DOI: 10.58254/viti.5.2024.15.165

УДК 004.056(53+57)

д-р техн. наук, професор Субач І. Ю. ORCID: 0000-0002-9344-713X
(ІСЗЗІ НТУУ «КПІ ім. Ігоря Сікорського»)
Власенко О. В. ORCID: 0000-0001-6671-870X (ВІТІ ім. Героїв Крут)

НЕЧІТКІ МОДЕЛІ ВИЯВЛЕННЯ КІБЕРІНЦИДЕНТІВ У БАЗАХ ДАНИХ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ СИСТЕМ ВІЙСЬКОВОГО ПРИЗНАЧЕННЯ

Захист баз даних інформаційно-комунікаційних систем військового призначення є надзвичайно важливим завданням у сучасній сфері кібербезпеки. Зростаючі загрози від кібератак, необхідність ефективного виявлення, протидії та запобігання їм вимагають застосування нових, більш ефективних моделей та методів. Основні недоліки існуючих моделей і методів включають недостатню чутливість до нових загроз, велику кількість помилок виявлення, низьку відповідь на нові загрози, можливість обходу захисних заходів та низьку масштабованість, що є ключовими викликами для подальшого вдосконалення та розвитку кібербезпеки. У статті проведено аналіз існуючих нечітких моделей виявлення кіберінцидентів, виокремлено їхні недоліки та наголошено на необхідності їхнього подальшого удосконалення та розвитку. Запропоновано удосконалену нечітку модель виявлення кіберінцидентів у базах даних інформаційно-комунікаційних систем військового призначення та удосконалену нечітку модель виявлення кіберінцидентів у базах даних інформаційно-комунікаційних систем військового призначення зі зваженими правилами, на основі розширення ознак кіберінцидентів шляхом отримання їх з різних рівнів кіберзахисту баз даних. До основних рівнів кіберзахисту баз даних потрібно віднести: рівень операційної системи, рівень мережі та рівень системи керування базами даних. Для усунення недоліків, пов'язаних з помилковим спрацьовуванням правил виявлення кіберінцидентів та складністю їх налаштування в умовах ландшафту кібератак, що динамічно змінюється, а також розмірністю бази знань системи управління інформацією та подіями безпеки, запропоновано нечітку модель виявлення кіберінцидентів у базах даних інформаційно-комунікаційних систем військового призначення із вагами антецедентів правил. Показано доцільність застосування розробленої моделі.

Ключові слова: *база даних, інформаційно-комунікаційна система, кіберзахист, кіберінцидент, SIEM-система, теорія нечітких множин, нечіткі правила.*

I. Subach, O. Vlasenko Fuzzy models for cyber incident detection in military information and communication systems databases

Protecting databases of military information and communication systems is an extremely important task in the modern cybersecurity sphere. Growing threats from cyberattacks, the need to effectively detect, counteract and prevent them require the use of new, more effective models and methods. The main disadvantages of existing models and methods include insufficient sensitivity to new threats, a large number of detection errors, low response to new threats, the possibility of bypassing protective measures, and low scalability, which are key challenges for further improvement and development of cybersecurity. The article analyzes the existing fuzzy models for detecting cyber incidents, identifies their shortcomings and emphasizes the need for their further improvement and development. An improved fuzzy model for detecting cyber incidents in databases of military information and communication systems and an improved fuzzy model for detecting cyber incidents in databases of military information and communication systems with weighted rules based on the expansion of cyber incident signs by obtaining them from different levels of cyber security of the data are proposed. The main levels of database cybersecurity include: the operating system level, the network level, and the database management system level. To eliminate the shortcomings associated with the false triggering of cyber incident detection rules and the complexity of their configuration in a dynamically changing cyberattack landscape, as well as the dimensionality of the knowledge base of the information and security event management system, a fuzzy model for detecting cyber incidents in databases of military information and communication systems with weights of rule antecedents is proposed. The expediency of applying the developed model is shown.

Keywords: *database, information and communication system, cyber protection, cyber incident, SIEM system, fuzzy set theory, fuzzy rules.*

Постановка завдання. З кожним днем цифрові технології підтверджують свою критичну роль у військовій справі, а отже питання кібербезпеки стає все більш актуальнішим. Через стрімкий розвиток інформаційних технологій та діджиталізацію процесів військової діяльності основою для координації військових операцій, прийняття стратегічних рішень та забезпечення національної безпеки стають інформаційно-комунікаційні системи військового

призначення (далі – ІКСВП) [1]. Зі зростанням значення ІКСВП у сучасній військовій сфері збільшується і загроза кібератак на них, що може знизити ефективність та надійність проведення військових операцій у цілому. Внаслідок цього особливої ваги набуває важливість розвитку методів та засобів кіберзахисту ІКСВП.

Сучасні кібератаки, використовуючи передові технології й методи, стають все більш складними та вишуканими, що надзвичайно ускладнює їх виявлення та запобігання їм. Зловмисники постійно вдосконалюють свої навички, використовуючи нові методи інтелектуального зламу, соціальної інженерії та експлуатацію нових вразливостей програмного забезпечення.

Тому, для ефективного захисту від сучасних кібератак необхідно постійно розвивати та вдосконалювати методи та засоби кіберзахисту, враховуючи нові виклики, що виникають в інформаційній інфраструктурі органів військового управління.

Аналіз наукових публікацій. Швидкий та постійний розвиток інформаційних технологій супроводжується ростом кількості кібератак, які ставлять під загрозу конфіденційність, цілісність та доступність інформації. У цьому контексті системи управління інформацією та подіями безпеки (далі – SIEM-системи) є одними з ключових інструментів в побудові екосистеми кіберзахисту ІКСВП [2]. SIEM-система поєднує в собі функціонал аналізу, виявлення та реагування на події безпеки, забезпечуючи комплексний підхід до захисту інформації та інфраструктури від кібератак. Ефективність SIEM-систем роблять їх важливим інструментом для будь-якого підрозділу, який прагне забезпечити безпеку ІКСВП.

У роботі [3] досліджено та аргументовано важливість використання інтелектуальних SIEM-систем для створення ефективної системи кіберзахисту інформаційно-комунікаційних систем (далі – ІКС). Архітектуру інтелектуальної SIEM-системи для виявлення кіберінцидентів у базах даних (далі – БД) наведено у публікації [4]. Ключовим компонентом SIEM-системи, який відповідає за обробку та аналіз подій безпеки, є підсистема аналізу даних. Вона включає в себе функції виявлення аномалій, ідентифікації потенційно небезпечних подій, а також кореляції даних із різних джерел. У цій підсистемі застосовуються унікальні високоспеціалізовані алгоритми та складні методи аналізу, які переважно недоступні та вважаються комерційною таємницею. Основними функціями, які покладаються на таку підсистему, є наступні:

Виявлення кібератак/кіберінцидентів. Адміністратори можуть налаштовувати правила, що визначають певні ситуації або шаблони поведінки користувачів, які можуть вказувати на потенційні загрози функціонуванню системи керування базами даних (далі – СКБД). Наприклад, використання несправжніх облікових записів, спроби неуспішного входу для роботи з БД системи, зміни конфігурації СКБД тощо, що допомагає виявляти кібератаки та кіберінциденти, пов’язані з роботою БД, а також надає інформацію про їхні наслідки.

Машинне навчання та аналіз відхилень. Деякі сучасні SIEM-системи використовують методи машинного навчання для виявлення аномальних патернів, які можуть вказувати на зловмисну діяльність.

Кореляція подій. Здійснюється аналіз і кореляція події з різних джерел для виявлення складних атак на БД або зловмисних дій, які можуть бути приховані в окремих записах.

Профілювання користувачів і програмних застосунків. Здійснення аналізу поведінки користувачів і програмних застосунків у мережі для виявлення їхніх незвичних або підозрілих дій по відношенню до БД та СКБД.

Шляхом аналізу інформації про події безпеки та використання методів інженерії знань і штучного інтелекту SIEM-система може прогнозувати майбутні кіберінциденти та допомагати у прийнятті оперативних, обґрунтованих рішень щодо запобігання їм [5]. Це досягається завдяки розробці та впровадженню нових або удосконалених моделей і методів виявлення та

ідентифікації кіберінцидентів, що відбуваються у СКБД та пов'язаних з несанкціонованим доступом до БД ІКСВП.

Існуючи у теперішній час найефективніші моделі та методи виявлення кіберінцидентів стають застарілими у швидкозмінному кіберпросторі. Стрімке зростання кількості та складності кіберзагроз вимагає постійної адаптації та покращення захисних стратегій [6].

Незважаючи на велику кількість наукових досліджень, присвячених використанню методів штучного інтелекту, машинного навчання та інших новітніх інформаційних технологій у SIEM-системах для виявлення кіберінцидентів [7–11], питання організації знань у базах знань інтелектуальних SIEM-систем та їх використання для ідентифікації кіберінцидентів, пов'язаних з БД ІКСВП, зокрема в умовах неповноти та невизначеності інформації, залишаються не повністю вирішеними.

