КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№4 (8), 2020
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2020.8.135148
УДК 004.056

**Лахно Валерій Анатолійович**
д.т.н., професор, зав. кафедрою комп'ютерних систем і мереж
Національний університет біоресурсів і природокористування України, Київ, Україна
ORCID: 0000-0001-9695-4543
*valss21@ukr.net*

**Касаткін Дмитро Юрійович**
к.пед.н., доцент, доцент кафедри комп'ютерних систем і мереж
Національний університет біоресурсів і природокористування України, Київ, Україна
ORCID: 0000-0002-2642-8908
*d.kasatkin@nubip.edu.ua*

**Блозва Андрій Ігорович**
к.пед.н., доцент, доцент кафедри комп'ютерних систем і мереж
Національний університет біоресурсів і природокористування України, Київ, Україна
ORCID: 0000-0002-4377-0916
*andriy.blozva@nubip.edu.ua*

**Місюра Максим Дмитрович**
к.т.н., доцент кафедри комп'ютерних систем і мереж
Національний університет біоресурсів і природокористування України, Київ, Україна
ORCID: 0000-0002-9061-3462
*mdm@nubip.edu.ua*

**Гусєв Борис Семенович**
к.т.н., доцент, доцент кафедри комп'ютерних систем і мереж
Національний університет біоресурсів і природокористування України, Київ, Україна
ORCID: 0000-0003-1658-7822
*gusevbs@nubip.edu.ua*

# ПРОЕКТУВАННЯ БАЗИ ЗНАНЬ ДЛЯ СИСТЕМ КІБЕРБЕЗПЕКИ НА ОСНОВІ МЕТОДУ ЗМІСТОВНОЇ ІДЕНТИФІКАЦІЇ ОБ'ЄКТІВ

**Анотація.** У статті викладені результати досліджень, виконаних в процесі проектування експертної системи (ЕС), призначеної для оцінки загроз інформаційної безпеки (ІБ) критично важливих об'єктів інформатизації (КВОІ). Запропоновано підхід до проектування експертної системи на основі силогізмів і логіки предикатів, а також методу змістовної ідентифікації об'єктів бази знань (БЗ). Суть методу полягає в тому, що кожному об'єкту БЗ проектованої ЕС, ставиться у відповідність кортеж ключових слів (КС), значимість яких визначається експертним шляхом. Таким чином, кожен об'єкт БЗ ставиться у відповідність елементу кінцевого нечіткого топологічного простору об'єктів БЗ. Змістовна ідентифікація проходить по відстані між об'єктами БЗ. Запропонований в роботі підхід, в порівнянні з рішеннями інших авторів, має низку переваг. А саме дозволяє: моделювати різноваріативні сценарії реалізації кіберзагроз для КВОІ і їх наслідки; визначати внесок від кожного з факторів або компонент архітектури ІБ КВОІ на загальну картину ймовірності реалізації кіберзагрози для КВОІ; моделювати взаємодію всіх факторів ІБ і при необхідності візуалізувати цю взаємодію; розраховувати і в подальшому ранжувати значення ймовірностей кіберзагроз для КВОІ для конкретних сценаріїв реалізації загроз; автоматизувати за рахунок застосування розробленого ПО процеси моделювання загроз і значно скоротити час на аудит загроз. Показано, що використання методу змістовної ідентифікації, дозволяє підвищити адекватність моделей обраної предметної області, а також запобігти помилкове введення в БЗ ЕС однакових за змістом суджень експертів і

цілей, зокрема при об'єднанні ієрархій цілей, сформованих різними експертними групами. Показано, що метод також може використовуватися для пошуку цілей ієрархії, точні формулювання яких, за ключовими словами, невідомі.

**Ключові слова:** інформаційна безпека; експертна система; база знань; силогізм; логіка предикатів; метод змістовної ідентифікації об’єктів.

## 1. ВСТУП

**Постановка проблеми.** Організація системи інформаційної безпеки (СІБ) сьогодні стає важливим стратегічним чинником розвитку для багатьох компаній і підприємств, інформаційні системи яких потрапляють під визначення критично важливих або КВІС [1, 2]. А оскільки складність архітектури більшості критично важливих об'єктів інформатизації (далі КВОІ) знижує результативність і достовірність звичайної експертної оцінки загроз і ризиків для інформаційної та кібербезпеки (далі, ІБ і КБ, відповідно) КВОІ, то на думку багатьох фахівців [3, 4], доцільно для вирішення цього завдання використовувати потенціал інтелектуалізованих систем підтримки прийняття рішень (СППР) або експертних систем (ЕС) [5]. Подібні СППР і ЕС, як показує досвід їх застосування [6, 7], особливо коли мова йде про слабоструктуровані ознаки загроз і кібератак, здатні в достатній мірі перебрати на себе рутинні завдання, що пов'язані з оцінкою поточних кіберзагроз та вразливостей, що дозволить розвантажити персонал служб ІБ, дозволивши їм зосередитися на більш пріоритетних завданнях щодо забезпечення сталого і стабільного функціонування корпоративних інформаційних систем, в тому, числі КВІС.

Вище викладене, дозволяє стверджувати, що тема дослідження, результати якого представлені в даній роботі, є актуальними.

**Аналіз останніх досліджень і публікацій.** Як було показано в роботах [8, 9] побудова надійної системи захисту інформації (СЗІ) для КВОІ часто залежить від правильного розпізнавання і подальшої оцінки загроз ІБ, з подальшою актуалізацією найбільш пріоритетних з точки зору ІБ і КБ. При цьому як було показано в [9-11] завдання по визначенню ймовірності реалізації кіберзагроз є одним із пріоритетних в процесі оцінювання ризиків для ІБ і КБ КВОІ. Однак, зауважимо, що багато з розглянутих вище робіт [9-14] не містять моделей для прогнозної оцінки розвитку ситуації, коли загроза реалізована. Більш того, в проаналізованих джерелах [2, 4, 7, 10, 11] мало розглядається можливість помилкового введення однакових або близьких за змістом суджень експертів. Зокрема, при об'єднанні кількох суджень, сформульованих експертними групами різної спеціалізації. Також можлива ситуація, коли різними формулюваннями описується одна і та ж ситуація (мета) [12-15]. Для обліку цих особливостей знань пропонується використовувати метод змістовної ідентифікації об'єктів баз знань систем підтримки прийняття рішень в ході процедури загроз для ІБ і КБ КВОІ [12, 16-18].

