

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА ТЕОРІЯ КОДУВАННЯ

УДК 004.8 + 004.056

Ю. В. Барішев, Н. Р. Кондратенко, В. В. Казміревський, Т. Г. Кирилащук

НЕЧІТКІ МНОЖИНИ ТИПУ-2 В ЗАДАЧАХ МОДЕЛЮВАННЯ ТА ОЦІНЮВАННЯ СТАНІВ КРИТИЧНИХ СИСТЕМ З НЕДОВИЗНАЧЕНИМИ ВХІДНИМИ ДАНИМИ ТА ВИКОРИСТАННЯМ ЕКСПЕРТІВ

Вінницький національний технічний університет, м. Вінниця

Анотація. Запропоновано методику використання нечітких множин типу-2 в задачах моделювання та оцінювання станів критичних систем з недовизначеними вхідними даними. Показано, що основою розв'язання задачі моделювання є побудова системи нечіткої логіки з інтервальними функціями належності типу-2. В роботі представлено задачу розвитку процесу оцінювання інтервального виходу нечіткої системи за допомогою експертів. Запропоновано підхід на основі нечітких множин для розв'язання задачі оцінювання та моделювання станів критичних систем. На прикладі енергетичних систем, де наявний високий ступінь невизначеності, показано, що основні фактори, які впливають на появу невизначеностей у вихідних даних таких систем, обумовлюються відсутністю достатньої інформації у відкритому джерелі та високою мінливістю загроз під впливом зростання темпів цифровізації робочих процесів. Наведено аналіз експертних оцінок інтервального виходу нечітких систем на прикладах моделювання складних об'єктів в різних прикладних галузях. В першому прикладі продемонстровано результати моделювання в природничій області з недовизначеними вхідними даними для оцінювання перспективності артезіанської свердловини, де кінцеву оцінку виставляє експерт. Другий приклад демонструє реалізацію інтервальної нечіткої моделі в завданні соціального спрямування, де моделюється проблема підбору персоналу в соціальних групах з точки зору професійної придатності. Третє завдання відноситься до моделювання в галузі медичного діагностування захворювань ендокринної системи. Оцінювання експертами результатів інтервального нечіткого моделювання в цій галузі дає можливість визначити стан захворювання людини на ендокринну патологію та призначити своєчасне лікування. Наведені приклади оцінювання інтервального виходу нечіткої системи з врахуванням думки експертів, підтверджують можливості для прийняття рішень, що є адекватними предметній області в умовах недовизначених вхідних даних. Наведено перспективи застосування запропонованих моделей для задач кібербезпеки критичних систем.

Ключові слова: нечіткі множини типу-1 та типу-2, інтервальна нечітка модель, функція належності, недовизначені вхідні дані, експертні знання, кібербезпека, критичні системи.

Abstract. A method of type-2 fuzzy sets implementation for critical systems' modeling and state assessment tasks with uncertain input data is proposed. It is shown that the basis for solving the modeling task is designing of a fuzzy logic system with interval membership functions of type-2. The paper presents the task of further developing the process of estimating the interval output of a fuzzy system with experts involvement. An approach based on fuzzy sets is proposed for solving the task of critical systems' modeling and states assessment. Using the example of energy grid systems, where a high degree of uncertainty is present, it is shown that the main factors that influence the appearance of uncertainty in the initial data set of such systems are caused by the lack of sufficient information in the open print and the high variability of threats under the influence of the growing pace of digitalization of business processes. An analysis of expert evaluations of the interval output of fuzzy systems based on examples of modeling complex objects in various fields of application is given. The first example demonstrates the results of modeling in the field of natural sciences with uncertain initial data for assessing the prospects of an artesian well, where the final assessment is made by an expert. The second example demonstrates the implementation of the interval fuzzy model in the task of social orientation, where the problem of recruiting personnel in social groups from the point of view of professional suitability is modeled. The third task refers to modeling in the field of medical diagnosis of diseases of the endocrine system. Evaluation by experts of the results of interval fuzzy modeling in this field makes it possible to determine the state of a person's disease for endocrine pathology and prescribe timely treatment. The given examples of evaluating the interval output of a fuzzy system, taking into account the opinion of experts, confirm the possibilities for making decisions that are adequate for the subject area in the conditions of uncertain input data. Prospects for the application of the proposed models for the problems of cyber security of critical systems are given.

Key words: type-1 and type-2 fuzzy sets, interval fuzzy model, belonging function, uncertain input data, expert knowledge.

DOI: <https://doi.org/10.31649/1999-9941-2023-57-2-13-24>.

Вступ

На даний час інформаційні системи в різних сферах людської діяльності (наука, техніка, економіка, медицина, промисловість та ін.) неможливо уявити без використання інтелектуальних технологій. Інформаційні системи використовують та накопичують знання, що представляються фахівцями з різних галузей. Ці знання надаються експертами на основі професійної мови та лінгвістичних висловлювань. Як правило, знання від експертів мають характер формулювань з нечіткими границями або висловленнями з багатозначною шкалою значень істинності. Тому сучасний підхід до розв'язання задач моделюванням станів складних систем з недовизначеними вхідними даними, пов'язують з використанням теорії нечітких множин та лінгвістичних змінних.

Теорія нечітких множин була вперше запропонована американським математиком Л. Заде в 1965 р. [1] і призначалася для подолання труднощів представлення неточних понять, аналізу і моделювання

систем, у яких бере участь людина. Поняття нечітких множин типу-2 (fuzzy sets type-2) було представлено в роботі з нечіткої логіки [2]. Подальший розвиток нечіткі множини першого та другого порядку (fuzzy sets type-1 та type-2) отримали в роботах відомих дослідників [3, 4]. У цих роботах береться до уваги те, що апарат нечітких множин типу-2 дає змогу будувати математичні моделі складних об'єктів та процесів, які здатні оперувати вхідними даними, що є малоінформативними або зашумленими. Важливою особливістю цих моделей є достатньо прозора побудова нечіткої моделі на основі причинно-наслідкових зв'язків, відсутність обмежень на складність моделі, оскільки моделюється відгук системи на вектор вхідних даних та непотрібна реалізація внутрішніх процесів, що відбуваються в системі. В рамках сучасного впровадження систем нечіткої логіки типу-2 є доцільним дослідження їх ефективності для моделювання станів, зокрема кібербезпеки, складних систем в умовах недовизначених вхідних даних та врахуванням точки зору експертів.