Метою статті є побудова моделі виявлення кіберінцидентів, пов'язаних з базами даних інформаційно-комунікаційних систем військового призначення на основі застосування моделей і методів теорії нечітких множин та лінгвістичних термів.

Виклад основного матеріалу дослідження. Аналіз моделей і методів нечіткої ідентифікації кіберінцидентів [12; 13], показує, що рішення задачі щодо їхнього розпізнавання полягає у знаходженні відображення (1):

$$F^* = (f_1^*, f_2^*, \dots, f_n^*) \to c_j \in C = (c_1, c_2, \dots, c_m), \quad (1)$$

де F^* – множина ознак кіберінциденту, пов'язаного з функціонуванням БД ІКСВП;

C – множина можливих кіберінцидентів.

Проте в [14] показано, що ефективність рішення задачі виявлення кіберінцидентів у БД ІКСВП підвищується шляхом застосування багаторівневого кіберзахисту БД від кібератак (рис. 1).

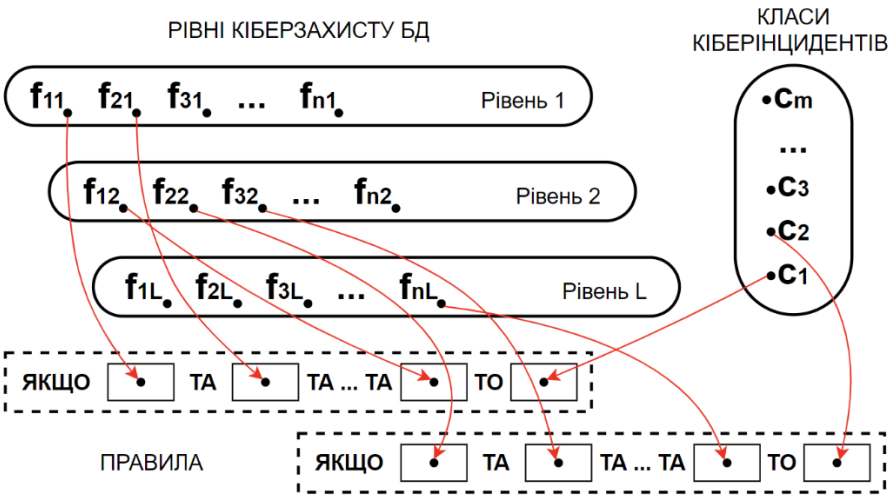

Рис. 1. Розширення ознак кіберінцидентів на основі багаторівневого кіберзахисту БД

Тоді задача розпізнавання кіберінцидентів у БД ІКСВП прийме вигляд (2):

$$F_l^* = (f_{1l}^*, f_{2l}^*, \dots, f_{nl}^*) \to c_j \in C = (c_1, c_2, \dots, c_m), \quad (2)$$

де F_l^* – множина ознак кіберінциденту, пов'язаного з функціонуванням БД ІКСВП, отриманих з різних рівнів кіберзахисту БД [14].

Область зміни ознак кіберінцидентів $f_{il} \in \left[\underline{f_{il}}, \overline{f_{il}}\right],\ i = 1 \dots n,\ l = 1 \dots L$, отриманих на різних рівнях кіберзахисту БД $l = 1 \dots L$ і вихідного значення ідентифікації $c_i \in \left[\overline{c_j}, \underline{c_j}\right]$, $j = 1 \dots m$.

Відповідно $\underline{f_{il}}, (\overline{f_{il}})$ – нижнє (верхнє) значення ознак кіберінциденту f_{il}, $\underline{c_j}$ $(\overline{c_j})$ – нижнє (верхнє) значення результату ідентифікації c_j.

Для вирішення задачі, вхідні і вихідні змінні розглядаються як лінгвістичні змінні, що задані на універсальних множинах [15; 16]:

$$f_{il} = \left[\underline{f_{il}}, \overline{f_{il}}\right],\ c_j = \left[\underline{c_j},\ \overline{c_j}\right]. \tag{3}$$

Для оцінки (3) доцільно використовувати якісні терми, які входять до множин термів:

$A_{il} = \{a_{il}^1, a_{il}^2, \dots, a_{il}^k\}$ – терм-множина змінної f_{il}, де a_{il}^k – k-й лінгвістичний терм змінної f_{il}, $k = 1 \dots k_i,\ i = 1 \dots n$ l-го рівня кіберзахисту БД $l = 1 \dots L$;

$\Delta = \{\delta_1, \delta_2, \dots, \delta_m\}$ – терм-множина змінної c_j, де $\delta_j,\ j = 1 \dots m$ – лінгвістичний терм змінної c_j, m – число можливих класів кіберінцидентів, пов'язаних із БД.

Таким чином, лінгвістичні терми $a_{il}^k \in A_{il}, k = 1 \dots k_i,\ i = 1 \dots n,\ l = 1 \dots L$ та $\delta_j \in \Delta$, $j = 1 \dots m$ можна розглядати як нечіткі множини, які задані на універсальних множинах f_{il}, c_j.

У свою чергу, нечіткі множини a_{il}^k та δ_j можна визначити так [17–19]:

$$\boldsymbol{a_{il}^k} = \int_{\underline{f_{il}}}^{\overline{f_{il}}} \boldsymbol{\mu^{\alpha_{il}^k}(f_{il})/f_{il}}\ , \tag{4}$$

$$\delta_j = \int_{\underline{c_j}}^{\overline{c_j}} \mu^{\delta_j}(c_j)/c_j\ , \tag{5}$$

де $\mu^{\alpha_{il}^k}(f_{il})$ – функція належності значення змінної $f_{il} \in \left[\underline{f_{il}}, \overline{f_{il}}\right],\ i = 1 \dots n,\ l = 1 \dots L$ терму $\alpha_{il}^k \in A_{il},\ k = 1 \dots k_i,\ i = 1 \dots n,\ l = 1 \dots L$;

$\mu^{\delta_j}(c_j)$ – функція належності значення змінної $c_j = \left[\underline{c_j}, \overline{c_j}\right]$ терму – класу кіберінциденту, пов'язаному з БД $\delta_j \in \Delta, = 1 \dots m$.

Зауважимо, що у виразах (4) та (5) знак інтегралу означає об'єднання пар $\mu(\omega)/\omega$ [20; 21]. Експертні дані можуть бути представлені у вигляді багатовимірної матриці знань про кіберінциденти (табл. 1). Вона має наступні властивості [15; 16]:

кожний рядок матриці є комбінацією вхідних значень ознак кіберінцидентів, що пов'язані з функціонуванням БД ІКСВП $f_{il},\ i = 1 \dots n, l = 1 \dots L$ та які належать різним рівням безпеки БД, яка віднесена експертом до одного з його класів δ_j, причому перші 11 рядків відповідають класу δ_1, а останні $1m$ рядків – класу δ_m;

перші n стовпчиків матриці відповідають вхідним значенням ознак кіберінцидентів з різних рівнів кіберзахисту БД $f_{il},\ i = 1 \dots n,\ l = 1 \dots L$, а $(n+1)$-ий стовпчик відповідає вихідному значенню – класу кіберінциденту c, пов'язаному з БД;

на перетині i-го стовпчика та jk_j-го рядку знаходиться елемент $\alpha_{il}^{jk_j}$, який відповідає лінгвістичній оцінці ознаки кіберінциденту f_{il} у рядку матриці jk_j, яка належить терм-множині відповідної ознаки f_{il}: $\alpha_{il}^{k} \in A_i,\ k = 1 \dots k_i,\ i = 1 \dots n,\ l = 1 \dots L.$

Таблиця 1

Багатовимірна таблиця ознак кіберінцидентів і класів, що їм відповідають

Номер вхідної комбінації значень	Ознаки кіберінцидентів, отримані з різних рівнів кіберзахисту БД						Клас кіберінциденту
	j_1	j_2	…	j_i	…	j_n	c
11	α_{1l}^{11}	α_{2l}^{11}	…	α_{il}^{11}	…	α_{nl}^{11}	δ_1
12	α_{1l}^{12}	α_{2l}^{12}	…	α_{il}^{12}	…	α_{nl}^{12}	
…	…	…	…	…	…	…	
1k₁	$\alpha_{1l}^{1k_1}$	$\alpha_{1l}^{1k_2}$	…	$\alpha_{il}^{1k_1}$	…	$\alpha_{nl}^{1k_1}$	
…	…	…	…	…	…	…	…
j1	α_{1l}^{j1}	α_{2l}^{j1}	…	α_{il}^{j1}	…	α_{nl}^{j1}	δ_j
j2	α_{1l}^{j2}	α_{2l}^{j2}	…	α_{il}^{j2}	…	α_{nl}^{j2}	
…	…	…	…	…	…	…	
jkⱼ	$\alpha_{1l}^{jk_j}$	$\alpha_{2l}^{jk_j}$	…	$\alpha_{il}^{jk_j}$	…	$\alpha_{nl}^{jk_j}$	
…	…	…	…	…	…	…	…
m_1	α_{1l}^{m1}	α_{2l}^{m1}	…	α_{il}^{m1}	…	α_{nl}^{m1}	δ_m
m_2	α_{1l}^{m2}	α_{2l}^{m2}	…	α_{il}^{m2}	…	α_{nl}^{m2}	
…	…	…	…		…		
mk_m	$\alpha_{1l}^{mk_m}$	$\alpha_{2l}^{mk_m}$	…	$\alpha_{il}^{mk_m}$	…	$\alpha_{nl}^{mk_m}$	