Таким чином, як показав аналіз аналогічних досліджень в даному напрямку [19-22], тематика роботи, пов'язана з удосконаленням методологічної бази розробки експертних систем і, зокрема, баз знань для них, в задачах оцінки загроз інформаційній безпеці різних об'єктів інформатизації, залишається релевантною і вимагає подальшого вивчення.

**Мета статті.** Метою даного дослідження є вдосконалення методології проектування баз знань (БЗ) для експертних і систем підтримки прийняття рішень при

оцінці кіберзагроз для КВОІ на основі силогізмів, логіки предикатів і методу змістовної ідентифікації об'єктів БЗ.

Для досягнення поставленої мети вирішуються завдання по:

- розробці методики моделювання суджень експертів при оцінці загроз ІБ і КБ КВОІ на основі силогізмів і логіки предикатів;
- розвитку методу змістовної ідентифікації об'єктів баз знань ЕС в ході процедури оцінювання загроз для ІБ і КБ КВОІ.

## 2. МЕТОДИ І МОДЕЛІ.

При проектуванні ЕС в задачах оцінки кіберзагроз та ризиків для ІБ КВОІ одним з природних етапів є формування аналітиками анкет або опитувальних листів, які містять типові питання, що задаються при аудиті ІБ. Наприклад, до таких питань можна віднести - чи мали місце інциденти, володіє організація або компанія конфіденційною інформацією тощо [8, 10, 23-26]. Далі експерт чи аналітик по ІБ формує унарні і / або бінарні висловлювання. Подібні висловлювання дозволяють будувати так звані сорити тобто ланцюг послідовних силогізмів. В контексті проектування бази знань (БЗ) для ЕС в задачах ІБ, силогізм - це двопосильний умовивід, який складається з атрибутивних висловлювань. Далі використовуючи апарат числення предикатів і семантичні мережі можна сформувати моделі для актуальних кіберзагроз конкретному КВОІ. Загалом ключові етапи алгоритму формування переліку кіберзагроз для КВОІ можна представити у вигляді схеми, яка показана на рис. 1.

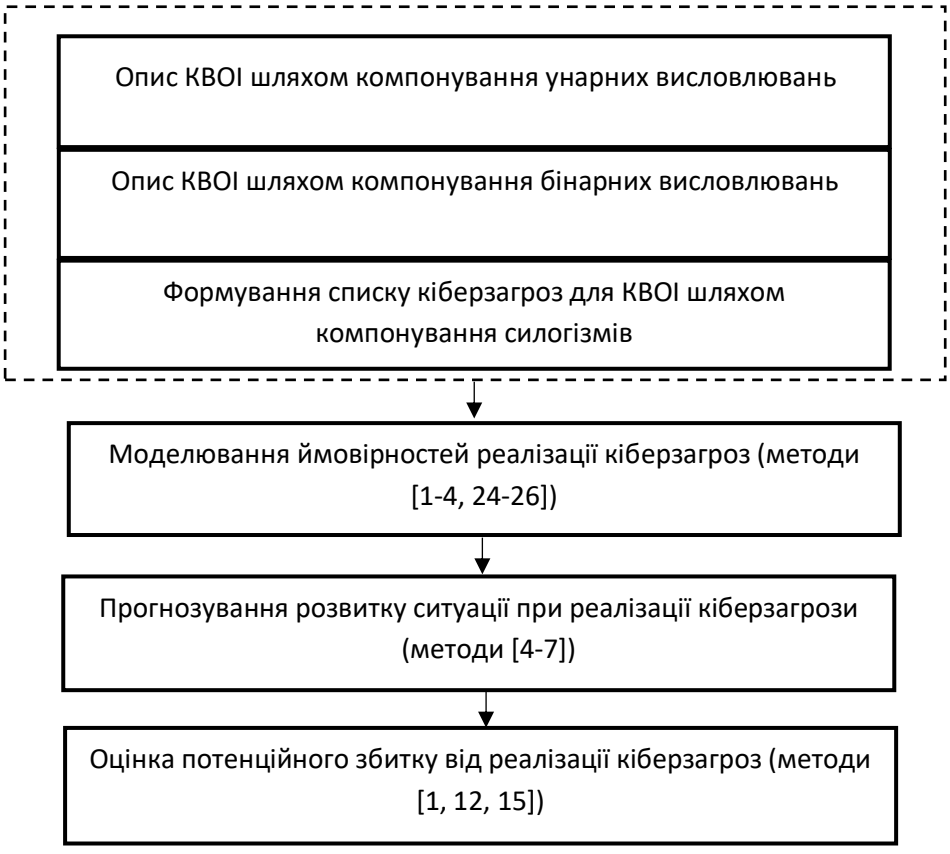


*Рис. 1 Загальна схема взаємодії алгоритмів оцінки загроз та ризиків ІБ КВОІ*

ISSN 2663 - 4023

При проектуванні БЗ використовувалися такі базові унарні висловлювання (в рамках статті наведені частково): 0 - виконання шкідливого програмного забезпечення (ПЗ); 1 - використання входів USB; 2 - використання гнучких дисків; 3 - наявність виходу в Інтернет; 4 - наявність виходу в локальну обчислювальну мережу (ЛОМ); 5 - наявність CD / DVD; 6 - відсутність оновлень антивірусів і сигнатур; 7 - відсутність системи комплексного ЗІ; 8 - відсутність антивіруса; 9 - відсутність інструкцій для відповідального за ІБ і КБ; 10 - відсутність в інструкціях для адміністратора ІБ; 11 - відсутність технологічних процесів ЗІ; 12 - відсутність інструкцій для засобів антивірусного ЗІ; 13 - відсутність акту по установці засобів ЗІ; 14 - витік ключів і атрибутів доступу; 15 - відсутність файлу (файлів) резервної копії; 16 - відсутність елементів в інструкціях користувачам; 17 - факт розголошення інформації; 18 - відсутність договору про нерозголошення інформації співробітниками; 19 - загроза зараження ЛОМ вірусним ПЗ; 20 - загроза перехоплення паролів в ЛОМ; 21 - відсутність фаєрволу; 22 - 100 і інше, включаючи резерв.