Аналіз поточного стану та постановка проблеми

В завданнях моделювання та оцінювання станів критичних систем поширеною є ситуація, коли необхідно приймати рішення, маючи суперечливі дані, що обумовлюється присутністю певного рівня «шуму» або пропусками в даних.

Для вирішення цієї проблеми використовуються інтервальні нечіткі множини, або нечіткі множини типу-2, коли як значення функцій належності подається не число, а інтервал. Особливість нечітких множин типу-2 полягає в тому, що в них ступінь належності є нечіткою множиною типу-1. Використання таких нечітких множин для розв'язання прикладних завдань, зумовлюється тим, що їх опис подається за допомогою нижньої та верхньої функції належності (membership function, MF), при цьому кожна з цих функцій представлялася нечіткою множиною першого порядку. Це сприяє появі інтервалу між цими двома функціями, цей інтервал назвали зоною невизначеності (footprint of uncertainty, FOU) [3, 4, 5]. В роботі [6], показано появу такої зони невизначеності при застосуванні функцій належності типу-2 в системах нечіткої логіки в умовах недовизначеності вхідних даних. В публікації [7] наведено дослідження властивостей нечітких множин типу-2 та їх переваги на відміну від нечітких множин типу-1 на прикладі роботи пошукових алгоритмів в завданнях оптимізації. Практичному використанню нечітких множин типу-2 присвячена робота [8], де показано використання інтервальної нечіткої системи типу-2 для дослідження адаптивного управління ковшним режимом виконавчих механізмів, для яких можливі відмови; справедливості такого підходу продемонстровано чисельним прикладом. Апарат інтервальних нечітких множин типу-2, було успішно використано в роботі [9], як математичний апарат, що дозволив провести моделювання станів складної природничої системи в умовах малоінформативних вхідних даних. В роботі викладено процес моделювання систем такого типу на прикладі оцінювання стану артезіанської свердловини в довільний момент часу в межах від початку гідрогеологічної розвідки до її повного завершення. Для дослідження артезіанської свердловини характерними факторами є важкодоступність експериментальних даних, неточність вимірювань, відсутність умов для безпосереднього спостереження об'єкта та інші. Ці фактори обумовлюють появу недовизначених вхідних даних про предметну галузь і роблять апарат нечітких множин типу-2 найкращим інструментом для опису предметної області. В роботі [10] продемонстровано розв'язок задачі моніторингу станів складної системи за допомогою узагальнювальної інтервальної нечіткої моделі за участі декількох експертів та показано процедура отримання кінцевого результату на основі теоретико-множинного підходу. В рамках задачі моделювання станів складної системи в роботах [11, 12] було представлено завдання медичного діагностування, де показано використання апарату інтервальних нечітких множин в умовах пропусків вхідних даних. Завдяки появі зони FOU вдалося описати неповноту та неточність даних в одній з найскладніших задач моделювання, а саме - прогнозування часових послідовностей із зашумленими даними. В роботі [13], викладено, що для успішного розв'язання задачі інтервального прогнозування часової послідовності за умови невизначеностей, для звуження інтервального прогнозу, будується узагальнювальна інтервальна нечітка модель типу-2, за допомогою якої вдається отримати більш вузький інтервальний прогноз.

В роботах [14, 15] наведено дослідження можливості використання інтервальних нечітких множин типу-2 та їх властивостей у відображенні предметної галузі соціального спрямування, де розглядається проблема з підбору персоналу у соціальних групах. Завдяки апарату інтервальних нечітких множин вдалося описати невизначеності, які виникають при проведенні тестування чи опитування людини та надати право експерту приймати кінцеве рішення стосовно вибору найкращого претендента на робоче місце.

В завданнях моделювання та оцінювання станів критичних систем привертають увагу енергетичні системи. Безпека енергетичних систем включає низку різномірних факторів, пов'язаних з балансуванням виробництва та споживання енергії, її розподілом, що пов'язане з ризиками для системи, навіть, за умов відсутності впливу зовнішніх чинників, навмисно направлених на її дестабілізацію таких, як кібератаки. У зв'язку зі збільшенням останніх, особливо на інфраструктуру України [16], перспективним виглядає застосування апарату нечітких множин для оцінювання кібербезпеки енергетичних систем. Це

обумовлюється, з одного боку, як було зазначено вище, зростанням кількості таких атак, та недостатністю кількості таких атак для формування адекватної статистичної вибірки [17]. Через те, що відомості щодо проведених атак певного виду можуть бути розділені роками, а при цьому інформаційні технології активно змінюються, більш віддалені в часі атаки менш інформативні, оскільки за цей проміжок часу змінюються програмні засоби та технології обробки інформації.

Безумовно, використання апарату інтервальних нечітких множин типу-2 в задачах, що пов'язані з моделюванням станів складних об'єктів в критичних системах, які представлені вище, повинно супроводжуватись дослідженнями адекватності нечітких моделей, які будуються на основі нечітких множин типу-2. В роботах [18, 19], подані дослідження на адекватність представлення предметній області нечіткими моделями, що будуються на інтервальних нечітких множинах типу-2.

Як показав наведений вище аналіз застосування інтервальних нечітких множин типу-2 є перспективним для моделювання станів об'єктів за умов недовизначених вхідних даних. Крім того, з аналізу випливає, що актуальною сферою застосування цього математичного апарату є критичні системи, зокрема стан їх кібербезпеки. Водночас відомі роботи сфокусовані на конкретних специфічних задачах, а тому є актуальним розширення можливостей відомих рішень для більш широкого кола задач в цій сфері застосування. В рамках викладених понять та аналізу предметної області, приймемо наступну мету та задачу дослідження.

Мета та задачі дослідження

Метою дослідження є розширення можливостей відомих нечітких моделей та методів прийняття рішень, що побудовані з використанням нечітких множин типу-2 в задачах моделювання та оцінювання станів критичних систем з недовизначеними вхідними даними та врахуванням думки експертів.

Для досягнення поставленої мети вирішувались такі задачі:

- розробити методику використання нечітких множин типу-2 в задачах моделювання та оцінювання станів критичних систем, пристосовану до роботи з недовизначеним вхідним вектором;
- запропонувати шляхи врахування експертних знань у процесі прийняття рішення;
- навести приклади можливості застосування запропонованої методики для різних задач моделювання стану критичних систем, зокрема стану кібербезпеки, як доказ концепції (proof of concept).

Виклад основного матеріалу

В подальшому викладі матеріалу розглянемо методику моделювання та оцінювання станів критичних систем із застосуванням нечітких множин типу-2, в рамках існування недовизначених вхідних даних, та проаналізуємо оцінювання кінцевого результату експертами.