Описана вище матриця знань про кіберінциденти, пов'язані з БД ІКСВП, може бути представлена у вигляді системи нечітких правил виду «ЯКЩО – ТО» [15; 16], які зв'язують значення вхідних ознак кіберінцидентів $f_{il},\ i = 1 \dots n,\ \ l = 1 \dots L$ з одним з їхніх можливих класів $\delta_j \in \Delta,\ j = 1 \dots m$:

$$
\begin{aligned}
&\text{ЯКЩО } (f_{il} = \alpha_{1l}^{11}) \text{ ТА } (f_{2l} = \alpha_{2l}^{11}) \text{ ТА…ТА } (f_{nl} = \alpha_{nl}^{11}) \text{ АБО}\\
&\quad (f_{il} = \alpha_{1l}^{12}) \text{ ТА } (f_{2l} = \alpha_{2l}^{12}) \text{ ТА…ТА } (f_{nl} = \alpha_{nl}^{12}) \text{ АБО}\\
&\quad (f_{il} = \alpha_{1l}^{1k_1}) \text{ ТА } (f_{2l} = \alpha_{2l}^{1k_1}) \text{ ТА…ТА } (f_{nl} = \alpha_{nl}^{1k_1}), \text{ ТО (с=}\delta_1\text{),…}\\
&\text{…, ЯКЩО } (f_{il} = \alpha_{1l}^{j1}) \text{ ТА } (f_{2l} = \alpha_{2l}^{j1}) \text{ ТА…ТА } (f_{nl} = \alpha_{nl}^{j1}) \text{ АБО}\\
&\quad (f_{il} = \alpha_{1l}^{j2}) \text{ ТА } (f_{2l} = \alpha_{2l}^{j2}) \text{ ТА…ТА } (f_{nl} = \alpha_{nl}^{j2}) \text{ АБО}\\
&\quad (f_{il} = \alpha_{1l}^{k_j}) \text{ ТА } (f_{2l} = \alpha_{2l}^{jk_j}) \text{ ТА…ТА } (f_{nl} = \alpha_{nl}^{jk_j}), \text{ ТО (с=}\delta_j\text{),…}\\
&\text{…, ЯКЩО } (f_{il} = \alpha_{1l}^{m1}) \text{ ТА } (f_{2l} = \alpha_{2l}^{m1}) \text{ ТА…ТА } (f_{nl} = \alpha_{nl}^{m1}) \text{ АБО}\\
&\quad (f_{il} = \alpha_{1l}^{m2}) \text{ ТА } (f_{2l} = \alpha_{2l}^{m2}) \text{ ТА…ТА } (f_{nl} = \alpha_{nl}^{m2}) \text{ АБО}\\
&\quad (f_{il} = \alpha_{1l}^{mk_m}) \text{ ТА } (f_{2l} = \alpha_{2l}^{mk_m}) \text{ ТА…ТА } (f_{nl} = \alpha_{nl}^{mk_m}), \text{ ТО (с=}\delta_m\text{),}
\end{aligned}
\tag{6}
$$

де α_{il}^{jk} – лінгвістична оцінка ознаки кіберінциденту $f_{il},\ i = 1 \dots n,\ l = 1 \dots L$ у рядку k j-ої диз'юнкції, що визначається на терм-множині $A_i = \{\alpha_{il}^1, \alpha_{il}^2, \dots, \alpha_{il}^{k_i}\}$;

$\delta_j \in \Delta,\ j = 1 \dots m$ – лінгвістична оцінка класу кіберінциденту, пов'язаного з БД, що визначається на терм-множині $\Delta = \{\delta_1, \delta_2, \dots, \delta_m\}$.

Таким чином, вираз (6), заданий у вигляді сукупності нечітких правил виду «ЯКЩО – ТО», які ґрунтуються на матриці знань про кіберінциденти (табл.1), являє собою

модель нечіткої ідентифікації кіберінцидентів, пов'язаних з БД ІКСВП інтелектуальною SIEM-системою з урахуванням різних рівнів кіберзахисту БД.

Якщо лінгвістичні оцінки α_{il}^{jk} змінних $f_{1l},\ f_{2l}, \dots ,\ f_{nl}$ та $\delta_j,\ j = 1 \dots m$ з (6) розглянути як нечіткі множини, що визначені на універсальних множинах $f_{il} = \left[\underline{f_{il}}, \overline{f_{il}}\right], c = \left[\underline{c_j}, \overline{c_j}\right]$, $i = 1 \dots n,\ j = 1 \dots m,\ \ l = 1 \dots L,$ то $\mu^{\alpha_{il}^{jk}}$ – функція належності ознаки кіберінциденту $f_{il} = \left[\underline{f_{il}}, \overline{f_{il}}\right]$ нечіткому терму $\alpha_{il}^{jk},\ i = 1 \dots n,\ j = 1 \dots m, k = 1 \dots k_i,\ \ \ l = 1 \dots L,$ а $\mu^{\delta_j}(f_{1l}, f_{2l}, \dots, f_{nl})$ – функція належності вектора ознак кіберінцидентів, пов'язаних з БД $F = \{f_{1l}, f_{2l}, \dots, f_{nl}\},\ i = 1 \dots n,$ *l=1…L*, значенню вихідної оцінки $c = \delta_j,\ j = 1 \dots m$ [15–17].

Зв'язок між ними визначається через НМЗ про кіберінциденти (табл. 1) та шляхом заміни лінгвістичних термів на їхні функції належності, а також заміни логічних операцій ТА чи АБО на операції ∧ та ∨ може бути представленим у наступному вигляді [15; 16]:

$$\begin{aligned} \mu^{\delta_j}(f_{1l}, f_{2l}, \dots, f_{nl}) = &\left[\mu^{\alpha_{1l}^{j1}}(f_{1l}) \wedge \mu^{\alpha_{2l}^{j1}}(f_{2l}) \wedge \dots \wedge \mu^{\alpha_{nl}^{j1}}(f_{nl})\right] \vee \\ \vee &\left[\mu^{\alpha_{1l}^{j2}}(f_{1l}) \wedge \mu^{\alpha_{2l}^{j2}}(f_{2l}) \wedge \dots \wedge \mu^{\alpha_{nl}^{j2}}(f_{nl})\right] \vee \\ \vee &\left[\mu^{\alpha_{1l}^{jk_j}}(f_{1l}) \wedge \mu^{\alpha_{2l}^{jk_j}}(f_{2l}) \wedge \dots \wedge \mu^{\alpha_{nl}^{jk_j}}(f_{nl})\right],\ j = 1 \dots m,\ l=1 \dots L. \end{aligned} \tag{7}$$

Шляхом згортання вираз (6) може бути представлено наступним чином:

$$\mu^{\delta_j}(f_{1l}, f_{2l}, \dots, f_{nl}) = \ \bigvee_{k=1}^{k_j}\left[\bigwedge_{i=1}^{n} \mu^{\alpha_{il}^{jk}}(f_{il})\right],\ j = 1 \dots m, l = 1 \dots L. \tag{8}$$

Така удосконалена модель може бути основою для розробки правило-орієнтованого методу виявлення кіберінцидентів, пов'язаних з БД SIEM-системою із врахуванням ознак кіберінцидентів, отриманих із різних рівнів кіберзахисту БД ІКСВП.

Головним недоліком запропонованої моделі (8) є те, що впевненість експертів у кожному правилі «ЯКЩО – ТО», які входять до нечіткої бази знань (6), може бути різною. Цей недолік може бути усунений шляхом введення ваги правила, яка й буде характеризувати значимість того чи іншого правила під час ідентифікації кіберінцидентів. Тут, ґрунтуючись на роботах [16–20], під вагою правила будемо розуміти число в інтервалі [0, 1], яке характеризує суб'єктивну міру впевненості експерта в тому чи іншому правилі.