Даний підхід розглянемо на прикладі побудови ланцюгу послідовних силогізмів, які оперують вихідними висловлюваннями в області мережевої безпеки.

Складемо найпростіші унітарні висловлювання, які описують характеристики об’єкту інформатизації: 1) $S0-$ аналізуємо КВОІ; 2) $S1-$КВОІ використовує підключення до Інтернету; 3) $S2-$ КВОІ не має загрози хакерських атак; 4) $S3-$ КВОІ схильний до загроз розкрадання інформації; 5) $S4-$КВОІ схильний до загроз, що пов’язані із знищенням інформації; 6) $S5-$ КВОІ схильний до загроз, що пов’язані із зараженням шкідливим ПЗ.

Складемо бінарні висловлювання, які будуть описувати вразливості КВОІ. Бінарні висловлювання складені на основі унітарних висловлювань і при цьому також використовувалась логіка кванторів, див. рис. 2: $A,S0,S1-$ КВОІ, що аналізується – це КВОІ, який використовує підключення до Інтернету. Або: $A,S1,S2-$ будь-який КВОІ, що використовує підключення Інтернету – це КВОІ, що має загрози хакерських атак; $A,S2,S3-$ КВОІ, що має загрози хакерських атак – це КВОІ, схильний до загрозам розкрадання інформації; $A,S2,S4-$ КВОІ, що має загрози хакерських атак – це КВОІ, схильний до загрозам знищення інформації; $A,S2,S5-$ КВОІ, схильний до загрозам хакерських атак – це КВОІ, схильний до загроз зараження шкідливим ПЗ.

Таким чином, бінарні висловлювання - це одне з посилань силогізму з відповідними кванторами. Силогізми можна поєднувати в різних комбінаціях. Кожне таке поєднання дозволяє отримувати новий висновок. Аналогічно можна поєднувати між собою отримані висновки. Подібні операції можна продовжувати до моменту поки перехід від сильних властивостей об'єкту аналізу, що змінюється (або в термінах силогізмів модуси), до слабких не перерве процеси синтезу нових висновків.

Так, приклад, показаний на рис. 2, дозволяє зробити очевидний висновок. Якщо автоматизувати процес синтезу зв'язків між силогізмами, то можна за допомогою ЕС визначати тільки безпосередньо перелік загроз для КВОІ, а й актуалізувати їх для поточного моменту часу. А на виході, отримавши перелік актуальних загроз, на основі застосування вже наявних моделей інших авторів [1-4, 14-17], або результатів наших попередніх досліджень [1-4, 5-7, 13], ми можемо приступити до визначення ймовірностей реалізації кіберзагроз для конкретного КВОІ [14].

ISSN 2663 - 4023

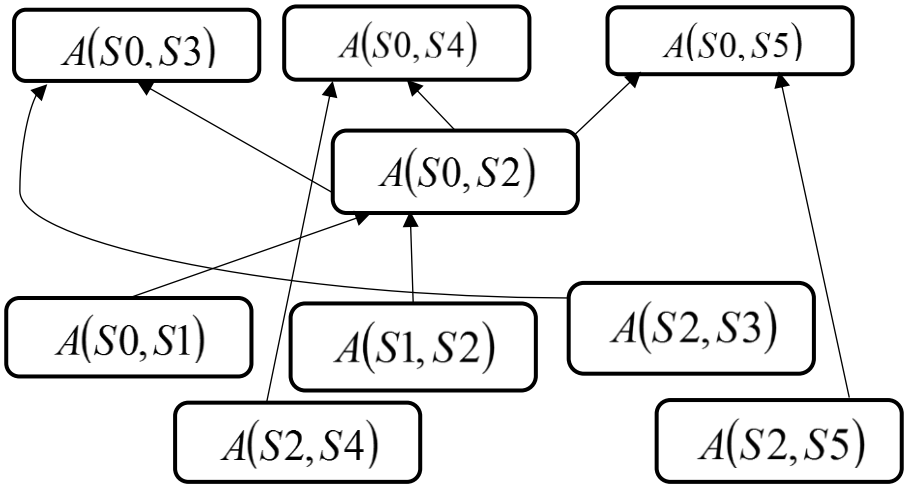


*Рис. 2 Приклад бінарних висловлювань в експертній системі*

Зауважимо, що в моделі подачі знань, що представлена на рис. 2, відсутня семантична метрика і не завжди може бути враховані внутрішні знання, що інтерпретуються, які стосуються нових загроз для ІБ і КБ КВОІ. Для обліку цих особливостей знань пропонується використовувати метод змістовної ідентифікації об'єктів баз знань ЕС.

Завдання змістовної ідентифікації об'єктів в БЗ можна звести до наступної постановки:

Дано:

1) Множина всіх об’єктів БЗ $G = \|G_i\|_{i=\overline{1,m}}$, де $m$ – кількість об’єктів в БЗ; кожний об’єкт в БЗ, відповідно, можна подати так $G_i = \|g_{ij}\|_{i=\overline{1,m};\ j=\overline{1,n}}$, де $g_{ij}$ – ваговий коефіцієнт ключового слова (КС) с номером $j$ в об’єкті $i$ БЗ СППР; $n$ – загальна кількість ключових слів, при цьому $0 \le g_{ij} \le 1$, $i = \overline{1,m};\ j = \overline{1,n}$, $\sum_{i=1}^{n} g_{ij} = 1$;

2) Симетрична матриця відмінностей КС між собою – $D$, де $D = \|d_{ij}\|_{i=\overline{1,n};\ j=\overline{1,n}}$, $0 \le d_{ij} \le 1 \ \&\ d_{ij} = d_{ji}, i = \overline{1,n};\ j = \overline{1,n}$;

3) Відносне відхилення – $\varepsilon$, де $0 < \varepsilon \le 1$.