Для з'ясування методики реалізації поставленого завдання, треба відмітити, що для завдань в такій постановці, моделі на основі нечітких множин типу-1 мають суттєве обмеження: вони не можуть напряму працювати на недовизначених вхідних даних, оскільки відомий механізм логічного висновування не спроможний визначити вихідне значення при неповному вхідному векторі. З огляду на це, для прийняття рішень в таких умовах, наведемо основні аспекти моделювання складних об'єктів в системах нечіткої логіки, що будуються з використанням нечітких множин типу-2, базові етапи моделювання є такими.

По-перше: при використанні апарату нечітких множин обов'язковою процедурою є формалізація причинно-наслідкових зв'язків між параметрами об'єкта, що досліджується, та його станом, за допомогою опису зв'язків на природній мові з використанням теорії лінгвістичних змінних у вигляді матриці знань. Вхідні і вихідні параметри можуть бути як кількісними, так і якісними. Кількісні параметри задаються діапазонами зміни, а якісні множинами можливих значень. Всі параметри розглядаються як лінгвістичні змінні [2-4].

По-друге: в умовах недовизначеності вхідних даних, коли можлива поява зашумлених, малоінформативних параметрів, нечітка система будується з використанням функцій належності, що реалізуються на традиційних типу-1 множинах та функцій належності - на інтервальних нечітких множинах типу-2. Стосовно методу генерації нечіткої системи, то найбільш поширеною є її генерування на основі заздалегідь підготовленої вибірки експериментальних даних, що показано нижче.

Нехай відома експериментальна вибірка X :

$$X = \{X_1, X_2, \dots, X_n\},$$

де $X_i = (x_{1i}, x_{2i}, \dots, x_{mi}, y_i), i = 1, \dots, n$; n – кількість експериментальних прикладів, m – кількість вхідних змінних, y – вихідна величина.

Процедура прийняття рішень реалізується за допомогою нечіткої логічної системи-класифікатора [5]. Вхідний вектор являє собою набір значень параметрів об'єкта $X^i = \{x_1^i, \dots, x_m^i\}$. База правил

формується з відомих пар (X_i, Y_i) , де вхідному вектору поставлено у відповідність лінгвістичну оцінку значення вихідної змінної Y , яка може надаватися експертом. Таким чином, кожен вхідний вектор експериментальних даних генерує одне правило. Антецеденти правил утворюються заміною значення x_j^i відповідним йому нечітким термом $A_{x_j}^i$, консеквентами є терм лінгвістичної змінної y , визначений експертом для вектора X^i :

$$R^i: IF x_1 \in A_{x_1}^i \wedge x_2 \in A_{x_2}^i \wedge \dots \wedge x_m \in A_{x_m}^i THEN y \in LV_y^k \in \{LV_1, \dots, LV_p\},$$

де x_i – вхідні змінні, y – вихідна змінна, $LV_y \in \{LV_1, \dots, LV_p\}$ – терм-множини вихідної змінної. Терм-множини вхідних та вихідної змінної описуються гаусівськими функціями належності

Таким чином, на основі набору значень (X, Y) , а також бази правил, синтезованої на основі набору експериментальних даних, спочатку будується нечітка логічна система типу-1 з чітким виходом Y , який задається за бальною шкалою. Після цього виконується перетворення результируючих функцій належності типу-1 на інтервальні функції належності типу-2. Опишемо нижче.

Розглянемо основні кроки побудови інтервальних функцій належності типу-2 за допомогою алгоритму [11, 12]. За цим алгоритмом одночасно здійснюється дві дії. За першою - відбувається зменшення вибірки, а саме, виключення певного правила та після цього розрахунок відгуку моделі по всій вибірці. Якщо похибка відпрацювання системи дорівнює 0, тоді кількість правил стає менше, якщо ні, виключене правило повертається до вибірки. За другою дією - відбувається перехід до інтервальної функції належності типу-2. Перед цієї дією вибираємо тип функції належності. В якості функції належності для вхідних змінних, беремо гаусову функцію типу:

$$\mu(x) = e^{-\left(\frac{x-b}{c}\right)^2},$$

Для опису інтервальних функцій належності оберемо модифіковану гаусову форму. Загальний вигляд такої інтервальної функції належності:

$$\mu(x) = e^{-\left(\frac{x-b}{[\min(c), \max(c)]}\right)^2},$$

де $[\min(c), \max(c)]$ – діапазон зміни параметру c гаусової функції належності.

Далі здійснюється перетворення функцій належності типу-1 на інтервальні функції належності типу-2. За вказаним алгоритмом визначається верхня та нижня межі діапазону параметру c . Множимо усі параметри функцій належності на невелику величину, наприклад, $\Delta=0.001$ та перевіряємо відгук системи на всій вибірці, якщо система не припускає помилки в рішенні поставленого завдання, процедура повторюється. Визначення змін функції належності завдяки перерахунку значення параметра c , показано на рисунку 1.

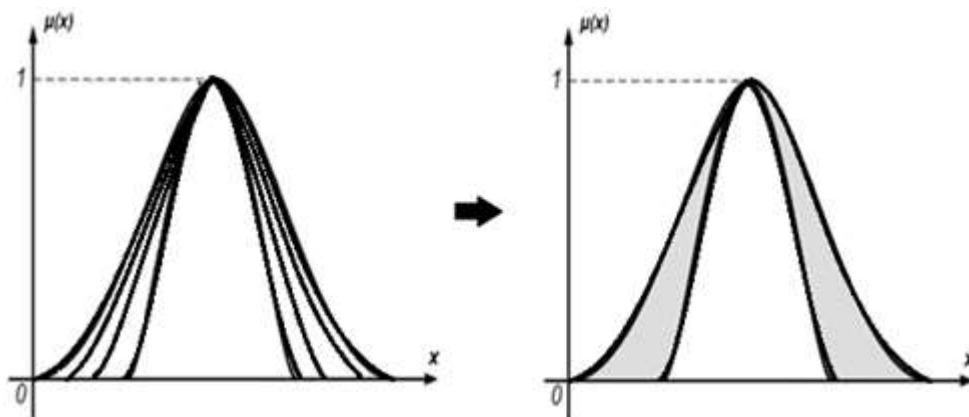


Рисунок 1 – Перехід від функції належності типу-1 до інтервальної типу-2

Побудова нечіткої моделі з функціями належності на інтервальних нечітких множинах типу-2 може мати певні особливості у випадку, коли експерт виключає малоінформативні параметри з початкового

вектору. Це означає появу пропусків в експериментальних даних та вхідному векторі. Тоді побудова нечіткої моделі з інтервальними функціями належності типу-2 виконується з корекцією наявної вибірки вихідних даних за алгоритмом [12]. А саме, видаляємо з початкової експериментальної вибірки стовпці зі значеннями змінних, які є малоінформативними. Оскільки відбувається корекція експериментальної вибірки, то будується нечітка система з інтервальними функціями належності по зменшеній експериментальній вибірці. Далі визначається вихідне значення та здійснюється тестування моделі на верифікованих даних.