Ураховуючи це, багатовимірна таблиця ознак кіберінцидентів, що відбуваються у БД ІКСВП і класів, що їм відповідають (табл. 1), з урахування впевненості експерта у тому чи іншому правилі, яка задається за допомогою ваги правила, прийме наступний вигляд (табл. 2):

Таблиця 2

Багатовимірна таблиця ознак кіберінцидентів у БД ІКСВП і класів, що їм відповідають з урахуванням ваги правил

Номер вхідної комбінації значень	Ознаки кіберінцидентів, отримані з різних рівнів кіберзахисту БД						Вага правила	Клас кіберінциденту
	f_{1l}	f_{2l}	…	f_{il}	…	f_{nl}	ω	c
11	α_{1l}^{11}	α_{2l}^{11}	…	α_{il}^{11}	…	α_{nl}^{11}	ω_{11}	δ_1
12	α_{1l}^{12}	α_{2l}^{12}	…	α_{il}^{12}	…	α_{nl}^{12}	ω_{12}	
…	…	…	…	…	…	…		
$1k_1$	$\alpha_{1l}^{1k_1}$	$\alpha_{1l}^{1k_2}$	…	$\alpha_{il}^{1k_1}$	…	$\alpha_{nl}^{1k_1}$	ω_{1k_1}	
…	…	…	…	…	…	…		…

Номер вхідної комбінації значень	Ознаки кіберінцидентів, отримані з різних рівнів кіберзахисту БД						Вага правила	Клас кіберінциденту
	f_{1l}	f_{2l}	…	f_{il}	…	f_{nl}	ω	c
$j1$	α_{1l}^{j1}	α_{2l}^{j1}	…	α_{il}^{j1}	…	α_{nl}^{j1}	ω_{j1}	δ_j
$j2$	α_{1l}^{j2}	α_{2l}^{j2}	…	α_{il}^{j2}	…	α_{nl}^{j2}	ω_{j2}	
…	…	…	…	…	…	…	…	
jk_j	$\alpha_{1l}^{jk_j}$	$\alpha_{2l}^{jk_j}$	…	$\alpha_{il}^{jk_j}$	…	$\alpha_{nl}^{jk_j}$	ω_{jk_j}	
…	…	…	…	…	…	…	…	…
m_1	α_{1l}^{m1}	α_{2l}^{m1}	…	α_{il}^{m1}	…	α_{nl}^{m1}	ω_{m1}	δ_m
m_2	α_{1l}^{m2}	α_{2l}^{m2}	…	α_{il}^{m2}	…	α_{nl}^{m2}	ω_{m2}	
…	…	…	…	…	…	…	…	
mk_m	$\alpha_{1l}^{mk_m}$	$\alpha_{2l}^{mk_m}$	…	$\alpha_{il}^{mk_m}$	…	$\alpha_{nl}^{mk_m}$	ω_{mk_m}	

Тоді, ґрунтуючись на підході з [15; 16] із врахуванням ваг правил, нечітка база знань, яка представлена сукупністю зважених нечітких правил «ЯКЩО – ТО», що зв'язують лінгвістичні оцінки ознак кіберінцидентів у БД ІКСВП з результатами їхньої ідентифікації, прийме наступний вигляд:

ЯКЩО ($f_{il} = \alpha_{1l}^{11}$) ТА ($f_{2l} = \alpha_{2l}^{11}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{11}$) з вагою w_{11} АБО
($f_{il} = \alpha_{1l}^{12}$) ТА ($f_{2l} = \alpha_{2l}^{12}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{12}$) з вагою w_{12} АБО
($f_{il} = \alpha_{1l}^{1k_1}$) ТА ($f_{2l} = \alpha_{2l}^{1k_1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{1k_1}$) з вагою w_{1k_1},
ТО (с=δ_1),…
…, ЯКЩО ($f_{il} = \alpha_{1l}^{j1}$) ТА ($f_{2l} = \alpha_{2l}^{j1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{j1}$) з вагою w_{j1} АБО
($f_{il} = \alpha_{1l}^{j2}$) ТА ($f_{2l} = \alpha_{2l}^{j2}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{j2}$) з вагою АБО
($f_{il} = \alpha_{1l}^{jk_j}$) ТА ($f_{2l} = \alpha_{2l}^{jk_j}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{jk_j}$) з вагою w_{jk_j},
ТО (с=δ_j),… (9)
…, ЯКЩО($f_{il} = \alpha_{1l}^{m1}$) ТА ($f_{2l} = \alpha_{2l}^{m1}$) ТА … ТА ($f_{nl} = \alpha_{nl}^{m1}$) з вагою w_{m1} АБО
($f_{il} = \alpha_{1l}^{m2}$) ТА ($f_{2l} = \alpha_{2l}^{m2}$) ТА … ТА ($f_{nl} = \alpha_{nl}^{m2}$) з вагою w_{m2} АБО
($f_{il} = \alpha_{1l}^{mk_m}$) ТА ($f_{2l} = \alpha_{2l}^{mk_m}$) ТА … ТА ($f_{nl} = \alpha_{nl}^{mk_m}$) з вагою w_{mk_m},
ТО (с=δ_m),

де α_{il}^{jk} – лінгвістична оцінка ознаки кіберінциденту у БД ІКСВП f_{il}, $i = 1 \dots n$; l=1…L у рядку k j-ої диз'юнкції, що визначається на терм-множині $A_i = \{\alpha_{il}^1, \alpha_{il}^2, \dots, \alpha_{il}^{k_i}\}$;

$\delta_j \in \Delta$, $j = 1 \dots m$ – лінгвістична оцінка класу кіберінциденту у БД ІКСВП, що визначається на терм-множині $\Delta = \{\delta_1, \delta_2, \dots, \delta_m\}$, w_{jk} – вага правила.

З урахуванням ваг правил, нечітка база знань (3.13) може бути представлена модифікованою системою нечітких рівнянь (3.11) наступним чином:

$$\mu^{\delta_j}(f_{1l}, f_{2l}, \dots, f_{nl}) = \vee_{k=1}^{k_j} \left\{ \omega_{jk} \left[\wedge_{i=1}^{n} \mu^{\alpha_{il}^{jk}}(f_{il}) \right] \right\}, j = 1 \dots m. \quad (10)$$

Шляхом заміни операцій $\wedge$ та $\vee$ на операції min та max [14–19], які їм відповідають, отримаємо модифіковану модель нечіткої ідентифікації кіберінцидентів у БД ІКСВП SIEM-системами (15) зі зваженими правилами:

$$\mu^{\delta_j}(f_{il}, f_{2l}, \dots, f_{nl}) = \max_{k = \overline{1, k_j}} \left\{ \omega_{jk} \min_{i = \overline{1, n}} \left\{ \mu^{\alpha_{il}^{jk}}(f_{il}) \right\}, j = 1 \dots m \right\}. \quad (11)$$

Така удосконалена модель дозволяє усунути недоліки, що були присутні в моделі (8), шляхом урахування впевненості експерта у тому чи іншому правилі.

Проте слід зауважити, що вплив окремих ознак кіберінцидентів (антецедентів) правила на результат ідентифікації, з точки зору впевненості експерта, ця модель не враховує. Крім того, сучасні бази знань SIEM-систем можуть містити десятки тисяч правил, тому на практиці експерту важко визначити вагу того чи іншого правила. Виходячи з цього, цілком доцільно удосконалити модель (11) шляхом надання ознакам кіберінцидентів вагових коефіцієнтів в межах правил, що визначаються експертами (табл. 3).

Таблиця 3

Багатовимірна таблиця ознак кіберінцидентів у БД ІКСВП і класів, що їм відповідають, з урахуванням ваги ознак кіберінцидентів у межах окремих правил