4) Необхідно визначити: які об’єкти БЗ є подібними для об'єкту $G_i$ в межах заданого відхилення $\varepsilon$.

Розглянемо сутність методу. При формулюванні мети експертом, для характеристики її змісту і забезпечення однозначної відповідності змісту уявленням експерта, задається кортеж КС з відповідними коефіцієнтами важливості. Цей кортеж КС з відповідними вагами може бути:

1. заданий експертним шляхом;
2. сформульований на основі нормативних документів в області інформаційної безпеки;
3. заснований на раніше побудованих графах бінарних висловлювань або ієрархіях;
4. заснований на різних змінних, які використовуються в моделях БЗ.
5. інше.

Кортеж КС об’єкту БЗ повинен бути мінімальній потужності і при цьому утримувати всю необхідну інформацію для його ідентифікації. Коефіцієнти важливості КС для об'єкта БЗ переважно [5,6] визначати експертним шляхом. Вагові коефіцієнти КС можуть визначатися матрицею парних порівнянь, або, наприклад, методами власного вектора [6], комбінаторними методами [11, 12], або іншими методами обробки експертної інформації [11, 12, 23-26]. Будується загальна для предметної області матриця відмінностей КС, що заснована на порівняннях всіх КС між собою по семантичному схожості.

Нехай $K$ – це множина всіх КС БЗ. Тоді побудований вищезгаданим чином об’єкт $G_i$, є нечіткою підмножиною [12, 14] множини $K$ $(G_i \subset K)$. Значення вагових коефіцієнтів $g_{ij}$ об’єкту $G_i$ задають таблицею значень, що відповідають функції належності $\mu_{G_i} \equiv g_{ij}$. Таким чином, об’єкту $G_i$ відповідає наступна сукупність пар:

$$\begin{aligned} &\left\{\left(K_j,\ \mu_{G_i}\left(K_j\right)\right)\middle| K_j \in K\right\} = \\ &= \left\{\left(K_j,\ g_{ij}\right)\middle| K_j \in K\right\},\ j = \overline{1,n}, \end{aligned} \tag{1}$$

де $K_j$ – $j$ – е КС.

Тоді $K_j$ можна розглядати як нечітка множина з наступною функцією належності $\mu_{K_j}$:

$$\mu_{K_j}\left(K_i\right) = \begin{cases} 0,\ i \neq j,\ i = \overline{1,n}; \\ 1,\ i = j,\ i = \overline{1,n}. \end{cases} \tag{2}$$

Візьмемо безліч добутків нечітких множин $K_j$ на відповідні коефіцієнти важливості КС $\left\{g_{ij} \times K_j,\ j = \overline{1,n},\ i = \overline{1,m}\right\}$ як бази [19] нечіткої топології [12, 14] $\tau$ на множину $K$. Згідно [12], база топології простору - це сукупність його відкритих підмножин. При цьому будь-яка відкрита множина - це сума деякого числа розглянутих підмножин. Топологію простору можна, наприклад, ставити вказуючи в цьому просторі деяку її базу. Тоді відповідно до [12], ця топологія збігається з сукупністю множин, що представлена як сума множин з цієї бази.

ISSN 2663 - 4023

Пара $(K,\tau)$ є кінцевим нечітким топологічним простором у відповідності [16, 18, 19]: $K-$множина; $\tau-$нечітка топологія для множини, тобто деяке сімейство його нечітких підмножин, яка задовольняє трьом наступним аксіомам:

$$\begin{aligned} &1)\ 0,1\in\tau;\\ &2)\ if\ U,V\in\tau\ then\ U\wedge V\in\tau;;\\ &3)\ if\ U_i\in\tau\ \ for\ everyone\ i\in I\ then\ \ \vee_i U_i\in\tau, \end{aligned} \qquad (3)$$

де $U,V-$елементи нечітких підмножин множини $K$.

Тоді, оскільки всі об'єкти БЗ представляються сумою деякого кінцевого числа елементів бази топології $\tau$, $G\subset\tau$, тобто, якщо кожний об'єкт БЗ належить топології, то й вся множина об'єктів БЗ належить їй. Всі об'єкти БЗ є відкритими множинами, оскільки вони належать топології [12]. Як було сказано вище, для подання змісту кожного об'єкта БЗ використовується кортеж КС з відповідними нормованими ваговими коефіцієнтами, тому всі об'єкти БЗ лежать на симплексі в просторі $(K,\tau)$.

Для визначення змістовної відстані між КС потрібно ввести метрику в описаному вище просторі об'єктів БЗ. Використовуємо підхід, що аналогічний знаходженню відстані по таблиці вимірювань з ієрархічного кластерного аналізу [7]. Наприклад, в [5] ми розглядали модель для процедури оцінювання інвестиційних проектів в сфері кібербезпеки, а конкретно модель для задачі вибору раціональної фінансової стратегії інвестора. Однак по мірі розвитку ЕС, оскільки вона носить відкритий характер, БЗ, і базу моделей можна доповнювати і іншими моделями, наприклад, такими [6,7]:

- конфліктної взаємодії сторін (інвесторів-гравців з боку захисту і з боку атаки), що функціонують в дискретні моменти часу;
- конфліктної взаємодії сторін, що функціонують неперервно в часі;
- інвестування в умовах неповної інформації про фінансовий стан другої сторони (стороні захисту заздалегідь не відомо якими фінансовими ресурсами володіє атакуюча сторона);
- інвестування з процедурою отримання додаткових даних інвесторами з боку захисту інформації;
- прийняття оптимального рішення щодо інвестування в системи кіберзахисту в умовах активної протидії з боку атакуючих;
- та інші.

Тоді фрагмент переліку ключових слів (КС) для БЗ і, відповідно об'єкти, можуть виглядати так: КС1 - дискретні моменти часу; КС2 - неперервно в часі; КС3 - не повна інформація; КС4 - додаткові дані; КС5 - активна протидія.