Таким чином, в основі методики побудови системи нечіткої логіки для моделювання станів складного об'єкта з недовизначеними вхідними даними є нечітка модель з інтервальними функціями належності типу-2. Структуру функціонування такої моделі зображено на рисунку 2.

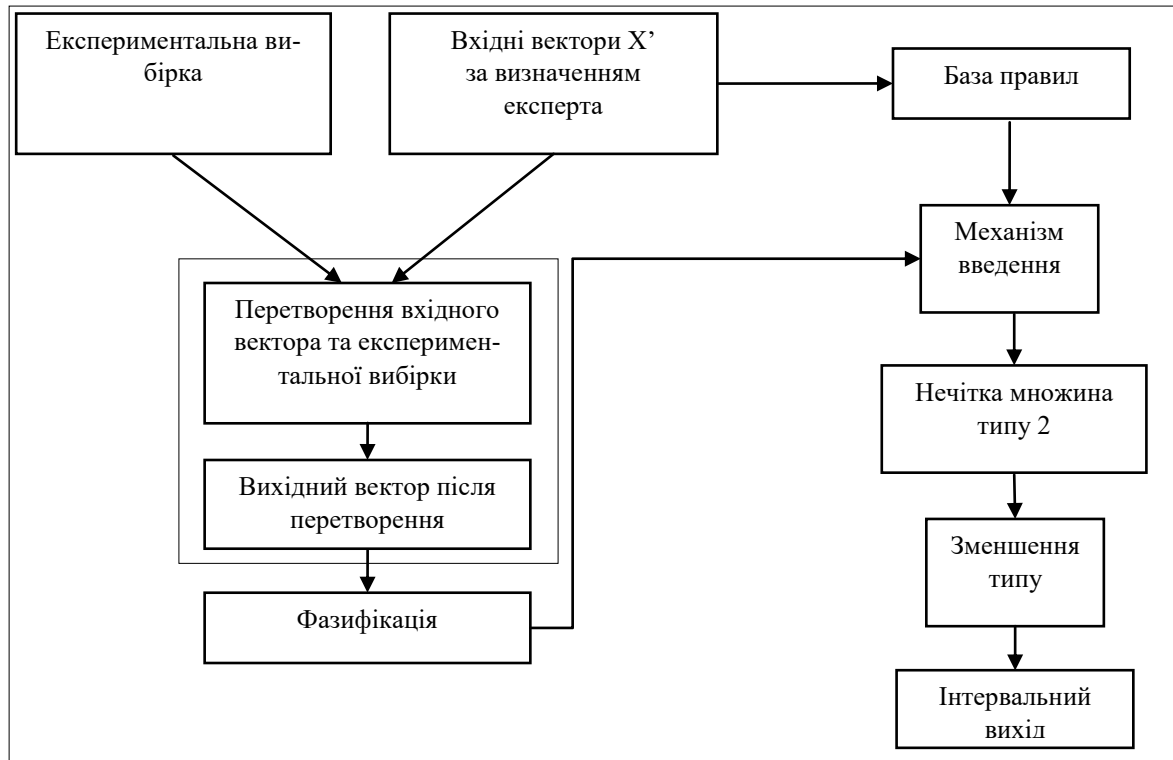


Рисунок 2 – Структура функціонування нечіткої моделі з інтервальними функціями належності

Для узагальненості оцінювання експертами вихідного інтервалу нечіткої моделі, представимо отримані результати з різних предметних галузей (наведено в таблиці 1) та скористаємось висновками, що надаються експертами.

Перша задача з технологічної галузі. Використовуємо інтервальну нечітку модель для моделювання та оцінювання перспективності артезіанської свердловини. Вихід моделі є інтервальним, кінцеву оцінку для стану свердловини виставляє експерт. Обґрунтування такого підходу полягає в тому, що свердловина – це складний природничий об'єкт, для якого не завжди можливо отримати повні вихідні дані; дослідження свердловини, а саме геологорозвідувальні роботи (ГРР) реалізуються поетапно на протязі 3-х років [9].

Друга задача має соціального спрямування, де моделюється проблема підбору персоналу в соціальних групах з точки зору професійної придатності. В таких завданнях вагому роль відіграє інформація, яку отримують в процесі опитування чи тестування. Оскільки процеси спілкування з людиною супроводжуються появою невизначеностей, таких як неясність, нечіткість в наслідок суб'єктивності людини у процесах мислення чи висловлювання, то використання нечітких систем з функціями належності на нечітких множинах типу-2 є очевидним [14].

Третя задача пов'язана з медичним діагностування захворювань ендокринної системи. Ці захворювання за оцінками фахівців [11,12] часто ховаються під маскою інших, та мають багато якісних ознак (сухість шкіри, сонливість, погіршення пам'яті та ін.). Останнє свідчить про наявні невизначеності вхідних даних. Для моделювання та оцінювання ступеня захворювання ендокринної системи людини, враховуючи існуючі невизначеності, використовується нечітка система з інтервальними функціями належності.

Основні відомості по наведеним завданням зведемо у таблицю 1.

Таблиця 1 – Основні відомості по наведеним завданням

№ задачі	Назва задачі	Кількість вхідних змінних	Бальна оцінка експерта за шириною інтервалу	Система якісних терм-множин для вихідної змінної. (Надано експертом).
1	Технологічне завдання	38	Одна вихідна величина – "Перспективність свердловини". Бальна оцінка: 0-10	Система якісних терм-множин для вихідної змінної: недостатня - Н, достатня - Д, висока - В.
2	Завдання соціального спрямування, оцінювання професійної придатності.	23	Одна вихідна величина – "Професійна придатність". Бальна оцінка: 0-10	Система якісних терм-множин: Низька (Н), НС – нижче середнього, С – середня, ВС – вище середнього, В – висока
3	Завдання медичної діагностики ендокринного захворювання.	39	Одна вихідна величина – "Важкість форми захворювання". Бальна оцінка: 0-4	Система якісних терм-множин: 1 – відсутність захворювання. 2 – другий ступінь захворювання; 3 – третій ступінь захворювання; 4 – тяжка форма захворювання.

