УДК 04.056.53:004.89:519.17 DOI: 10.31673/2412-4338.2026.029105

Субач Ігор Юрійович
доктор технічних наук, професор, завідувач кафедри
Інститут спеціального зв'язку та захисту інформації Національного технічного університету України Київський політехнічний інститут імені Ігоря Сікорського, Київ, Україна
ORCID ID: 0000-0002-9344-713X
igor_subach@ukr.net

Копич Данило Олексійович
аспірант
Інститут спеціального зв'язку та захисту інформації Національного технічного університету України Київський політехнічний інститут імені Ігоря Сікорського, Київ, Україна
ORCID ID: 0009-0005-9809-546X
danyla.kopych@gmail.com

МЕТОД ВИЯВЛЕННЯ КІБЕРІНЦИДЕНТІВ У SIEM НА ОСНОВІ НЕЧІТКОГО ГІПЕРГРАФОВОГО ПОДАННЯ ПОДІЙ БЕЗПЕКИ ТА ЛІНІЙНОЇ МОДЕЛІ ІНЦИДЕНТНОСТІ

Анотація. *У статті запропоновано метод виявлення кіберінцидентів у журналах подій систем управління інформацією та подіями безпеки (SIEM), що ґрунтується на нечіткому гіперграфовому поданні подій безпеки та використанні лінійної моделі інцидентності події – приналежності події безпеки до кіберінциденту. На відміну від традиційних підходів, що використовують жорсткі кореляційні правила (сигнатурний аналіз) або прості статистичні пороги, запропонований метод дозволяє враховувати багатосутнісну природу подій, їх структурні взаємозв'язки та невизначеність інтерпретації. В основі підходу лежить подання подій безпеки у вигляді зважених гіперребер, що поєднують множину сутностей інформаційної інфраструктури (користувачів, хостів, процесів, IP-адрес та технік кібератак). Для кожної події обчислюється локальна інцидентність на основі лінійної моделі, яка агрегує такі характеристики, як рівень аномальності (відхилення від базової поведінки), критичність спрацювання правил безпеки, контекстну узгодженість, семантичну значущість та історичну підозрілість сутності. Процес виявлення кіберінциденту реалізовано шляхом пошуку зв'язних підструктур у гіперграфі, які формуються на основі часової близькості подій та нетривіального перетину множин їхніх сутностей. Кіберінцидент у межах методу інтерпретується як зв'язна підструктура у гіперграфі подій безпеки, для якої агрегована інцидентність перевищує задане порогове значення. Запропонований метод забезпечує системний перехід від ізольованого аналізу окремих подій до цілісного виявлення структурно пов'язаних сукупностей подій безпеки, що відповідають розвитку кібератаки. Наукова новизна отриманого результату полягає у поєднанні лінійної моделі інцидентності події з гіперграфовою кореляцією подій безпеки та формалізації кіберінциденту як зв'язної підструктури у нечіткому гіперграфі.*

Ключові слова: *кіберзахист, SIEM, кіберінцидент, нечіткий гіперграф, кореляція подій, інцидентність події.*

Ihor Subach
doctor of technical science, associate professor, head of department
Institute of special communications and information security
National technical university of Ukraine Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine
ORCID ID: 0000-0002-9344-713X
igor_subach@ukr.net

Danylo Kopych
postgraduate student
Institute of special communications and information security
National technical university of Ukraine Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine
ORCID ID: 0009-0005-9809-546X
danyla.kopych@gmail.com

METHOD FOR DETECTION OF CYBER INCIDENTS IN SIEM BASED ON FUZZY HYPERGRAPHIC REPRESENTATION OF SECURITY EVENTS AND LINEAR INCIDENT MODEL