Головним моментом в кластерному аналізі параметрів об'єкта захисту вважається вибір метрики, від якої залежить кінцевий варіант розбиття об'єктів на групи при заданому алгоритмі розбиття. Вибір метрики впливає на форму кластерів [7]. Для кожної конкретної задачі метрика обирається з урахуванням головних цілей дослідження. Також до уваги приймається фізична та статистична природа інформації, що використовується.

Для кожного об'єкту БЗ, як одиницю виміру кожної характеристики візьмемо відповідні значення вагових коефіцієнтів КС $g_{ij}, i=\overline{1,n}$, як значущості (вагомості)

ISSN 2663 - 4023

кожної з характеристик. Відповідно, сумарна відмінність кожного КС від інших КС в БЗ складає $\sum_{i=1}^{n} d_{ij}$. Змістовна відстань $\rho$ між кожними двома об'єктами БЗ $G_k$ і $G_l$ можна знайти, наприклад, як зважену відстань Хеммінга:

$$\rho(G_k, G_l) = \sum_{i=1}^{n} \left( |g_{ki} - g_{li}| \sum_{j=1}^{n} d_{ij} \right). \quad (4)$$

Таким чином, можна побудувати кінцевий топологічний простір об'єктів БЗ (Чангівський підхід), як описано в [12, 13] та на ньому використовувати прийняту метрику.

Тепер обчислимо вектор відстаней від об'єкту $G_t$ до інших об'єктів БЗ – $\nu_t$:

$$\nu_t = \left\| \rho(G_k, G_l) \right\|_{i=\overline{1,m}}. \quad (5)$$

Для нормування поділимо знайдений вектор $\nu_t$ на його максимальний елемент $\max\limits_{i=\overline{1,m}} \{\nu_t\} \neq 0$ і отримаємо вектор $\nu_t^{norm}$:

$$\nu_t^{norm} = \frac{\nu_t}{\max\limits_{i=\overline{1,m}} \{\nu_t\}}. \quad (6)$$

Тоді множина $X_\varepsilon$ містить шукані об'єкти БЗ, що подібні $G_t$ в межах заданого відносного відхилення $\varepsilon$:

$$X_\varepsilon = \left\{ G_t \middle| \nu_i^{norm} \leq \varepsilon, i = \overline{1,m} \right\}. \quad (7)$$

Слід зазначити, що інформацію для матриці відмінностей КС між собою можна отримати одним з 3-х альтернативних способів:

1. за допомогою підходу, використаного в [6, 18] для оцінки компетентності експертів в групі. Матриця $D$ в цьому випадку виходить експертним шляхом. Цей підхід не є найзручнішим, адже, внаслідок великої кількості КС і психофізіологічних обмежень експерта доведеться розбивати безліч КС на групи для порівняння. Відповідно, знадобиться більший обсяг роботи експертів і засоби на їх оплату;
2. шляхом подання безлічі КС як семантичної мережі. Відмінність КС між собою в цьому випадку відповідає відстаням між її вузлами;
3. шляхом заповнення матриці з використанням ймовірнісної моделі виявлення зв'язків між поняттями (як поняття беруться КС) за допомогою інформаційно-аналітичних систем на основі обробки інформаційного потоку, сформованого, наприклад, під час пошуку в Інтернеті.

В результаті обчислення змістовних відстаней для кожної пари об'єктів БЗ отримаємо симетричну матрицю змістовних відстаней між усіма об'єктами БЗ. Далі, використовуючи цю матрицю, можна проводити кластеризацію об'єктів БЗ за змістом, що, зокрема, дасть можливість побачити і встановити відсутні зв'язку між подібними за

змістом об’єктами, з якими, наприклад, працювали різні за спеціалізацією і кваліфікації експерти.

При встановленні впливів в ієрархії цілей, коли експерту потрібно вибрати всі підцілі, що впливають на деякі цілі, метод може використовуватися для пошуку формулювань цілей з відомою експерту семантикою, але з невідомим формулюванням. У раніше побудованих графах або ієрархіях можна знаходити помилково введені семантично однакові цілі, а також формулювати відсутні, але необхідні за змістом цілі ієрархії. Також можливе використання методу в якості додаткового інструменту для аналізу адекватності БЗ предметної області - інформаційна безпека і кібербезпека.

Накопичені в процесі роботи методу дані (а саме КС і матрицю відмінностей КС) можна використовувати при обліку компетентності експертів (оскільки облік компетентності експертів в малих експертних групах є необхідною умовою забезпечення достовірності колективних рішень [21-23]), а також для формування експертних груп у відповідній предметній області.

На основі викладеного в статті підходу, проектується БЗ для СППР і ЕС в задачах кібербезпеки. Більш детально розроблювальне програмне забезпечення, яке засноване на подібному підході для секторальної ЕС в задачах аналізу ІБ і КБ КВОІ було описано в роботах [5, 6, 18, 19].

## 3. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

У процесі досліджень були отримані такі результати:

- розроблена методика моделювання суджень експертів при оцінці загроз інформаційної та кібербезпеки для критично важливих об'єктів інформатизації (КВОІ) на основі силогізмів і логіки предикатів, що дозволяє моделювати міркування експертів в ході аудиту ІБ КВОІ;
- запропонований метод змістовної ідентифікації об'єктів бази знань (БЗ) для експертної системи в задачах оцінки загроз інформаційній безпеці КВОІ.

Суть методу полягає в тому, що кожному об'єкту БЗ ставиться у відповідність кортеж ключових слів, значимість яких визначається експертним шляхом. Таким чином, кожен об'єкт БЗ ставиться у відповідність елементу кінцевого нечіткого топологічного простору об'єктів БЗ. Змістовна ідентифікація проходить по відстані між об'єктами БЗ, яка є метрикою в цьому просторі. Показано, що використання запропонованого методу змістовної ідентифікації, дозволяє підвищити адекватність моделей обраної предметної області, а також запобігти помилкове введення в БЗ ЕС однакових за змістом суджень експертів і цілей, зокрема при об'єднанні ієрархій цілей, сформованих різними експертними групами. Показано, що запропонований метод також може використовуватися для пошуку цілей ієрархії, точні формулювання яких, за ключовими словами, невідомі.