Для оцінювання експертами результатів моделювання за фрагментами роботи нечіткої системи з інтервальним виходом в першому та другому завданні скористаємось таблицями 2 та 3. В таблиці 2 показано результати роботи моделі на недовизначених гідрогеологічних даних, наведено приклад визначення вихідного значення на основі тестової вибірки за умови наявності пропусків у вхідних даних узгоджених з експертом. В таблиці 3 - результати визначення професійної придатності окремих претендентів на робоче місце за допомогою нечіткої моделі.

Таблиця 2 – Результати роботи моделі на недовизначених гідрогеологічних даних

№	Етап (ГРР)	Визначених ознак/всього	Вихід моделі	Оцінка експерта – перспективність свердловини
1	1	10/38	[1.69; 3.12]	1/ Недостатня (Н)
	2	11/38	[1.51; 2.01]	
	3	11/38	[3.11; 3.76]	
2	1	24/38	[5.68; 6.32]	5/ Достатня (Д)
	2	24/38	[5.59; 7.26]	
	3	24/38	[5.75; 7.61]	
3	1	9/38	[6.41; 7.74]	8/ Достатня або Висока (Д,В)
	2	16/38	[5.85; 8.58]	
	3	32/38	[6.23; 8.47]	

Таблиця 3 – Результати визначення професійної придатності окремих претендентів на робоче місце за допомогою нечіткої моделі

№	x_1	x_2	x_3	...	x_{16}	x_{17}	x_{18}	x_{19}	x_{20}	x_{21}	x_{22}	x_{23}	Вихід моделі	Оцінка експерта
1	9	3	21	...	11	11	9	20	8	24	16	11	[0.69; 2.85]	НС
2	18	9	22	...	3	25	6	9	24	10	19	18	[0.46; 2.90]	Н,НС
3	3	12	12	...	4	5	10	19	14	2	16	14	[4.12; 8.33]	НС÷ВС
4	24	0	0	...	10	0	0	30	0	0	15	30	[5.01; 9.62]	В

Аналіз оцінок у вигляді лінгвістичних змінних, виставлених експертом, показує, що для підвищення якості прийняття рішень, як в першому так і в другому завданні, експерт звертає увагу на значення правої границі інтервального виходу нечіткої моделі типу-2. Також на думку експертів впливає і ширина самого інтервального виходу, завелика ширина вказує на існуючі невизначеності у вихідних даних. Зображено на рисунку 3 (а – задача 1, б – задача 2).

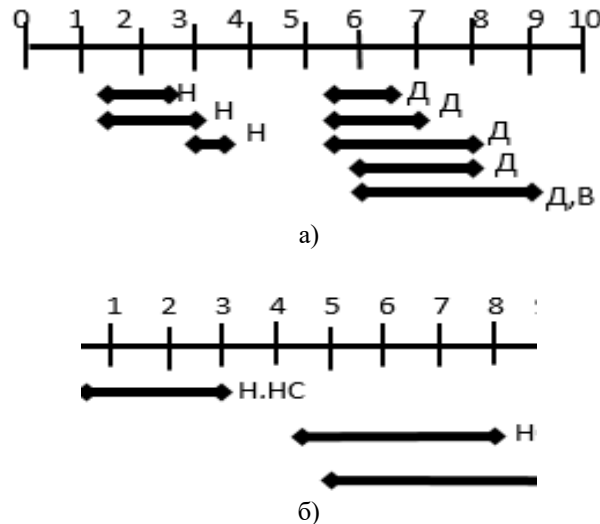


Рисунок 3 – Розташування лінгвістичних оцінок нечіткої моделі (а – задача 1, б – задача 2)

Узагальнюючи отримані оцінки експертів можливо стверджувати, що завдяки експертному оцінюванню, можливо досягнути наступне:

- в першій задачі (технологічна) - вдається скоротити матеріальні витрати на розробку свердловин та обмежити час на подальші дослідження, тобто інтервальна нечітка модель виступає в ролі порадики у прийнятті рішень стосовно перспективи подальшого дослідження свердловин;

- в другій задачі висновок експерта дозволяє вибрати найкращих претендентів на працевлаштування за рахунок того, що враховується величина невизначеності лінгвістичної оцінки «професійної придатності», яку надає експерт. При однаковому рівні «придатності» за значенням крайньої правої границі, він надає перевагу кандидату з меншим значенням невизначеності, яка обчислюється як різниця між правою та лівою границями інтервального виходу.

Треба відзначити особливість оцінювання експертом інтервального виходу нечіткої моделі в задачі медичного діагностування. Наведемо приклади оцінювання інтервального виходу експертом.

Перший випадок: інтервальний вихід показує межі одного рівня, наприклад, [2.1373; 2.2341]. Другий випадок: інтервальний вихід показує межі декількох рівнів, наприклад, [0.6760; 3.1621].

Оцінювання експертом для першого випадку: інтервальний вихід знаходиться у межах одного рівня, значення верхньої і нижньої межі інтервалу відповідають одному і тому ступеню захворювання – другому, це «середній ступень» важкості захворювання. Оцінювання експертом для другого випадку: інтервальний вихід показує межі декількох рівнів. Висновок експерта такий: значення верхньої і нижньої межі вихідного інтервалу відповідають різному ступеню захворювання. У такому випадку кінцевим рішенням будуть всі типи захворювання, які потрапляють у вихідний діапазон, тобто захворювання є але для встановлення ступеня його важкості необхідні додаткові дослідження. Таким чином, в даному завданні головним результатом роботи експерта є можливість визначити наявність захворювання ендокринної системи та призначити своєчасне лікування, яке є важливим для життя людини.

Запропонований підхід як на основі нечітких множин, так і на основі нечітких множин типу-2 дозволяє розв'язувати задачі оцінювання та моделювання для інших критичних систем, наприклад, енергетичних, де наявний високий ступінь невизначеності через такі фактори:

- відсутність достатньої інформації у відкритому друці про аналогічні підприємства інших країн, що не сприяє формування статистичної вибірки;

- унікальність та специфічність деяких критичних систем, які можуть не мати аналогів та існувати в одному екземплярі;

- висока мінливість структури загроз, що обумовлюється зростанням темпів цифровізації виробничих процесів.

