of information security.

Conclusions

As a result of the research, a key component of intellectualized information technology for predictive assessment of the reliability of text messages, based on fuzzy logic tools, is developed. The proposed approach allows to effectively formalize subjective expert judgments and adapt the analysis process to the conditions of semantic uncertainty characteristic of a natural language. The basis of the implemented technology is a system of fuzzification, aggregation, logical inference and defuzzification, which provides multi-level processing of text information taking into account linguistic variables and their values in the term-space.

An information database is designed that establishes relationships between input parameters, their linguistic representation and permissible ranges of values. A method of fuzzy logical inference is developed, which is based on a knowledge matrix and a hierarchy of the process of forming a message reliability indicator and allows calculating the reliability level using a system of logical equations. The target indicator value is determined using the center of mass method, which provides an integrated quantitative assessment of the message reliability.

As a result, a structural model of a key component of information technology is formed, which can be used to assess the veracity of news content at the stage of its appearance or distribution in the information space, which opens up prospects for its integration into information monitoring systems and countering disinformation. This approach allows not only to carry out multifactor analysis of messages, but also to effectively generalize assessments from different sources. The proposed technology is characterized by a high degree of versatility, since its architecture allows the adaptation to different subject areas, genre features of texts and the specificity of information sources. This creates broad prerequisites for practical implementation – in particular, in the tasks of automated monitoring of news flows, verification of information in social media, as well as in the field of information security.

References

1. D. K. Dixit, A. Bhagat and D. Dangi. Automating fake news detection using PPCA and levy flight-based LSTM. *Soft Comput.*, num. 26, num. 22 (2022) 12545–12557. <https://doi.org/10.1007/s00500-022-07215-4>.
2. Y. Wang, L. Wang, Y. Yang and Y. Zhang. Detecting fake news by enhanced text representation with multi-EDU-structure awareness. *Expert Syst. Appl.*, num. 206 (2022). <https://doi.org/10.1016/j.eswa.2022.117781>.
3. D. Michail, N. Kanakaris and I. Varlamis. Detection of fake news campaigns using graph convolutional networks. *Int. J. Inf. Manag. Data Insights*, num. 2 (2022). <https://doi.org/10.1016/j.jjime.2022.100104>.
4. S. Rastogi, D. Bansal. Disinformation detection on social media: An integrated approach. *Multimed. Tools Appl.*, num. 81, num. 28 (2022) 40675–40707. <https://doi.org/10.1007/s11042-022-13129-y>.
5. Rastogi, S., Bansal, D. A review on fake news detection 3T's: typology, time of detection, taxonomies. *Int. J. Inf. Secur.* 22 (2023) 177–212. <https://doi.org/10.1007/s10207-022-00625-3>.
6. Ines Kozuh, Peter Čakš Social Media Fact-Checking: The Effects of News Literacy and News Trust on the Intent to Verify Health-Related Information *October Healthcare*, (2023) 11(20):2796 DOI: 10.3390/healthcare11202796.
7. Lyudmila Shesterkina, Lidiya Lobodenko, Anna Krasavina, Arina Marfityna Fact-Checking and Information Verification in the Context of Journalism Education (March 2021). *Theoretical and Practical Issues of Journalism* 10(1):94-108 DOI: 10.17150/2308-6203.2021.10(1).
8. Nikolaou, N., & Karypis, G. Automated fake news detection: A survey. *Social Network Analysis and Mining*, 10(1) (2020) 1-29. DOI: 10.6017/ital.v39i4.12483.
9. Cheng Xu and M-Tahar Kechadi. 2023. Fuzzy Deep Hybrid Network for Fake News Detection. In *The 12th International Symposium on Information and Communication Technology (SOICT 2023)*, December 07–08 (2023), Ho Chi Minh, Vietnam. ACM, New York, NY, USA 8 Pages. <https://doi.org/10.1145/3628797.3628971>.
10. Memon, I., Shaikh, R. A., Hasan, M. K., Hassan, R., Haq, A. U., & Zainol, K. A. (2020). Protect mobile travelers information in sensitive region based on fuzzy logic in IoT technology. *Security and Communication Networks*, (1), (2020) 8897098. DOI:10.1155/2020/8897098.
11. Dumitrescu, C., Ciotirae, P., & Vizitiu, C. Fuzzy logic for intelligent control system using soft computing applications. *Sensors*, 21(8), (2021) 2617. <https://doi.org/10.3390/s21082617>.
12. Daradkeh, Y. I., & Tvoroshenko, I. Technologies for making reliable decisions on a variety of effective factors using fuzzy logic. *International Journal of Advanced Computer Science and Applications*, 11(5), (2020). DOI: 10.14569/IJACSA.2020.0110507.
13. NORDBERG, Pontus; KÄVRETTAD, Joakim; NOHLBERG, Marcus. Automatic detection of fake news. In: 6th International Workshop on Socio-Technical Perspective in IS Development, virtual conference in Grenoble, France, June 8-9, 2020. CEUR-Ws, (2020). p. 168-179.
14. Z. Brzhevskaya, H. Haydur, N. Dovzhenko, A. Anosov. Criteria for monitoring the reliability of information in the information space. *Cyber security: education, science, technology*, No. 1(5), 2019, pp. 105-112. (2019). <https://doi.org/10.28925/2663-4023.5.52.60>.

15. Marcheniuk, M. S., Kozachok, V. A., Bogdanov, O. M., Brzhevska, Z. M. Analysis of methods of detecting disinformation in social networks using machine learning. *Cyber security: education, science, technology*, No. 2(22) (2023). pp. 148-155. <https://doi.org/10.28925/2663-4023.2023.22.148155>.
16. B. V. Durnyak, I. V. Pikh, V. M. Senkivskyi. (2022). Theoretical foundations of the information concept of forming and assessing the quality of publishing and printing processes. Monograph. – Lviv: Ukrainian Academy of Printing. 356 p. URL: <https://biblio.uad.lviv.ua>.
17. Vsevolod Senkivskyi, Iryna Pikh, Alona Kudriashova, Roman Andriiv, Nazarii Senkivskyi. Evaluation of methods for intellectual analysis of manipulative content in the information space. The 6th International Workshop on Intelligent Information Technologies & Systems of Information Security, 2025, Khmelnytskyi, Ukraine. Pp. 139-155. <https://ceur-ws.org/Vol-3963/paper12.pdf>

Roman Andriiv Роман Андріїв	Teacher of mathematics, physics and computer science of the cyclical commission of physical and mathematical disciplines and information technologies Separate structural unit «Berezhany Professional College of NUBiP of Ukraine» Berezhany, Ternopil region, Ukraine e-mail: roman.andriiv@icloud.com https://orcid.org/0009-0001-2450-5439 Scopus Author ID: 59548768000	Викладач математики, фізики та інформатики циклової комісії фізико-математичних дисциплін та інформаційних технологій Відокремлений структурний підрозділ «Бережанський фаховий коледж НУБіП України», Тернопільська обл., Україна
Vsevolod Senkivskyi Всеволод Сеньківський	Doctor of Technical Sciences, Professor, Professor of Computer Technologies in Publishing and Printing Processes Department, Lviv Polytechnic National University, Lviv, Ukraine, e-mail: vsevolod.m.senkivskyi@lpnu.ua https://orcid.org/0000-0002-4510-540X Scopus Author ID: 57208667450	Доктор технічних наук, професор, професор кафедри комп'ютерних технологій у видавничо-поліграфічних процесах, Національний університет «Львівська політехніка», Львів, Україна.