## 4. ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Запропонований в роботі підхід, в порівнянні з рішеннями інших авторів [1, 4, 14, 16], на наш погляд має низку переваг. А саме дозволяє: моделювати різноваріативні сценарії реалізації кіберзагроз для КВОІ і їх наслідки; визначати внесок від кожного з факторів або компонент архітектури ІБ КВОІ на загальну картину ймовірності реалізації кіберзагрози для КВОІ; моделювати взаємодію всіх факторів ІБ і при необхідності візуалізувати цю взаємодію; розраховувати і в подальшому ранжувати

значення ймовірностей кіберзагроз для КВОІ для конкретних сценаріїв реалізації загроз; автоматизувати за рахунок застосування розробленого ПО процеси моделювання загроз і значно скоротити час на аудит загроз.

На теперішньому етапі досліджень певним недоліком роботи є невеликий обсяг тестових вибірок для перевірки працездатності ПЗ [5, 6, 18, 19]. Однак у міру збільшення кількості результатів експериментальних тестових перевірок, проектованої нами секторальної експертної системи в задачах аудиту загроз для КВОІ, ми плануємо усунути цей недолік.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Davies, J., Milward, D., Wang, C. W., & Welch, J. (2015). Formal model-driven engineering of critical information systems. Science of Computer Programming, 103, 88-113. DOI: https://doi.org/10.1016/j.scico.2014.11.004

[2] Trauth, E. (2017). A research agenda for social inclusion in information systems. ACM SIGMIS Database: the Database for Advances in Information Systems, 48(2), 9-20. DOI: https://doi.org/10.1145/3084179.3084182

[3] Shahbazian, E., & Rogova, G. (2016, November). Critical Aviation Information Systems Cybersecurity. In Meeting Security Challenges Through Data Analytics and Decision Support (Vol. 47, p. 308). IOS Press.

[4] Paradice, D., Freeman, D., Hao, J., Lee, J., & Hall, D. (2018). A Review of Ethical Issue Considerations in the Information Systems Research Literature. Foundations and Trends® in Information Systems, 2(2), 117-236. DOI: http://dx.doi.org/10.1561/2900000012

[5] Akhmetov, B., Lakhno, V., Malyukov, V., Sarsimbayeva, S., Zhumadilova, M., Kartbayev, T. (2019). Decision support system about investments in smart city in conditions of incomplete information, International Journal of Civil Engineering and Technology, 10 (2), pp. 661-670.

[6] Akhmetov, B., Lakhno, V., Akhmetov, B., Alimseitova, Z. (2019). Development of sectoral intellectualized expert systems and decision making support systems in cybersecurity, Advances in Intelligent Systems and Computing, 860, pp. 162-171. DOI: https://doi.org/10.1007/978-3-030-00184-1_15

[7] Lakhno, V., Zaitsev, S., Tkach, Y., Petrenko, T. (2019). Adaptive expert systems development for cyber attacks recognition in information educational systems on the basis of signs' clustering, Advances in Intelligent Systems and Computing, 754, pp. 673-682. DOI: https://doi.org/10.1007/978-3-319-91008-6_66

[8] Li, K., Wen, H., Li, H., Zhu, H., & Sun, L. (2018, October). Security OSIF: Toward Automatic Discovery and Analysis of Event Based Cyber Threat Intelligence. In 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI) (pp. 741-747). IEEE. DOI: https://doi.org/10.1109/SmartWorld.2018.00142

[9] Moulin, M., Eyisi, E., Shila, D. M., & Zhang, Q. (2018, October). Automatic Construction of Attack Graphs in Cyber Physical Systems Using Temporal Logic. In MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM) (pp. 933-938). IEEE. DOI: https://doi.org/10.1109/MILCOM.2018.8599799

[10] Kashyap, A. K., & Wetherilt, A. (2019, May). Some principles for regulating cyber risk. In AEA Papers and Proceedings (Vol. 109, pp. 482-87). DOI: https://doi.org/10.1257/pandp.20191058

[11] Mishina, Y., Takaragi, K., & Umezawa, K. (2018, October). A Method of Threat Analysis for Cyber-Physical System using Vulnerability Databases. In 2018 IEEE International Symposium on Technologies for Homeland Security (HST) (pp. 1-7). IEEE. DOI: https://doi.org/10.1109/THS.2018.8574154

[12] Chang, Chin-Liang. "Fuzzy topological spaces." *Journal of mathematical Analysis and Applications* 24.1 (1968): 182-190.

[13] Azad, K. K. "On fuzzy semicontinuity, fuzzy almost continuity and fuzzy weakly continuity." *Journal of Mathematical Analysis and Applications* 82.1 (1981): 14-32. DOI: https://doi.org/10.1016/0022-247X(81)90222-5

[14] Lowen, R. "Fuzzy topological spaces and fuzzy compactness." *Journal of Mathematical analysis and applications* 56.3 (1976): 621-633. DOI: https://doi.org/10.1016/0022-247X(76)90029-9

[15] Moldoveanu, Mihnea C., Joel AC Baum, and Tim J. Rowley. "Information regimes, information strategies and the evolution of interfirm network topologies." *Multi-level issues in organizational behavior and strategy*. Emerald Group Publishing Limited, 2003. 221-264. DOI: https://doi.org/10.1016/S1475-9144(03)02014-9

[16] Wu, Ing-Long, and Han-Chang Lin. "A strategy-based process for implementing knowledge management: An integrative view and empirical study." *Journal of the American Society for Information Science and Technology* 60.4 (2009): 789-802. DOI: https://doi.org/10.1002/asi.20999

[17] Pal, Ranjan, and Pan Hui. "Modeling internet security investments: Tackling topological information uncertainty." *International Conference on Decision and Game Theory for Security*. Springer, Berlin, Heidelberg, 2011. DOI: https://doi.org/10.1007/978-3-642-25280-8_18

[18] Kasabov, Nikola K. *Foundations of neural networks, fuzzy systems, and knowledge engineering*. Marcel Alencar, 1996.