Зокрема перспективу застосування ми вбачаємо в таких напрямках діяльності:

- оцінювання кібербезпеки енергетичних систем;

- виявлення шаблонів шкідливого програмного забезпечення;
- дослідження безпеки програмного коду.

Якщо розглядати математичний опис моделі в теоретико-множинному вигляді як автомат, то він буде мати такий вигляд:

$$M = \langle X, Y, S, P \rangle,$$

де X — множина векторів вхідних даних, Y — множина вихідних даних, S — множина внутрішніх станів, P — множина перетворень, що відбуваються в моделі.

Через недостатню задокументованість у відкритому друці атак на енергетичні системи задача оцінювання стану кібербезпеки для них має недовизначений характер. Більше того, роздільність в часі для таких атак за умов мінливості розвитку інформаційних технологій робить більш віддалені в часі атаки менш інформативними, оскільки за цей проміжок часу змінюються програмні засоби та технології обробки інформації. Як було показано вище, останній недолік може бути компенсований при використанні наявної статистики та залученні експертів. Вхідними даними такої моделі будуть виступати лінгвістичні параметри загроз

$$X_{grid} = \langle threat\ 1, threat\ 2, \dots, threat\ n \rangle.$$

Очікуваним виходом моделі буде лінгвістична оцінка поточного стану та їх ступінь належності

$$Y_{grid} = \langle threatLevel\ 1, threatLevel\ 2, \dots, threatLevel\ m \rangle.$$

Оскільки з плином часу та появою відомостей щодо нових атак параметри моделі потрібно буде уточнювати, тому внутрішнім станом моделі будуть саме ці параметри

$$S_{grid} = \langle params\ 1, params\ 2, \dots, params\ k \rangle.$$

Прикладом таких параметрів можуть є правила логічного висновування. Множина перетворень, які виконуються в моделі — це перетворення обробки вхідних даних, отриманих методом експертного оцінювання та/або аналізу наявних статистичних даних до єдиного вигляду, придатного для обробки в моделі.

При застосуванні в напрямку виявлення шаблонів шкідливого програмного забезпечення входами моделі X фрагменти ділянок пам'яті (файли, сектори завантаження, реєстр операційної системи тощо). Оскільки кількість входів може варіюватись для різних видів шкідливого програмного забезпечення та типу сигнатур (суцільна або розмита) [20], тому пропонується подавати на кожен вхід елементарну одиницю, а саме код операції та аргументи цієї операції:

$$X_{malware} = \langle operation\ 1, operation\ 2, \dots, operation\ n \rangle.$$

Вихідними даними Y — результати аналізу, тобто це буде лінгвістичний терм, який позначатиме вид шкідливого програмного забезпечення $\langle malwareType\ 1, malwareType\ 2, \dots, malwareType\ m \rangle$, на який схожі фрагменти, отримані на вході, або терм, що позначатиме відсутність ознак шкідливого програмного забезпечення

$$Y_{malware} = \langle malwareType\ 1, malwareType\ 2, \dots, malwareType\ m \rangle \cup \langle noMalwareIndicators \rangle.$$

Внутрішніми станами S стане набір правил висновування, побудований на основі опитування експертів та відомих баз знань антивірусного програмного забезпечення

$$S_{malware} = \langle rule\ 1, rule\ 2, \dots, rule\ k \rangle.$$

Множина перетворень передбачатиме обробку вхідних даних, їх фазифікацію та знаходження для них відповідного правила з-поміж множини $S_{malware}$.

Перевага такого застосування полягає в тому, що у зв'язку з широкоживаною практикою поліморфізму поміж шкідливим програмних забезпеченням, яка, зокрема, широко використовується

розробниками вірусів, задача виявлення ознак шкідливого програмного забезпечення повинна враховувати всі можливі модифікації. Саме тому кількість можливих видів, наприклад, одного вірусу в зараженому файлі, нелінійно зростає при лінійному зростанні кількості шляхів виконання кожної з процедур вірусу (розмноження, інфікування, виконання шкідливої дії, шифрування власного тіла тощо). Так, наприклад, при наявності 2 сценаріїв розмноження, 2 сценаріїв інфікування та 2 сценаріїв виконання шкідливої дії кількість варіантів вірусу становить $2^3 = 8$. При додаванні по одному сценарію до кожної групи, кількість варіантів зростає до $3^3 = 27$. На практиці, у кожного вірусу може бути більше трьох груп сценаріїв, відповідно кількість модифікацій (штамів) вірусів може сягати тисяч різновидів. Відповідно застосування нечітких множин дозволить в деяких випадках виявляти нові модифікації, в той час як традиційні методи сигнатурного аналізу можуть виявляти лише раніше відомі модифікації [20].

Крім пошуку шаблонів шкідливого коду у вже розроблених програмах, перспективною є модифікація моделі, коли як вхід будуть подаватись фрагменти сирцевого коду (source code), для випадків, коли аналіз відбувається в межах статичного дослідження безпеки програм (Static Application Security Testing) під час розробки програм [21].

Висновки

Запропоновано методіку використання нечітких множин типу-2 в завданнях моделювання та оцінювання станів складних об'єктів з недовизначеними вхідними даними. Показано етап оцінювання кінцевого результату інтервальних нечітких моделей з врахуванням думки експертів. Продемонстровано на прикладах прикладних задач, що на відміну від моделювання на основі нечітких множин типу 1, результатом роботи яких є ступінь належності в формі єдиного числа, нечіткі множини типу 2 дозволяють отримати на виході інтервал можливих значень вихідної лінгвістичної змінної. Ці особливості дають змогу експертам приймати більш адекватні предметній області рішення в умовах недовизначених вхідних даних.

Наведено приклади задач з різних напрямів галузі кібербезпеки (оцінювання стану кібербезпеки, аналіз шкідливого програмного забезпечення, тестування кібербезпеки програм), які є перспективними для подальшого застосування запропонованої розширеної методіки. На основі математичного апарату цифрових автоматів показано можливість її використання для розв'язання цих задач.