Номер вхідної комбінації значень	Ознаки кіберінцидентів, отримані з різних рівнів кіберзахисту БД										Клас кібер-інциденту
	f_{1l}	Ω_{1l}	f_{2l}	Ω_{2l}	…	f_{il}	Ω_{il}	…	f_{nl}	Ω_{nl}	c
11	α_{1l}^{11}	$\Omega_{1l}{}^{11}$	α_{2l}^{11}	$\Omega_{2l}{}^{11}$	…	α_{il}^{11}	$\Omega_{il}{}^{11}$	…	α_{nl}^{11}	$\Omega_{nl}{}^{11}$	δ_1
12	α_{1l}^{12}	$\Omega_{1l}{}^{12}$	α_{2l}^{12}	$\Omega_{2l}{}^{12}$	…	α_{il}^{12}	$\Omega_{il}{}^{12}$	…	α_{nl}^{12}	$\Omega_{nl}{}^{12}$	
…	…		…		…	…		…	…		
1k₁	$\alpha_{1l}^{1k_1}$	$\Omega_{1l}{}^{1k}{}_1$	$\alpha_{1l}^{1k_2}$	$\Omega_{2l}{}^{1k}{}_2$	…	$\alpha_{il}^{1k_1}$	$\Omega_{il}{}^{1k}{}_1$	…	$\alpha_{nl}^{1k_1}$	$\Omega_{nl}{}^{1k}{}_1$	
…	…		…		…	…		…	…		…
j1	α_{1l}^{j1}	$\Omega_{1l}{}^{j1}$	α_{2l}^{j1}	$\Omega_{2l}{}^{j1}$	…	α_{il}^{j1}	$\Omega_{il}{}^{j1}$	…	α_{nl}^{j1}	$\Omega_{nl}{}^{j1}$	δ_j
j2	α_{1l}^{j2}	$\Omega_{1l}{}^{j2}$	α_{2l}^{j2}	$\Omega_{2l}{}^{j2}$	…	α_{il}^{j2}	$\Omega_{il}{}^{j2}$	…	α_{nl}^{j2}	$\Omega_{nl}{}^{j2}$	
…	…		…		…	…		…	…		
jk$_j$	$\alpha_{1l}^{jk_j}$	$\Omega_{1l}{}^{jk}{}_j$	$\alpha_{2l}^{jk_j}$	$\Omega_{2l}{}^{jk}{}_j$	…	$\alpha_{il}^{jk_j}$	$\Omega_{il}{}^{jk}{}_j$	…	$\alpha_{nl}^{jk_j}$	$\Omega_{nl}{}^{jk}{}_j$	
…	…		…		…	…		…	…		…
m₁	α_{1l}^{m1}	$\Omega_{1l}{}^{m1}$	α_{2l}^{m1}	$\Omega_{2l}{}^{m1}$	…	α_{il}^{m1}	$\Omega_{il}{}^{m1}$	…	α_{nl}^{m1}	$\Omega_{nl}{}^{m1}$	δ_m
m₂	α_{1l}^{m2}	$\Omega_{1l}{}^{m2}$	α_{2l}^{m2}	$\Omega_{2l}{}^{m2}$	…	α_{il}^{m2}	$\Omega_{il}{}^{m2}$	…	α_{nl}^{m2}	$\Omega_{nl}{}^{m2}$	
…	…		…		…		…	…			
mk$_m$	$\alpha_{1l}^{mk_m}$	$\Omega_{1l}{}^{mk}{}_m$	$\alpha_{2l}^{mk_m}$	$\Omega_{2l}{}^{mk}{}_m$	…	$\alpha_{il}^{mk_m}$	$\Omega_{il}{}^{mk}{}_m$	…	$\alpha_{nl}^{mk_m}$	$\Omega_{nl}{}^{mk}{}_m$	

Тоді, з врахуванням ваг ознак кіберінцидентів у БД ІКСВП, НБЗ, яка представлена сукупністю нечітких правил «ЯКЩО – ТО», що зв'язують лінгвістичні оцінки ознак кіберінцидентів з результатами їхньої ідентифікації, прийме наступний вигляд:

ЯКЩО ($f_{il} = \alpha_{1l}^{11}$з вагою $\Omega_{1l}{}^{11}$) ТА ($f_{2l} = \alpha_{2l}^{11}$ з вагою $\Omega_{2l}{}^{11}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{11}$з вагою $\Omega_{nl}{}^{11}$) АБО

($f_{il} = \alpha_{1l}^{12}$з вагою $\Omega_{1l}{}^{12}$) ТА ($f_{2l} = \alpha_{2l}^{12}$з вагою $\Omega_{2l}{}^{12}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{12}$з вагою $\Omega_{nl}{}^{12}$) АБО

($f_{il} = \alpha_{1l}^{1k_1}$з вагою $\Omega_{1l}{}^{1k}{}_1$) ТА ($f_{2l} = \alpha_{2l}^{1k_1}$з вагою $\Omega_{2l}{}^{1k}{}_1$) ТА…ТА ($f_{nl} = \alpha_{nl}^{1k_1}$з вагою $\Omega_{nl}{}^{1k}{}_1$), ТО (c=δ_1),…

…, ЯКЩО ($f_{il} = \alpha_{1l}^{j1}$з вагою $\Omega_{1l}{}^{j1}$) ТА ($f_{2l} = \alpha_{2l}^{j1}$з вагою $\Omega_{2l}{}^{j1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{j1}$з вагою $\Omega_{nl}{}^{j1}$) АБО

($f_{il} = \alpha_{1l}^{j2}$з вагою $\Omega_{1l}{}^{j2}$) ТА ($f_{2l} = \alpha_{2l}^{j2}$з вагою $\Omega_{2l}{}^{j2}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{j2}$з вагою $\Omega_{nl}{}^{j2}$) АБО

($f_{il} = \alpha_{1l}^{jk_j}$з вагою $\Omega_{1l}{}^{jk}{}_j$) ТА ($f_{2l} = \alpha_{2l}^{jk_j}$з вагою $\Omega_{2l}{}^{jk}{}_j$) ТА…ТА ($f_{nl} = \alpha_{nl}^{jk_j}$з вагою $\Omega_{nl}{}^{jk}{}_j$), ТО (c=δ_j),…

…, ЯКЩО ($f_{il} = \alpha_{1l}^{m1}$ з вагою $\Omega_{1l}{}^{m1}$) ТА ($f_{2l} = \alpha_{2l}^{m1}$з вагою $\Omega_{2l}{}^{m1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{m1}$з вагою $\Omega_{nl}{}^{m1}$) АБО

($f_{il} = \alpha_{1l}^{m2}$з вагою Ω_{il}^{m2})ТА($f_{2l} = \alpha_{2l}^{m2}$з вагою Ω_{2l}^{m2})ТА…ТА($f_{nl} = \alpha_{nl}^{m2}$з вагою Ω_{nl}^{m2})

АБО

($f_{il} = \alpha_{1l}^{mk_m}$з вагою $\Omega_{il}^{mk_m}$)ТА($f_{2l} = \alpha_{2l}^{mk_m}$з вагою $\Omega_{2l}^{mk_m}$)ТА…ТА($f_{nl} = \alpha_{nl}^{mk_m}$з вагою $\Omega_{nl}^{mk_m}$),

ТО (c=δ_m), (12)

де α_{il}^{jk} – лінгвістична оцінка ознаки кіберінциденту у БД ІКСВП $f_{il},\ i = 1 \dots n$; l=1…L у рядку k j-ої диз'юнкції, що визначається на терм-множині $A_i = \{\alpha_{il}^1, \alpha_{il}^2, \dots, \alpha_{il}^{k_i}\}$;

$\delta_j \in \Delta,\ j = 1 \dots m$ – лінгвістична оцінка класу кіберінциденту у БД ІКСВП, що визначається на терм-множині $\Delta = \{\delta_1, \delta_2, \dots, \delta_m\}$;

Ω_{nl}^{mk} – вага ознак кіберінцидентів у межах правил.

З урахуванням ваг ознак кіберінцидентів у межах правил, нечітка база знань (11) може бути представлена модифікованою системою нечітких рівнянь (12) наступним чином:

$$\mu^{\delta_j}(f_{1l}, f_{2l}, \dots, f_{nl}) = \bigvee_{k=1}^{k_j} \left\{ \bigwedge_{i=1}^{n} \left[\mu^{\alpha_{il}^{jk}}(f_{il})\ \Omega_{il}^{jk} \right] \right\}, j = 1 \dots m. \quad (13)$$

Шляхом заміни операцій ∧ та ∨на операції min та max, які їм відповідають, отримаємо модифіковану модель нечіткої ідентифікації кіберінцидентів SIEM-системами (13) зі зваженими ознаками правил:

$$\mu^{\delta_j}(f_{il}, f_{2l}, \dots, f_{nl}) = \underset{k = \overline{1, k_j}}{max} \left\{ \underset{i = \overline{1, n}}{min} \left\{ \mu^{\alpha_{il}^{jk}}(f_{il})\ \Omega_{il}^{jk} \right\},\ j = \overline{1, m} \right\}. \quad (14)$$

Така модель дозволяє певною мірою усунути недоліки, що були присутні в моделі (11), шляхом урахування впевненості експерта у тій чи іншій ознаці кіберінциденту у правилі.

Як приклад, розглянемо кіберінцидент, який має бути розглянутий офіцером безпеки, щодо кількості невдалих спроб підключення до БД з однієї або декількох IP-адрес.

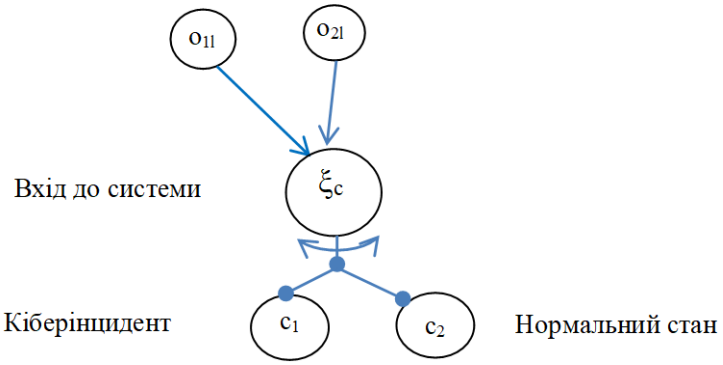

Рис. 2. Дерево логічного виводу для кіберінциденту

На рисунку 2 наведено дерево логічного виводу для цього кіберінциденту.