[19] Raban, Y., & Hauptman, A. (2018). Foresight of cyber security threat drivers and affecting technologies. foresight, 20(4), 353-363. DOI: https://doi.org/10.1108/FS-02-2018-0020

[20] Lakhno, V., Kasatkin, D., Kozlovskyi, V., Petrovska, S., Boiko, Y., Kravchuk, P., Lishchynovska, N. (2019). A model and algorithm for detecting spyware in medical information systems, International Journal of Mechanical Engineering and Technology, (1), pp. 287-295.

[21] Mishina, Y., Takaragi, K., & Umezawa, K. (2018, October). A Method of Threat Analysis for Cyber-Physical System using Vulnerability Databases. In 2018 IEEE International Symposium on Technologies for Homeland Security (HST) (pp. 1-7). IEEE. DOI: https://doi.org/10.1109/THS.2018.8574154

[22] Petrenko, S. (2018). Possible Scientific-Technical Solutions to the Problem of Giving Early Warning. In Big Data Technologies for Monitoring of Computer Security: A Case Study of the Russian Federation (pp. 175-218). Springer, Cham. DOI: https://doi.org/10.1007/978-3-319-79036-7_4

[23] Evangelopoulou, M., & Johnson, C. W. (2015, June). Empirical framework for situation awareness measurement techniques in network defense. In 2015 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA) (pp. 1-4). IEEE.

[24] Herley, C., & Van Oorschot, P. C. (2017, May). Sok: Science, security and the elusive goal of security as a scientific pursuit. In 2017 IEEE Symposium on Security and Privacy (SP) (pp. 99-120). IEEE. DOI: https://doi.org/10.1109/SP.2017.38

[25] Akhmetov, B., Lakhno, V. (2018). System of decision support in weaklyformalized problems of transport cybersecurity ensuring, Journal of Theoretical and Applied Information Technology, 96 (8), pp. 2184-2196.

[26] Akhmetov, B., Lakhno, V., Boiko, Y., Mishchenko, A. (2017). Designing a decision support system for the weakly formalized problems in the provision of cybersecurity, Eastern-European Journal of Enterprise Technologies, 1 (2-85), pp. 4-15.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (8), 2020
ISSN 2663 - 4023

**Valerii A. Lakhno**
Dr. Tech. Sc., Professor, Head of the Departament of Computer System and Networks
National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine
ORCID: 0000-0001-9695-4543
*valss21@ukr.net*

**Dmytro Y. Kasatkin**
Cand. Pedag. Sc. (Ph.D.), Docent, Associate Professor at the Department of Computer System and Networks
National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine
ORCID: 0000-0002-2642-8908
*d.kasatkin@nubip.edu.ua*

**Andii I. Blozva**
Cand. Pedag. Sc. (Ph.D.), Docent, Associate Professor at the Department of Computer System and Networks
National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine
ORCID: 0000-0002-4377-0916
*andriy.blozva@nubip.edu.ua*

**Maksym D. Misiura**
Cand. Tech. Sc. (Ph.D), Associate Professor at the Department of Computer System and Networks
National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine
ORCID: 0000-0002-9061-3462
*mdm@nubip.edu.ua*

**Borys S. Husiev**
Cand. Tech. Sc. (Ph.D), Docent, Associate Professor at the Department of Computer System and Networks
National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine
ORCID: 0000-0003-1658-7822
*gusevbs@nubip.edu.ua*

# DESIGN OF KNOWLEDGE BASE FOR CYBER SECURITY SYSTEMS ON THE BASIS OF SUBJECT IDENTIFICATION METHOD

**Abstract.** The article presents the results of research performed in the process of designing an expert system (ES) designed to assess the threats to information security (IS) of critical information facilities (CIF). The approach to designing of expert system on the basis of syllogisms and logic of predicates, and also a method of meaningful identification of objects of knowledge base (KB) is offered. The essence of the method is that each object of the database of the projected EU, is matched by a tuple of keywords (ToK), the significance of which is determined by experts. Thus, each database object is placed in accordance with the element of the finite fuzzy topological space of the database objects. Meaningful identification takes place on the distance between the objects of the database. The approach proposed in the work, in comparison with the decisions of other authors, has a number of advantages. Namely, it allows: to model different variants of cyber threat scenarios for CIF and their consequences; determine the contribution of each of the factors or components of the architecture of the IS CIF to the overall picture of the probability of a cyber threat to the CIF; model the interaction of all IS factors and, if necessary, visualize this interaction; calculate and further rank the values of cyber threat probabilities for CIF for specific threat scenarios; automate the processes of threat modeling through the use of developed software and significantly reduce the time for audit of threats. It is shown that the use of the method of meaningful identification allows to increase the adequacy of the models of the selected subject area, as well as to prevent erroneous introduction of the same judgments of experts and goals in the EU database, in particular by combining hierarchies of goals formed by different expert groups. It is shown that the method can also be used to find the goals of the hierarchy, the exact wording of which, according to keywords, is unknown.

**Keywords:** informational security; expert system; knowledge base; syllogism; predicate logic; method of meaningful identification of objects.

## REFERENCES

[1] Davies, J., Milward, D., Wang, C. W., & Welch, J. (2015). Formal model-driven engineering of critical information systems. Science of Computer Programming, 103, 88-113. DOI: https://doi.org/10.1016/j.scico.2014.11.004

[2] Trauth, E. (2017). A research agenda for social inclusion in information systems. ACM SIGMIS Database: the Database for Advances in Information Systems, 48(2), 9-20. DOI: https://doi.org/10.1145/3084179.3084182

[3] Shahbazian, E., & Rogova, G. (2016, November). Critical Aviation Information Systems Cybersecurity. In Meeting Security Challenges Through Data Analytics and Decision Support (Vol. 47, p. 308). IOS Press.