Abstract. *The article proposes a method for detecting cyber incidents in event logs of security information and event management systems (SIEM), which is based on a fuzzy hypergraph representation of security events and the use of a linear model of event incidence - the belonging of a security event to a cyber incident. Unlike traditional approaches that use strict correlation rules (signature analysis) or simple statistical thresholds, the proposed method allows taking into account the multi-entity nature of events, their structural relationships and interpretation uncertainty. The approach is based on the representation of security events in the form of weighted hyperedges that connect a set of information infrastructure entities (users, hosts, processes, IP addresses and cyberattack techniques). For each event, the local incidence is calculated based on a linear model that aggregates such characteristics as the level of anomaly (deviation from basic behavior), the criticality of security rule triggering, contextual coherence, semantic significance, and historical suspiciousness of the entity. The process of detecting a cyber incident is implemented by searching for connected substructures in the hypergraph, which are formed based on the temporal proximity of events and the non-trivial intersection of the sets of their entities. A cyber incident within the method is interpreted as a connected substructure in the hypergraph of security events, for which the aggregated incidence exceeds a given threshold value. The proposed method provides a systematic transition from isolated analysis of individual events to holistic detection of structurally related sets of security events corresponding to the development of a cyber attack. The scientific novelty of the obtained result lies in the combination of a linear event incidence model with a hypergraph correlation of security events and the formalization of a cyber incident as a connected substructure in a fuzzy hypergraph.*

Keywords: *cyber defense, SIEM, cyber incident, fuzzy hypergraph, event correlation, event incidence.*

1. Вступ

У сучасних умовах функціонування інформаційно-комунікаційних систем (ІКС) засоби кіберзахисту типу SIEM формують великі масиви подієвих даних, що відображають як нормальну діяльність інформаційної інфраструктури, так і ознаки потенційно небезпечних впливів. Водночас ізольований аналіз окремої події, як правило, не забезпечує достатньої інформативності для обґрунтованого висновку щодо наявності кіберінциденту. Практика свідчить, що кібератаки реалізуються як узгоджені ланцюги подій, пов’язаних між собою за часовими, просторовими та логічними ознаками, утворюючи складні поведінкові шаблони.

Існуючі методи виявлення кіберінцидентів у SIEM здебільшого спираються на правила кореляції, сигнатурний аналіз або використання фіксованих порогових критеріїв. Проте такі підходи характеризуються обмеженою здатністю до відображення багатовимірної природи подій безпеки, недостатньо враховують складні структурні залежності між ними, а також практично не адаптовані до роботи з невизначеними та неповними даними. Додатково, застосування жорстких порогів зумовлює дискретний характер прийняття рішень, що зводиться до бінарної класифікації, яка не відображає формування кіберінцидентів у реальних умовах.

У попередніх дослідженнях [1] було запропоновано модель оцінювання інцидентності події безпеки, яка враховує сукупність статистичних, структурних, семантичних та історичних характеристик і дозволяє отримати градуйовану оцінку її належності до кіберінциденту. Проте зазначена модель орієнтована на аналіз окремих подій і не визначає механізму переходу від локальної оцінки інцидентності до виявлення кіберінцидентів як складних структурно пов’язаних процесів.

2. Постановка проблеми

У зв’язку з цим виникає наукове завдання щодо розроблення методу виявлення кіберінцидентів, який би, з одного боку, спирався на модель локальної інцидентності події, а з іншого – враховував структурні взаємозв’язки між подіями безпеки та дозволяв виявляти кіберінциденти як цілісні утворення.

Перспективним підходом до розв’язання цього завдання є використання гіперграфових моделей, які дозволяють природним чином відображати багатосутнісні взаємодії між подіями безпеки. Поєднання гіперграфового подання подій із нечіткою оцінкою їх інцидентності створює підґрунтя для побудови методу, що забезпечує виявлення кіберінцидентів як зв’язних підструктур у гіперграфі подій безпеки.

Отже, актуальним є завдання розроблення методу виявлення кіберінцидентів у SIEM на основі нечіткого гіперграфового подання подій безпеки та лінійної моделі інцидентності події, який забезпечує перехід від аналізу окремих подій до виявлення їх структурно пов’язаних сукупностей.

3. Аналіз останніх досліджень і публікацій

Задача ідентифікації кіберінцидентів у системах забезпечення кібербезпеки є предметом інтенсивних досліджень як у наукових роботах, так і в прикладних розробках. Наявні підходи до її розв’язання доцільно класифікувати за кількома ключовими напрямами, серед яких виділяють сигнатурні, статистичні, графоорієнтовані та нечіткі методи [2]-[8].