Для ідентифікації зазначеного кіберінциденту доцільно дослідити його ознаки у вигляді вхідних лінгвістичних змінних:

numb_attempts (o_{1l}) – кількість невдалих спроб вводу пароля для входу до БД, який відноситься безпосередньо до рівня кіберзахисту БД [15];

count_ip (o_{2l}) – кількість IP-адрес, з яких вводився пароль, що відповідає рівню кіберзахисту мережі.

Для опису подій, які відбуваються в системі з БД, використаємо наступні їхні типи та позначимо через c_1 та c_2.

Таблиця 4

Тип події, що відбувається в системі

Подія	Зміст події
C_1	Кіберінцидент
C_2	Нормальний стан системи

Структура дерева логічного виводу відповідає відношенню (14):

$$c = \xi_c(o_{1l}, o_{2l}). \quad (14)$$

Для оцінки значень лінгвістичних змінних o_{1l} та o_{2l} застосовується єдина шкала якісних термів: Н – низька, В – висока. Кожний з цих термів задається відповідною функцією належності.

Таблиця 5

Нечіткий опис станів

Ознаки кіберінциденту		Тип кіберінциденту
numb_attempts (o_{1l})	*count_ip* (o_{2l})	
Н	Н	C_2
Н	В	
В	Н	C_1
В	В	

Дослідження ознак проведемо на основі функцій належності з наступними значеннями:
numb_attempts (o_{1l}) – {«Н-низька [1,2,3]», «В-висока [3,4,5]»} на універсумі [1,5];
count_ip (o_{2l}) – {«Н-низька [1,2]», »В-висока [2,3,4,5]»} на універсумі [1,5].

Значення функції належності для деякого фіксованого вектора вхідних даних, що відповідають лінгвістичній змінній – *numb_attempts* (o_{1l}), наведено на рисунку 3.

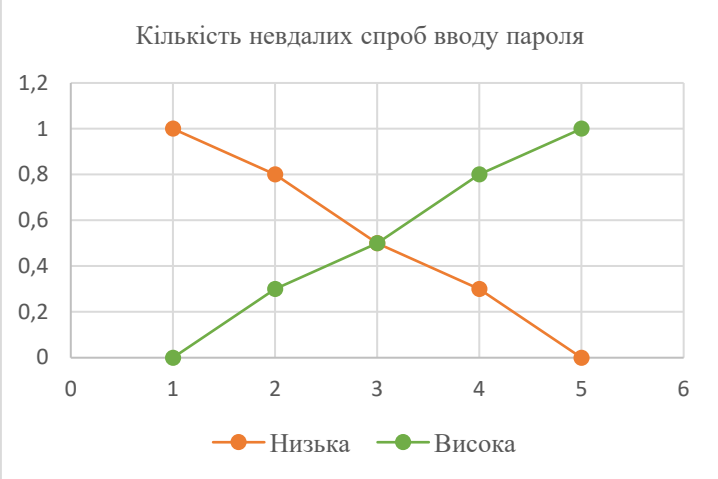

Рис. 3. Значення функції належності лінгвістичної змінної «Кількість невдалих спроб вводу пароля»

Значення функції належності для деякого фіксованого вектора вхідних даних, що відповідають лінгвістичній змінній – *count_ip* (o_{2l}), наведено на рисунку 4.

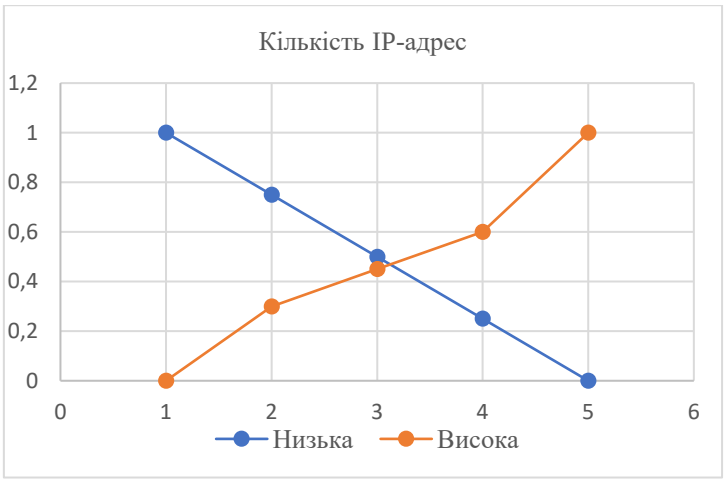

Рис. 4. Значення функції належності лінгвістичної змінної «Кількість IP-адрес, з яких вводився пароль»

З метою порівняння раніше розглянутих моделей (8), (11) та (14), розглянемо деякі варіанти спрацювання нечітких правил. Функції належності, що відповідають цим ознакам, наведено в таблиці 6 та в таблиці 7.

Таблиця 6

Результати обчислень функцій належності вхідного вектора

Лінгвістична змінна	o_{il}	$\mu^H(o_{il})$	$\mu^B(o_{il})$
numb_attempts (o_{1l})	1	1	0
	2	0,8	0,3
	3	0,5	0,5
	4	0,3	0,8
	5	0	1

Таблиця 7

Результати обчислень функцій належності вхідного вектора

Лінгвістична змінна	o_{il}	$\mu^H(o_{il})$	$\mu^B(o_{il})$
count_ip (o_{2l})	1	1	0
	2	0,75	0,3
	3	0,5	0,45
	4	0,25	0,6
	5	0	1

Для визначення максимальної відповідності досліджуваних ознак до шаблону подій, які відбуваються в системі, обчислюємо функції належності на основі (8), а логічні операції кон'юнкції та диз'юнкції заміняємо на нечіткі кон'юнкцію та диз'юнкцію на основі максимінного підходу:

$$\mu(a) \wedge \mu(b) = \min[\mu(a), \mu(b)],$$
$$\mu(a) \vee \mu(b) = \max[\mu(a), \mu(b)].$$

Результати розрахунків для моделі (8) будуть наступними:
Кількість спроб підключення до БД – 2 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}8 \wedge 0{,}75) \vee (0{,}8 \wedge 0{,}3) = 0{,}75,$$

$$\mu^{c_1} = (0{,}3 \wedge 0{,}75) \vee (0{,}3 \wedge 0{,}3) = 0{,}3.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}5 \wedge 0{,}75) \vee (0{,}5 \wedge 0{,}3) = 0{,}5,$$
$$\mu^{c_1} = (0{,}5 \wedge 0{,}75) \vee (0{,}5 \wedge 0{,}3) = 0{,}5.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}3 \wedge 0{,}75) \vee (0{,}3 \wedge 0{,}3) = 0{,}3,$$
$$\mu^{c_1} = (0{,}8 \wedge 0{,}75) \vee (0{,}8 \wedge 0{,}3) = 0{,}75.$$

Кількість спроб підключення до БД – 5 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0 \wedge 0{,}75) \vee (0 \wedge 0{,}3) = 0,$$
$$\mu^{c_1} = (1 \wedge 0{,}75) \vee (1 \wedge 0{,}3) = 0{,}75.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}5 \wedge 0{,}5) \vee (0{,}5 \wedge 0{,}45) = 0{,}5,$$
$$\mu^{c_1} = (0{,}5 \wedge 0{,}5) \vee (0{,}5 \wedge 0{,}45) = 0{,}5.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}3 \wedge 0{,}5) \vee (0{,}3 \wedge 0{,}45) = 0{,}3,$$
$$\mu^{c_1} = (0{,}8 \wedge 0{,}5) \vee (0{,}8 \wedge 0{,}45) = 0{,}5.$$

Кількість спроб підключення до БД – 5 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0 \wedge 0{,}5) \vee (0 \wedge 0{,}45) = 0,$$
$$\mu^{c_1} = (1 \wedge 0{,}5) \vee (1 \wedge 0{,}45) = 0{,}5.$$

Результати розрахунків, опираючись на модель (12), допустимо випадок, коли значення ваги правил співпадає і має – 0,8:

Кількість спроб підключення до БД – 2 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}8 \wedge 0{,}75) * 0,8 \vee (0{,}8 \wedge 0{,}3) * 0,8 = 0{,}75,$$
$$\mu^{c_1} = (0{,}3 \wedge 0{,}75) * 0,8 \vee (0{,}3 \wedge 0{,}3) * 0,8 = 0{,}3.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}5 \wedge 0{,}75) * 0,8 \vee (0{,}5 \wedge 0{,}3) * 0,8 = 0{,}4,$$
$$\mu^{c_1} = (0{,}5 \wedge 0{,}75) * 0,8 \vee (0{,}5 \wedge 0{,}3) * 0,8 = 0{,}4.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}3 \wedge 0{,}75) * 0,8 \vee (0{,}3 \wedge 0{,}3) * 0,8 = 0{,}24,$$
$$\mu^{c_1} = (0{,}8 \wedge 0{,}75) * 0,8 \vee (0{,}8 \wedge 0{,}3) * 0,8 = 0{,}6.$$