Список літератури

- [1] Zadeh L.A. Fuzzy sets / Zadeh L.A. // Inform. And Control. – 1965. – V8. – P. 339 - 353.
- [2] Zadeh L.A. Fuzzy sets as a basis for theory of possibility // Fuzzy sets and systems 100 supplements. – 1999. – P. 9 - 34.
- [3] Mendel, J. M. Uncertainty, fuzzy logic, and signal processing [Text] // Signal Processing Journal. 2000. – V. 80. – P. 913–933.
- [4] Liang, Q. Interval Type-2 fuzzy logic systems: theory and design [Text] / Q. Liang, J. M. Mendel // IEEE Trans. on Fuzzy Syst. – 2000. –V. 8. – P. 535–550.
- [5] Zeng, J. Type-2 Fuzzy sets for pattern classification: A review [Text] / J. Zeng, Z. Q. Liu // Proceedings of the IEEE Symposium on Foundations of computational intelligence. – 2007. – P. 193–200.
- [6] Nataliia R. Kondratenko Applications Type-2 Membership Functions in Fuzzy Logic Systems Under Conditions of Uncertainty Input Data Proceedings of XVI International Conference Measurement and Control in Complex System (MCCS-2022) <https://doi.org/10.31649/mccs2022.02>
- [7] Guerrero, Maribel. Comparative Study between Type-1 and Interval Type-2 Fuzzy Systems in Parameter Adaptation for the Cuckoo Search Algorithm [Text] / Maribel Guerrero, Fevrier Valdez and Oscar Castillo // Symmetry.- 2022, 14, 2289. <https://doi.org/10.3390/sym>
- [8] Zhina Zhang &Yugang Niu. Adaptive sliding mode control for interval type-2 stochastic fuzzy systems subject to actuator failures [Text] / Zhina Zhang &Yugang Niu // International Journal of systems Science. - 2018. <https://doi.org/10.1080/00207721.2018.1534027>
- [9] Kondratenko, N. Interval Fuzzy Modeling of Complex Systems under Conditions of Input Data Uncertainty [Text] / N. Kondratenko, O. Snihur // Eastern-European Journal of Enterprise Technologies. – 2016. – V. 4/4 (82). – P. 20–28.
- [10] Kondratenko, N. Interval type-2 generalizing fuzzy model for monitoring the states of complex systems using experts knowledge [Text] / N. Kondratenko, O. Snihur, R. Kondratenko // System Research And Information Technologies. – 2023. – No.2. DOI: <https://doi.org/10.20535/SRIT.2308-8893.2023.2.05>
- [11] Кондратенко Н.Р. Використання нечітких баз знань з функціями належності типу-2 у медичній діагностиці / Н.Р. Кондратенко // Матеріали статей Міжнародної науково-практичної

конференції «Актуальні задачі медичної, біологічної фізики та інформатики». - 2022р.-
Вінниця:

<https://drive.google.com/file/d/1icajVT7OKyVxlfXZd1czwhS13EHtmUj8/view?usp=sharing>

- [12] Кондратенко Н.Р., Зелінська Н.Б., Куземко С.М. Нечіткі логічні системи з врахуванням пропусків в експериментальних даних // Наукові вісті НТУУ «КПІ». - 2004. – № 5. – С. 37–41.
- [13] Кондратенко Н.Р., Чеборака О.В. Дослідження можливостей узагальнювальної інтервальної типу-2 нечіткої моделі для прогнозування часових послідовностей // Вісник Вінницького політехнічного інституту. - 2008. – № 6. – С. 22–27.
- [14] Кондратенко, Н.Р. Нечіткі моделі в задачах підбору персоналу при формуванні соціальних груп [Текст] / Н.Р.Кондратенко, С.В. Лужецький, О.В. Чеборака //Системні дослідження та інформаційні технології. – 2011. - № 3.- С.56-62.
- [15] Кондратенко, Н. Р. Підвищення адекватності нечітких моделей за рахунок використання нечітких множин типу 2 [Текст] // Наукові вісті національного технічного університету України «Київський політехнічний інститут». – 2014. – №6.
- [16] I. Kotsiuba, I. Skarga-Bandurova, A. Giannakoulis, M. Chaikin and A. Jevremovic. Technique for Finding and Investigating the Strongest Combinations of Cyberattacks on Smart Grid Infrastructure, 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 4265-4272, doi: 10.1109/BigData47090.2019.9006335.
- [17] T. Pléta, M. Tvaronavičienė, S. Della Casa, K. Agafonov. Cyber-attacks to critical energy infrastructure and management issues: overview of selected cases. Insights into Regional Development, 2020, 2 (3), pp. 703 - 715. URL: <https://hal.science/hal-03271856/> (accessed: 08.09.2023).
- [18] Кондратенко Н.Р., Куземко С. М. Нечіткі логічні системи з використанням нечітких множин загального типу // Наукові вісті НТУУ «КПІ». - 2004. – № 1. – С. 16–21.
- [19] Кондратенко, Н. Р. Дослідження адекватності інтервальних нечітких моделей типу-2 в задачах ідентифікації складних об'єктів [Текст] / Н. Р. Кондратенко, О. О. Снігур // Системні дослідження та інформаційні технології. – 2019. – № 4. – С. 94–104.
- [20] Льїн М. І., Якобчук Д. І. Зворотна розробка та аналіз шкідливого програмного забезпечення: Лабораторний практикум: навчальний посібник для студентів спеціальностей 125 «Кібербезпека», 113 «Прикладна математика». Київ: НТУУ «КПІ імені Ігоря Сікорського», 2020. 117 с.
- [21] T. D. Oyetoyan, B. Milosheska, M. Grini, D. S. Cruze. Myths and Facts About Static Application Security Testing Tools: An Action Research at Telenor Digital. Agile Processes in Software Engineering and Extreme Programming 19th International Conference, XP 2018 Porto, Portugal, May 21–25, 2018 Proceedings. pp. 86-103 URL: <https://library.oapen.org/bitstream/handle/20.500.12657/27860/1002144.pdf?sequence=1#page=98> (accessed: 08.09.2023).

Стаття надійшла: 15.04.2023.