Сигнатурні та правило-орієнтовані механізми, що широко інтегровані в сучасні SIEM-рішення [8]-[10] (зокрема, платформи на кшталт Wazuh), базуються на використанні наперед визначених шаблонів поведінки та наборів кореляційних залежностей. Хоча такі інструменти демонструють високу результативність при виявленні вже відомих сценаріїв кібератак, їх застосування обмежене у випадках появи нових або трансформованих загроз. Крім того, подібні підходи, як правило, не забезпечують повноцінного врахування складних багаторівневих взаємозв'язків між подіями безпеки.

Методи, що спираються на статистичний аналіз і виявлення аномалій, орієнтовані на фіксацію відхилень від типової поведінки системи. Зокрема, у дослідженні [2] запропоновано графові підходи до детекції аномалій, які дозволяють враховувати структурні властивості даних. Водночас такі рішення переважно зосереджені на ідентифікації окремих аномальних елементів і не формалізують кіберінцидент як інтегровану сукупність взаємопов'язаних подій.

Останнім часом дедалі більшого поширення набувають графові моделі для аналізу подієвих потоків у системах безпеки, оскільки вони дають змогу враховувати взаємодії між об'єктами [4]-[5]. Проте класичні графові структури обмежені представленням лише попарних зв'язків. З метою подолання цього обмеження використовуються гіперграфові підходи, які забезпечують можливість моделювання взаємодій за участю декількох сутностей одночасно. Дослідження у цьому напрямі демонструють їхню ефективність для опису складних залежностей у кіберпросторі, однак у більшості випадків такі моделі не враховують невизначеність, притаманну інтерпретації подій безпеки.

Для моделювання невизначеності у системах виявлення вторгнень широко використовуються нечіткі методи. Зокрема, у роботах, присвячених нечітким системам виявлення кібератак [3], показано можливість інтеграції експертних знань і нечітких оцінок для підвищення точності їх виявлення. Однак у більшості випадків нечітка логіка застосовується до аналізу окремих подій або ознак і не поєднується зі структурним аналізом взаємозв'язків між подіями.

Таким чином, аналіз існуючих підходів свідчить про те, що:

- сигнатурні методи не забезпечують адаптивності до нових атак;
- статистичні та методи виявлення аномалій недостатньо враховують структуру взаємозв'язків подій безпеки;
- графові підходи обмежені попарними зв'язками або не враховують невизначеність;
- нечіткі методи не інтегровані зі структурними моделями подій безпеки.

Отже, актуальним є завдання розроблення методу, який поєднує:

- гіперграфове подання подій безпеки як засіб моделювання багатосутнісних взаємодій;
- нечітку (градуйовану) оцінку інцидентності подій безпеки;
- механізм виявлення кіберінцидентів як структурно пов'язаних сукупностей подій безпеки.

Запропонований у даній роботі метод спрямований на подолання зазначених обмежень шляхом інтеграції лінійної моделі інцидентності події безпеки [1] з їхньою кореляцією у SIEM на основі нечіткого гіперграфового представлення.

4. Мета і задачі дослідження

Метою роботи є розроблення методу виявлення кіберінцидентів у SIEM на основі нечіткого гіперграфового подання подій безпеки, що ґрунтується на лінійній моделі інцидентності події та забезпечує виявлення кіберінцидентів як структурно пов'язаних сукупностей подій.

Для досягнення поставленої мети було вирішено такі взаємопов'язані задачі:

- 1. Формалізовано представлення подій безпеки у вигляді зважених гіперребер нечіткого гіперграфу та застосовано лінійну модель інцидентності події безпеки для оцінювання ступеня її належності до множини подій, що можуть формувати кіберінцидент.
- 2. Розроблено механізм встановлення зв'язків між подіями безпеки на основі часової близькості та нетривіального перетину множин сутностей, з яких вони складаються.
- 3. Запропоновано процедуру виділення зв'язних компонент нечіткого гіперграфа подій безпеки як кандидатів на кіберінциденти.
- 4. Розроблено метод агрегування інцидентності подій безпеки у межах зв'язної компоненти та критерій прийняття рішення щодо наявності кіберінциденту шляхом переходу від локальної оцінки інцидентності окремих подій до виявлення кіберінцидентів як цілісних структурних утворень.