Кількість спроб підключення до БД – 5 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0 \wedge 0{,}75) * 0,8 \vee (0 \wedge 0{,}3) * 0,8 = 0,$$
$$\mu^{c_1} = (1 \wedge 0{,}75) * 0,8 \vee (1 \wedge 0{,}3) * 0,8 = 0{,}6.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}5 \wedge 0{,}5) * 0,8 \vee (0{,}5 \wedge 0{,}45) * 0,8 = 0{,}4,$$
$$\mu^{c_1} = (0{,}5 \wedge 0{,}5) * 0,8 \vee (0{,}5 \wedge 0{,}45) * 0,8 = 0{,}4.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}3 \wedge 0{,}5) * 0,8 \vee (0{,}3 \wedge 0{,}45) * 0,8 = 0{,}24,$$
$$\mu^{c_1} = (0{,}8 \wedge 0{,}5) * 0,8 \vee (0{,}8 \wedge 0{,}45) * 0,8 = 0{,}4.$$

Кількість спроб підключення до БД – 5 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0 \wedge 0{,}5) * 0,8 \vee (0 \wedge 0{,}45) * 0,8 = 0,$$
$$\mu^{c_1} = (1 \wedge 0{,}5) * 0,8 \vee (1 \wedge 0{,}45) * 0,8 = 0{,}4.$$

Припустимо, що були введені наступні значення ваги антецедентів:

кількість невдалих спроб вводу пароля = низька – 0,8;

кількість невдалих спроб вводу пароля = висока – 1;

кількість IP-адрес, з яких вводився пароль = низька – 0,8;

кількість IP-адрес, з яких вводився пароль = висока – 0,9.

Отже, результати розрахунків на основі моделі (14):

Кількість спроб підключення до БД – 2 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}8 * 0{,}8 \wedge 0{,}75 * 0{,}8) \vee (0{,}8 * 0{,}8 \wedge 0{,}3 * 0{,}9) = 0{,}6,$$

$$\mu^{c_1} = (0{,}3 * 1 \wedge 0{,}75 * 0{,}8) \vee (0{,}3 * 1 \wedge 0{,}3 * 0{,}9) = 0{,}27.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}5 * 0{,}8 \wedge 0{,}75 * 0{,}8) \vee (0{,}5 * 0{,}8 \wedge 0{,}3 * 0{,}9) = 0{,}4,$$

$$\mu^{c_1} = (0{,}5 * 1 \wedge 0{,}75 * 0{,}8) \vee (0{,}5 * 1 \wedge 0{,}3 * 0{,}9) = 0{,}5.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}3 * 0{,}8 \wedge 0{,}75 * 0{,}8) \vee (0{,}3 * 0{,}8 \wedge 0{,}3 * 0{,}9) = 0{,}24,$$

$$\mu^{c_1} = (0{,}8 * 1 \wedge 0{,}75 * 0{,}8) \vee (0{,}8 * 1 \wedge 0{,}3 * 0{,}9) = 0{,}6.$$

Кількість спроб підключення до БД – 5 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0 * 0{,}8 \wedge 0{,}75 * 0{,}8) \vee (0 * 0{,}8 \wedge 0{,}3 * 0{,}9) = 0,$$

$$\mu^{c_1} = (1 * 1 \wedge 0{,}75 * 0{,}8) \vee (1 * 1 \wedge 0{,}3 * 0{,}9) = 0{,}6.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}5 * 0{,}8 \wedge 0{,}5 * 0{,}8) \vee (0{,}5 * 0{,}8 \wedge 0{,}45 * 0{,}9) = 0{,}4,$$

$$\mu^{c_1} = (0{,}5 * 1 \wedge 0{,}5 * 0{,}8) \vee (0{,}5 * 1 \wedge 0{,}45 * 0{,}9) = 0{,}405.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}3 * 0{,}8 \wedge 0{,}5 * 0{,}8) \vee (0{,}3 * 0{,}8 \wedge 0{,}45 * 0{,}9) = 0{,}24,$$

$$\mu^{c_1} = (0{,}8 * 1 \wedge 0{,}5 * 0{,}8) \vee (0{,}8 * 1 \wedge 0{,}45 * 0{,}9) = 0{,}405.$$

Кількість спроб підключення до БД - 5 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0 * 0{,}8 \wedge 0{,}5 * 0{,}8) \vee (0 * 0{,}8 \wedge 0{,}45 * 0{,}9) = 0,$$

$$\mu^{c_1} = (1 * 1 \wedge 0{,}5) \vee (1 * 1 \wedge 0{,}45 * 0{,}9) = 0{,}5.$$

Внаслідок проведених розрахунків (табл. 8) та аналізу виявлено, що у випадку застосування моделей (8) та (11) у процесі прийняття рішень виникають ситуації, коли кілька правил активуються одночасно (в поточних розрахунках це відображається, коли кількість невдалих спроб входу відповідає трьом (o_{1l})). Це призводить до невизначеності у прийнятті рішень офіцерами безпеки щодо ідентифікації кіберінциденту, пов'язаного з БД.

Для усунення цієї проблеми запропоновано використовувати ваги антецедентів (логічних умов у правилі), що дозволяє враховувати важливість кожного антецедента при прийнятті рішення. Застосування ваг антецедентів дозволить уникнути ситуацій, коли правила з однаковою вагою конфліктують між собою, та надати особам, що приймають рішення (ОПР), рекомендації для прийняття обґрунтованих рішень у критичних ситуаціях.

Таблиця 8

Порівняння результатів розрахунків

Значення o_{2l}	Значення o_{1l}	Модель (8)		Модель (11)		Модель (14)	
		μ^{c_2}	μ^{c_1}	μ^{c_2}	μ^{c_1}	μ^{c_2}	μ^{c_1}
2	2	0,75 +	0,3	0,5 +	0,3	0,6 +	0,27
	3	0,5 ?	0,5 ?	0,4 ?	0,4 ?	0,4	0,5 +
	4	0,3	0,75 +	0,24	0,6 +	0,24	0,6 +
	5	0	0,75 +	0	0,6 +	0	0,6 +
3	3	0,5 ?	0,5 ?	0,4 ?	0,4 ?	0,4	0,405 +
	4	0,3	0,5 +	0,24	0,4 +	0,24	0,405 +
	5	0	0,5 +	0	0,6+	0	0,5 +

Використання ваг антецедентів у моделі (14) ідентифікації кіберінцидентів SIEM-системою (рис. 5) призвело до покращення порівняно з моделлю (8) та моделлю з вагами правил (12).

Так, з рисунку 5 видно, що під час трьох невдалих спроб введення пароля до БД, у випадку використання моделі (14), усувається невизначеність в результатах ідентифікації кіберінциденту/нормального стану системи при заданих значеннях функцій належності нечітких термів (див. рис. 3). Зокрема, вочевидь буде прийнято рішення щодо віднесення цієї події до нормального стану функціонування інформаційної системи, до складу якої входить БД, що у багатьох випадках відповідає дійсності. Проте це не можна сказати у випадку застосування моделей (8) та (11).