References

- [1] Zadeh L.A. Fuzzy sets / Zadeh L.A. // Inform. And Control. – 1965. – V8. – P. 339 - 353.
- [2] Zadeh L.A. Fuzzy sets as a basis for theory of possibility / Zadeh L.A. // Fuzzy sets and systems 100 supplements. – 1999. – P. 9 - 34.
- [3] Mendel, J. M. Uncertainty, fuzzy logic, and signal processing [Text] // Signal Processing Journal. 2000. – V. 80. – P. 913–933.
- [4] Liang, Q. Interval Type-2 fuzzy logic systems: theory and design [Text] / Q. Liang, J. M. Mendel // IEEE Trans. on Fuzzy Syst. – 2000. –V. 8. – P. 535–550.
- [5] Zeng, J. Type-2 Fuzzy sets for pattern classification: A review [Text] / J. Zeng, Z. Q. Liu // Proceedings of the IEEE Symposium on Foundations of computational intelligence. – 2007. – P. 193–200.
- [6] Nataliia R. Kondratenko Applications Type-2 Membership Functions in Fuzzy Logic Systems Under Conditions of Uncertainty Input Data Proceedings of XVI International Conference Measurement and Control in Complex System (MCCS-2022) <https://doi.org/10.31649/mccs2022.027>
- [7] Guerrero, Maribel. Comparative Study between Type-1 and Interval Type-2 Fuzzy Systems in Parameter Adaptation for the Cuckoo Search Algorithm [Text] / Maribel Guerrero, Fevrier Valdez and Oscar Castillo // Symmetry.- 2022, 14, 2289. <https://doi.org/10.3390/sym>

- [8] Zhina Zhang &Yugang Niu. Adaptive sliding mode control for interval type-2 stochastic fuzzy systems subject to actuator failures [Text] / Zhina Zhang &Yugang Niu // International Journal of systems Science. - 2018. <https://doi.org/10.1080/00207721.2018.1534027>
- [9] Kondratenko, N. Interval Fuzzy Modeling of Complex Systems under Conditions of Input Data Uncertainty [Text] / N. Kondratenko, O. Snihur // Eastern-European Journal of Enterprise Technologies. – 2016. – V. 4/4 (82). – P. 20–28.
- [10] Kondratenko, N. Interval type-2 generalizing fuzzy model for monitoring the states of complex systems using experts knowledge [Text] / N. Kondratenko, O. Snihur, R. Kondratenko // System Research And Information Technologies. – 2023. – No. 2. DOI: <https://doi.org/10.20535/SRIT.2308-8893.2023.2.05>
- [11] Kondratenko N.R. Vykorystannya nechitkykh baz znanj z funktsiyamy nalezhnosti typu-2 u medychnyi diahnostytsi / N.R. Kondratenko // Materialy mizhnar. nauk.-prakt. konf., «Aktual'ni zadachi medychnoyi biolohichnoyi fizyky ta informatyky»: - 2022p.:.- Vinnytsya, rezhym dostupu: <https://drive.google.com/file/d/1icajVT7OKyVxlfXZd1czwhS13EHtmUj8/view?usp=sharing>
- [12] Kondratenko, N. R. Fuzzy Logic Systems with Allowance for the Blank in Experimental Data Taken [Text] / N. R. Kondratenko, N. B. Zelinsjka, S.M. Kuzemko // Naukovi visti NTUU KPI.- 2004.- No. 5. P. 37–41.
- [13] Kondratenko, N. R. Study of Aggregating Interval Type-2 Fuzzy Models Capabilities for Forecasting of Time Series [Text] / N.R. Kondratenko, O. V. Cheboraka, S.M.Kuzemko // Visnyk VPI. - 2010. – No. 4. – P. 22–27.
- [14] Kondratenko, N.R. Fuzzy models in staff recruitment problems during the social groups forming [Text] / N.R.Kondratenko, S.V. Luzhetskiy, O.V. Cheboraka // System Research And Information Technologies. – 2011. – No 3. – P. 56–62.
- [15] Kondratenko, N. R. Improving Adequacy of Type-2 Fuzzy Models by Using Type-2 Fuzzy Sets [Text] / N. Kondratenko // Naukovi visti NTUU KPI. – 2014. – No. 6. – P. 56–61.
- [16] I. Kotsiuba, I. Skarga-Bandurova, A. Giannakoulis, M. Chaikin and A. Jevremovic. Technique for Finding and Investigating the Strongest Combinations of Cyberattacks on Smart Grid Infrastructure, 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 4265-4272, doi: 10.1109/BigData47090.2019.9006335.
- [17] T. Plėta, M. Tvaronavičienė, S. Della Casa, K. Agafonov. Cyber-attacks to critical energy infrastructure and management issues: overview of selected cases. Insights into Regional Development, 2020, 2 (3), pp. 703 - 715. URL: <https://hal.science/hal-03271856/> (accessed: 08.09.2023).
- [18] Kondratenko, N.R. Fuzzy Logic Systems with the use of general type fuzzy sets [Text] / N.R. Kondratenko, S.M. Kuzemko // Naukovi visti NTUU KPI, - 2004.- No.1.- P.16-21.
- [19] Kondratenko, N. R. Investigating adequacy of interval type-2 fuzzy models in complex objects identification problems [Text] / N. R. Kondratenko, O. O. Snihur // System Research And Information Technologies. – 2019. – NO 4. – P. 94–104.
- [20] Ilin, M., Yakobchuk, D. Reverse Engineering and Malware Analyses: Laboratory works' manual for students of 125 «Cybersecurity», 113 «Applied Mathematics» curricula. Kyiv, NTUU «Igor Sikorsky Kyiv Polytechnic Institute», 2020. 117 p.
- [21] T. D. Oyetoyan, B. Miloshevska, M. Grini, D. S. Cruze. Myths and Facts About Static Application Security Testing Tools: An Action Research at Telenor Digital. Agile Processes in Software Engineering and Extreme Programming 19th International Conference, XP 2018 Porto, Portugal, May 21–25, 2018 Proceedings. pp. 86-103 URL: <https://library.oapen.org/bitstream/handle/20.500.12657/27860/1002144.pdf?sequence=1#page=98> (accessed: 08.09.2023).

Відомості про авторів

Барішев Юрій Володимирович – к. т. н., доцент кафедри захисту інформації Вінницького національного технічного університету, м. Вінниця.

Кондратенко Наталія Романівна – к. т. н., професор кафедри захисту інформації Вінницького національного технічного університету, м. Вінниця.

Казміревський Віталій Віталійович – аспірант кафедри захисту інформації, факультет інформаційних технологій та комп'ютерної інженерії Вінницького національного технічного університету, м. Вінниця.

Кирилащук Тетяна Геннадіївна – асистент кафедри захисту інформації, факультет інформаційних технологій та комп'ютерної інженерії Вінницького національного технічного університету, м. Вінниця.

Y. V. Baryshev, N. R Kondratenko, V. V. Kazmirevskyi, T. G. Kyrylashchuk
**TYPE-2 FUZZY SETS IN THE TASKS OF MODELLING AND
ESTIMATING OF CRITICAL SYSTEMS'S STATES WITH
UNCERTAIN INPUT DATA AND THE USAGE OF EXPERTS**

Vinnitsia National Technical University, Vinnitsia

5