5. Результати дослідження

Для формалізації подій безпеки та їх структурних взаємозв'язків у SIEM пропонується використати апарат нечітких гіперграфів. Такий підхід дозволяє враховувати багатосутнісну природу подій, а також невизначеність їх інтерпретації. Основна ідея підходу полягає у наступному [1].

Нехай множина подій безпеки задається як: $E = \{e_1, e_2, ..., e_m\}$, а множина сутностей інформаційної системи $V = \{v_1, v_2, ..., v_n\}$, де елементи множини V відповідають сутностям, що беруть участь у подіях безпеки (користувачі, хости, IP-адреси, процеси, техніки атак тощо).

Кожна подія $e \in E$ інтерпретується як гіперребро, що поєднує підмножину сутностей: $V(e) \subseteq V$.

Таким чином, події безпеки формують гіперграф: $H = (V, E)$, у якому вершини відповідають сутностям, а гіперребра – подіям.

На відміну від класичних графових моделей, де зв'язки встановлюються лише між парами вершин, гіперграф дозволяє моделювати взаємодію довільної кількості сутностей у межах однієї події, що є характерним для журналів SIEM.

Для врахування невизначеності інтерпретації подій вводиться функція інцидентності [1]:

$$\mu_E : E \to [0,1], \tag{1}$$

де $\mu_E(e)$ – інтерпретується як вага гіперребра та відображає рівень інцидентності відповідної події (приналежності події безпеки до кіберінциденту).

Запропонована модель дозволяє перейти від ізольованого аналізу подій до дослідження їх структурних взаємозв'язків у рамках єдиного формалізму.

5.1 Нечітка міра інцидентності події безпеки

Нехай $e \in E$ – подія безпеки, що представлена у вигляді гіперребра нечіткого гіперграфа. Для кількісної оцінки ступеня її належності до кіберінциденту вводиться функція інцидентності (1).

Ґрунтуючись на [1], значення $\mu_E(e)$ визначається на основі сукупності характеристик події, що відображають різні аспекти її ризику, а саме: A_e – аномальність події; R_e – критичність правила; C_e – контекстна узгодженість; T_e – семантична значущість; H_e – історична підозрілість.

Зазначені характеристики нормуються до інтервалу [0, 1] та формують вектор ознак:

$$\chi_e = (A_e, R_e, C_e, T_e, H_e) \tag{2}$$

У даній роботі інцидентність події визначається за допомогою лінійної моделі, запропонованої у [1]:

$$\mu_E(e) = \omega_A A_e + \omega_R R_e + \omega_C C_e + \omega_T T_e + \omega_H H_e, \tag{3}$$

де $\omega_A, \omega_A, \omega_C, \omega_T, \omega_H$ – вагові коефіцієнти, що задовольняють умову нормування: $\sum \omega_i = 1$.

Вагові коефіцієнти відображають відносну важливість відповідних характеристик та можуть визначатися: експертним шляхом; на основі статистичного аналізу історичних даних; шляхом оптимізації показників якості виявлення кіберінцидентів.

Отже, кожна складова моделі відповідає окремому виміру оцінювання події: аномальність A_e характеризує відхилення події від типового профілю поведінки; критичність R_e відображає апріорну небезпечність події відповідно до правил SIEM; контекстна узгодженість C_e визначає ступінь включеності події у структурно пов'язану сукупність подій; семантична значущість T_e відображає відповідність події відомим технікам атак; історична підозрілість H_e враховує накопичену активність пов'язаних сутностей [1].

Таким чином, функція $\mu_E(e)$ інтегрує статистичний, експертний, структурний, семантичний та часовий аспекти аналізу подій безпеки.