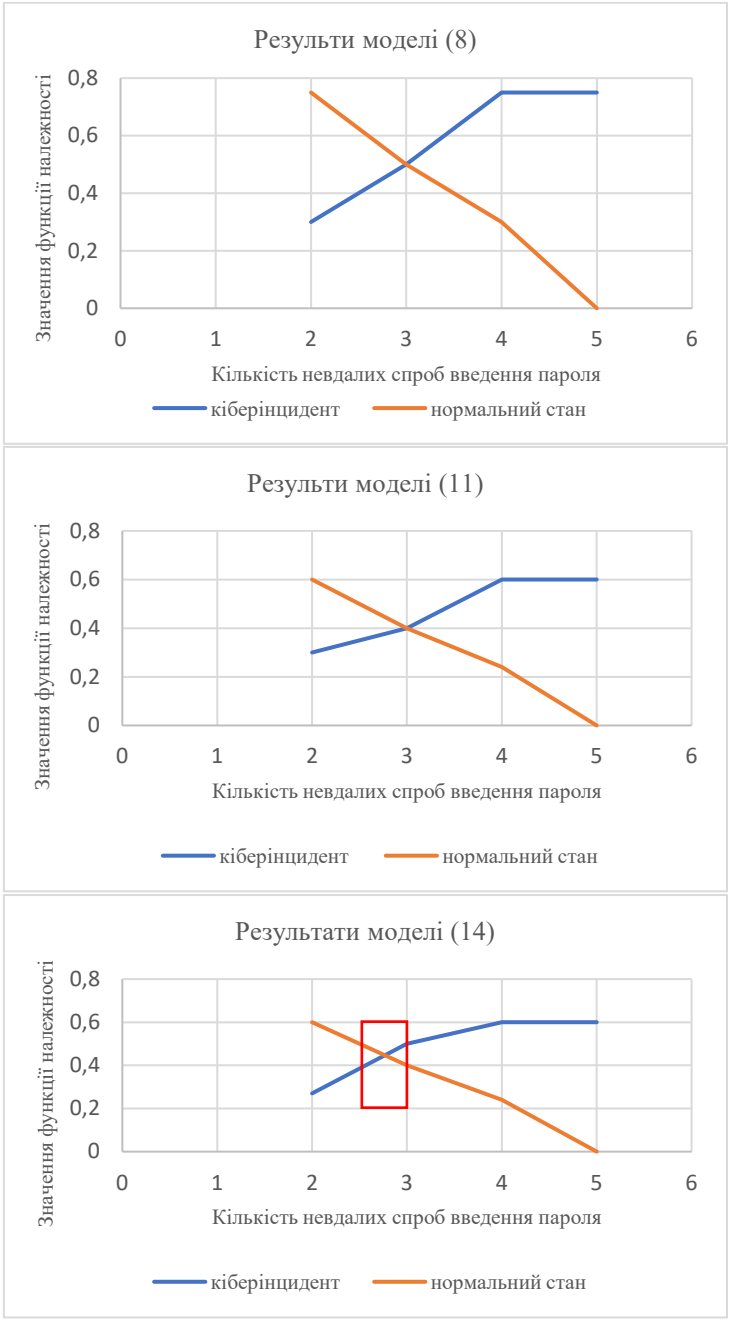

Рис. 5. Результати ідентифікації кіберінцидентів/нормального стану роботи системи у випадках застосування моделей (8), (11) та (14)

Це свідчить про доцільність у деяких випадках врахування ваг антецедентів в усуненні проблеми неоднозначності, що виникає при спрацюванні протилежних правил. Отже, результати підтверджують, що модель з використанням ваг антецедентів, при певних умовах,

може бути більш адекватною порівняно з іншими розглянутими моделями виявлення (ідентифікації) кіберінцидентів у БД ІКСВП. Це може мати важливі наслідки для розробки ефективних систем кіберзахисту та забезпечення кіберзахисту БД в ІКСВП.

Висновки. Удосконалення та впровадження нечітких моделей виявлення кіберінцидентів у БД суттєво покращує ефективність застосування систем управління інформацією та подіями безпеки в контурі захисту ІКСВП в умовах неповноти та невизначеності інформації про них. Аналіз існуючих моделей показав їхні недоліки, пов'язані зі складністю вирішення задачі ідентифікації кіберінцидентів у деяких випадках, зокрема у випадку невдало налаштованих функцій належності нечітких термів, що потребує додатково уточнення їх по експериментальним даним. Проте в умовах ландшафту кібератак, що постійно змінюється, це не завжди є можливим. Крім цього успішне застосування розглянутих моделей є можливим на основі удосконалення їх шляхом розширення простору ознак кіберінцидентів, отриманих з різних рівнів кіберзахисту БД.

Ще одним із недоліків, пов'язаним із застосування нечіткої моделі ідентифікації кіберінцидентів зі зваженими правилами, є великий обсяг бази знань, що призводить до складності рішення цієї задачі, а також значних часових та обчислювальних витрат.

Виходом з цієї ситуації, у деяких випадках, може стати застосування нечіткої моделі ідентифікації кіберінцидентів зі зваженими антецедентами (ознаками кіберінцидентів) правил, яка запропонована в роботі.

Перспективним напрямком подальших досліджень є розробка методу визначення ваги антецедентів нечітких правил.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Герасимов Б. М., Субач І. Ю., Хусаінов П. В., Міщенко В. О. Аналіз задач моніторингу інформаційних мереж та методів підвищення ефективності їх функціонування. *Сучасні інформаційні технології у сфері безпеки та оборони.* 2008. № 3 (3). С. 24–27.

2. Гнатюк С. О. Система корелювання подій та управління інцидентами кібербезпеки на об'єктах критичної інфраструктури. *Кібербезпека: освіта, наука, техніка*. 2023. Т. 3. № 19. С. 176–196. URL: https://doi.org/10.28925/2663-4023.2023.19.176196.

3. Субач І., Кубрак В., Микитюк А. Архітектура та функціональна модель перспективної проактивної інтелектуальної системи SIEM-системи для кіберзахисту об'єктів критичної інфраструктури. *Information Technology and Security*. 2019. № 7 (2). P. 208–215. URL: https://doi.org/10.20535/2411-1031.2019.7.2.190570.

4. Субач І., Власенко О. Архітектура інтелектуальної SIEM-системи для виявлення кіберінцидентів у базах даних інформаційно-телекомунікаційних системах військового призначення. *Збірник наукових праць ВІТІ.* 2023. № 4. С. 82–92. URL: https://doi.org/10.58254/viti.4.2023.07.82.

5. Seyed M. Z. Analysis of Security Information and Event Management (SIEM) – Evasion and Detection Methods. Tallinn University of Technology, Faculty of Information Technology, Tallinn, Estonia, Master Thesis, 2016.

6. Granadillo, Gustavo Gonzalez. Security Information and Event Management (SIEM): Analysis, Trends, and Usage in Critical Infrastructures. *Sensors (Basel, Switzerland)* 21. 2021: n. pag. DOI: 10.3390/s21144759.

7. Suarez-Tangil Guillermo, Palomar Esther, Ribagorda Arturo, Sanz Ivan. *Providing SIEM systems with self-adaptation*. 2015.

8. Anastasov Igor, Davcev Danco. SIEM implementation for global and distributed environments. *Computer Applications and Information Systems* (WCCAIS). 2014 World Congress, 2014.

9. Rafał Leszczyna, Michał R. Wróbel. Evaluation of Open Source SIEM for Situation Awareness Platform in the Smart Grid Environment. *Factory Communication Systems* (WFCS), IEEE World Conference on, 2015.

10. Hanemann A., Marcu P. Algorithm Design and Application of Service Oriented Event Correlation, In Proceedings of Conference BDIM 2008, 3rd IEEE/IFIP International Workshop on Business-Driven IT Management. 2011. P. 61–70.

11. Elshoush H., Osman I. M. Alert correlation in collaborative intelligent intrusion detection systems. A survey. Applied Soft Computing, 2011. P. 4349–4365.

12. Субач І., Фесьоха В. Модель виявлення аномалій в інформаційно-телекомунікаційних мережах органів військового управління на основі нечітких множин та нечіткого логічного виводу. *Збірник наукових праць ВІТІ*. 2017. № 3. С. 158–164. URL: http://nbuv.gov.ua/UJRN/Znpviti_2017_3_21.

13. Subach I., Fesokha V. Model of detecting cybernetic attacks on information-telecommunication systems based on description of anomalies in their work by weighed fuzzy rules. *Collection «Information Technology and Security»*, 2017. № 5 (2). P. 145–152. URL: https://doi.org/10.20535/2411-1031.2017.5.2.136984.

14. Субач І., Власенко О. Інформаційні технології захисту баз даних від кібератак в інформаційних системах військового призначення. *Collection «Information Technology and Security»*. 2022. № 10 (2). С. 177–193. URL: https://doi.org/10.20535/2411-1031.2022.10.2.270412.

15. Rotshtein A. P. Medical diagnostics using fuzzy logic. Vinnitsa: Continent-PRIM, 1996. 132 p.

16. Rothstein A. Intelligent identification technologies: fuzzy sets, genetic algorithms, neural networks. Vinnytsia: UNIVERSUM, 1999.

17. RothsteinO., Chernovolyk G., Laryushkin E. Method of constructing membership functions of fuzzy sets. Bulletin of VPI, 1996. Vol. 3. P. 72–75.

18. Mityushkin Y., Mokin B., Rothstein O. Soft Computing: identification of patterns of fuzzy knowledge bases: a monograph. Vinnytsia: UNIVERSUM-Vinnytsia, 2002.

19. Rotshtein A. Design and Tuning of Fuzzy Rule-Based Systems for Medical Diagnosis. In N.-H. Teodorescu (ed): Fuzzy and Neuro - Fuzzy Systems in Medicine. CRC Press. 1998. P. 243–289.

20. Zaichenko Y. P. Operations Research: Fuzzy Optimisation. Vyshcha Shkola, 1991.

21. Borisov A. N., Krumberg O. A., Fedorov I. P. Decision-Making on the Basis of Fuzzy Models: Examples of Use. Zinatne. 1990.