[4] Paradice, D., Freeman, D., Hao, J., Lee, J., & Hall, D. (2018). A Review of Ethical Issue Considerations in the Information Systems Research Literature. Foundations and Trends® in Information Systems, 2(2), 117-236. DOI: http://dx.doi.org/10.1561/2900000012

[5] Akhmetov, B., Lakhno, V., Malyukov, V., Sarsimbayeva, S., Zhumadilova, M., Kartbayev, T. (2019). Decision support system about investments in smart city in conditions of incomplete information, International Journal of Civil Engineering and Technology, 10 (2), pp. 661-670.

[6] Akhmetov, B., Lakhno, V., Akhmetov, B., Alimseitova, Z. (2019). Development of sectoral intellectualized expert systems and decision making support systems in cybersecurity, Advances in Intelligent Systems and Computing, 860, pp. 162-171. DOI: https://doi.org/10.1007/978-3-030-00184-1_15

[7] Lakhno, V., Zaitsev, S., Tkach, Y., Petrenko, T. (2019). Adaptive expert systems development for cyber attacks recognition in information educational systems on the basis of signs' clustering, Advances in Intelligent Systems and Computing, 754, pp. 673-682. DOI: https://doi.org/10.1007/978-3-319-91008-6_66

[8] Li, K., Wen, H., Li, H., Zhu, H., & Sun, L. (2018, October). Security OSIF: Toward Automatic Discovery and Analysis of Event Based Cyber Threat Intelligence. In 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI) (pp. 741-747). IEEE. DOI: https://doi.org/10.1109/SmartWorld.2018.00142

[9] Moulin, M., Eyisi, E., Shila, D. M., & Zhang, Q. (2018, October). Automatic Construction of Attack Graphs in Cyber Physical Systems Using Temporal Logic. In MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM) (pp. 933-938). IEEE. DOI: https://doi.org/10.1109/MILCOM.2018.8599799

[10] Kashyap, A. K., & Wetherilt, A. (2019, May). Some principles for regulating cyber risk. In AEA Papers and Proceedings (Vol. 109, pp. 482-87). DOI: https://doi.org/10.1257/pandp.20191058

[11] Mishina, Y., Takaragi, K., & Umezawa, K. (2018, October). A Method of Threat Analysis for Cyber-Physical System using Vulnerability Databases. In 2018 IEEE International Symposium on Technologies for Homeland Security (HST) (pp. 1-7). IEEE. DOI: https://doi.org/10.1109/THS.2018.8574154

[12] Chang, Chin-Liang. "Fuzzy topological spaces." *Journal of mathematical Analysis and Applications* 24.1 (1968): 182-190.

[13] Azad, K. K. "On fuzzy semicontinuity, fuzzy almost continuity and fuzzy weakly continuity." *Journal of Mathematical Analysis and Applications* 82.1 (1981): 14-32. DOI: https://doi.org/10.1016/0022-247X(81)90222-5

[14] Lowen, R. "Fuzzy topological spaces and fuzzy compactness." *Journal of Mathematical analysis and applications* 56.3 (1976): 621-633. DOI: https://doi.org/10.1016/0022-247X(76)90029-9

[15] Moldoveanu, Mihnea C., Joel AC Baum, and Tim J. Rowley. "Information regimes, information strategies and the evolution of interfirm network topologies." *Multi-level issues in organizational behavior and strategy*. Emerald Group Publishing Limited, 2003. 221-264. DOI: https://doi.org/10.1016/S1475-9144(03)02014-9

[16] Wu, Ing-Long, and Han-Chang Lin. "A strategy-based process for implementing knowledge management: An integrative view and empirical study." *Journal of the American Society for Information Science and Technology* 60.4 (2009): 789-802. DOI: https://doi.org/10.1002/asi.20999

[17] Pal, Ranjan, and Pan Hui. "Modeling internet security investments: Tackling topological information uncertainty." *International Conference on Decision and Game Theory for Security*. Springer, Berlin, Heidelberg, 2011. DOI: https://doi.org/10.1007/978-3-642-25280-8_18

[18] Kasabov, Nikola K. *Foundations of neural networks, fuzzy systems, and knowledge engineering*. Marcel Alencar, 1996.
[19] Raban, Y., & Hauptman, A. (2018). Foresight of cyber security threat drivers and affecting technologies. foresight, 20(4), 353-363. DOI: https://doi.org/10.1108/FS-02-2018-0020
[20] Lakhno, V., Kasatkin, D., Kozlovskyi, V., Petrovska, S., Boiko, Y., Kravchuk, P., Lishchynovska, N. (2019). A model and algorithm for detecting spyware in medical information systems, International Journal of Mechanical Engineering and Technology, (1), pp. 287-295.
[21] Mishina, Y., Takaragi, K., & Umezawa, K. (2018, October). A Method of Threat Analysis for Cyber-Physical System using Vulnerability Databases. In 2018 IEEE International Symposium on Technologies for Homeland Security (HST) (pp. 1-7). IEEE. DOI: https://doi.org/10.1109/THS.2018.8574154
[22] Petrenko, S. (2018). Possible Scientific-Technical Solutions to the Problem of Giving Early Warning. In Big Data Technologies for Monitoring of Computer Security: A Case Study of the Russian Federation (pp. 175-218). Springer, Cham. DOI: https://doi.org/10.1007/978-3-319-79036-7_4
[23] Evangelopoulou, M., & Johnson, C. W. (2015, June). Empirical framework for situation awareness measurement techniques in network defense. In 2015 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA) (pp. 1-4). IEEE.
[24] Herley, C., & Van Oorschot, P. C. (2017, May). Sok: Science, security and the elusive goal of security as a scientific pursuit. In 2017 IEEE Symposium on Security and Privacy (SP) (pp. 99-120). IEEE. DOI: https://doi.org/10.1109/SP.2017.38
[25] Akhmetov, B., Lakhno, V. (2018). System of decision support in weaklyformalized problems of transport cybersecurity ensuring, Journal of Theoretical and Applied Information Technology, 96 (8), pp. 2184-2196.
[26] Akhmetov, B., Lakhno, V., Boiko, Y., Mishchenko, A. (2017). Designing a decision support system for the weakly formalized problems in the provision of cybersecurity, Eastern-European Journal of Enterprise Technologies, 1 (2-85), pp. 4-15.