Не важко помітити, що запропонована функція має такі властивості, як: обмеженість: $\mu_E(e) \in [0,1]$; лінійність: забезпечує простоту обчислення та інтерпретації; монотонність: збільшення будь-якого параметра не зменшує значення інцидентності; адитивність внесків: кожна характеристика робить незалежний вклад у загальну оцінку.

З іншого боку, функція інцидентності $\mu_E(e)$ визначає вагу гіперребра у нечіткому гіперграфі, використовується як складова узагальненої функції зв'язності та є базовим елементом подальшого агрегування інцидентності у межах зв'язних компонент.

Таким чином, введена нечітка міра інцидентності події забезпечує зв'язок між локальними характеристиками подій безпеки та глобальною (у межах нечіткого гіперграфу) структурною оцінкою кіберінцидентів.

5.2 Означення кіберінциденту у нечіткому гіперграфі подій безпеки

Нехай задано нечіткий гіперграф подій безпеки [1]:

$$H = (V, E, \mu_E), \tag{4}$$

де V – множина сутностей, E – множина подій безпеки (гіперребер), $\mu_E : E \to [0,1]$ – функція інцидентності події безпеки.

Нехай на множині подій E задано відношення зв'язності $\sim$, яке визначає структурний зв'язок між подіями.

Означення 1. Підмножина подій $C \subseteq E$ називається зв'язною компонентою гіперграфа H, якщо вона є максимальною за включенням та для будь-яких $e_i, e_j \in C$ існує послідовність подій:

$$\{e_{k_l}\}_{l=0}^{p} \subseteq C, \tag{5}$$

у якій початковим елементом є e_i, а кінцевим – e_j, і для всіх $l = 0,..., p-1$ виконується умова: $e_{k_l} \sim e_{k_{l+1}}$.

Означення 2. Агрегованою інцидентністю подій зв'язної компоненти $C \subseteq E$ будемо називати величину $I(C)$, яка розраховується наступним чином:

$$I(C) = \frac{\sum_{e \in C} \mu_E(e)\omega(e)}{\sum_{e \in C} \omega(e)}, \tag{6}$$

де $\omega(e) > 0$ – вагова функція події безпеки.

Означення 3. Кіберінцидентом називається зв'язна компонента $C \subseteq E$ нечіткого гіперграфа $H = (V, E, \mu_E)$, для якої агрегована інцидентність задовольняє умову:

$$I(C) \geq I_0, \tag{7}$$

де $I_0 \in [0,1]$ – порогове значення, що визначає межу прийняття рішення щодо належності компоненти гіперграфа до кіберінцидентів.

Виходячи з цього, суть запропонованого методу полягає у формалізації процесу виявлення кіберінцидентів як задачі виділення зв'язних підмножин подій у нечіткому гіперграфі, побудованому на основі потоків подій безпеки SIEM, із подальшим оцінюванням їх інцидентності за допомогою лінійної моделі (3) та порогового критерію I_0.

Нехай задано множину подій безпеки $E = \{e_1, e_2, ..., e_m\}$ та множину сутностей $V = \{v_1, v_2, ..., v_n\}$. Необхідно знайти відображення:

$$\Phi : (E, V) \to INC = \{C \subseteq E : C - зв'язна\ компонента, I(C) \geq I_0\}, \quad (8)$$

де INC – множина виявлених кіберінцидентів.

Основними етапами методу є побудова нечіткого гіперграфу подій безпеки, виділення та аналіз його зв'язних компонент та прийняття рішення про наявність кіберінциденту. Опишемо ці етапи більш детально (див. рис. 1).

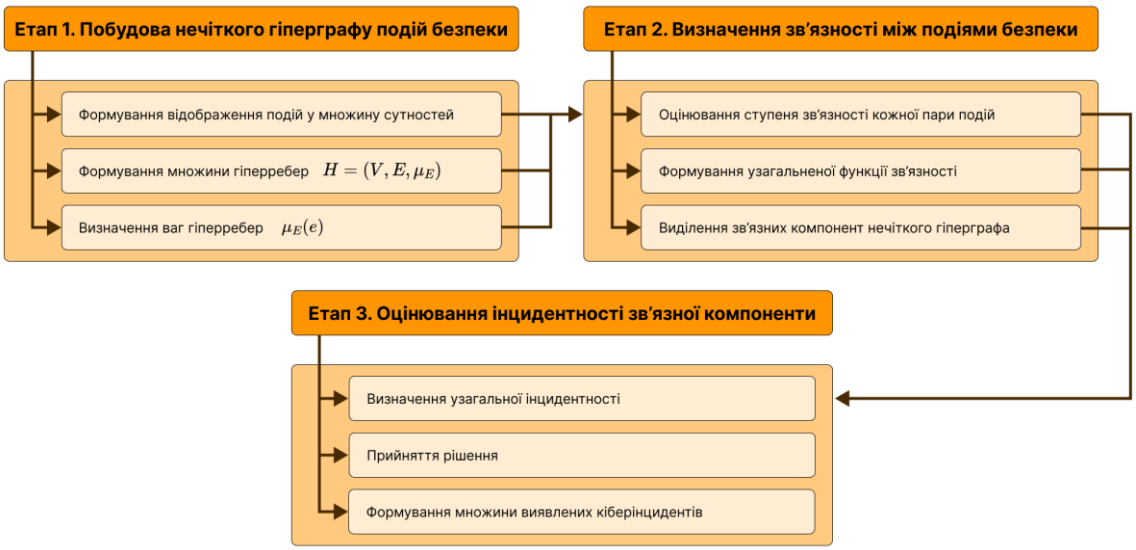

Рис. 1. Схема методу виявлення кіберінцидентів у SIEM на основі нечіткого гіперграфового подання подій безпеки та лінійної моделі інцидентності

Етап 1. Побудова нечіткого гіперграфу подій безпеки. Здійснюється шляхом відображення кожної події у підмножину сутностей інформаційної інфраструктури. На відміну від класичних задач аналізу графів, де структура задана наперед, у запропонованому підході гіперграф формується безпосередньо з подій безпеки, що є частиною методу виявлення кіберінцидентів.

Спочатку здійснюється формування відображення подій у множину сутностей:

$$\psi : E \to 2^V, \quad (9)$$

де $\psi(e) = V(e) \in V$.

Тобто кожній події ставиться у відповідність множина сутностей, що беруть участь у ній.

Тоді множина вершин гіперграфа формується як об'єднання всіх сутностей:

$$V = \bigcup_{e \in E} V(e), \quad (10)$$

Далі здійснюється формування множини гіперребер. Множина гіперребер визначається як:

$$E_H = \{V(e) \mid e \in E\}, E_H \subseteq 2^V, \quad (11)$$

Кожне гіперребро $e \equiv V(e)$ відповідає одній події безпеки.

Після цього визначаються ваги гіперребер – здійснюється обчислення локальної інцидентності події безпеки на основі (2), (3). Тобто кожному гіперребру присвоюється вага: $w_{H(e)} = \mu_E(e)$ та формується нечіткий гіперграф (4).

Етап 2. Визначення зв'язності між подіями безпеки.

Для кожної пари подій $e_i, e_j \in E$ оцінюється ступінь їх зв'язності на основі сукупності ознак, що характеризують: спільність сутностей, залучених до подій; часову близькість подій; узгодженість їх інцидентності.

На основі зазначених ознак формується узагальнена функція зв'язності $R(e_i, e_j) \in [0,1]$, яка відображає ступінь взаємозалежності подій та яка може бути реалізована як монотонна функція агрегування відповідних ознак:

$$R(e_i, e_j) = \alpha R_V(e_i, e_j) + \beta R_T(e_i, e_j) + \gamma R_\mu(e_i, e_j), \tag{12}$$

де $\alpha, \beta, \gamma \in [0,1]$ – вагові коефіцієнти, причому $\alpha + \beta + \gamma = 1$. Вони визначають внесок кожної складової: α – важливість спільних сутностей; β – важливість часової близькості; γ – важливість узгодженості інцидентності. Вагові коефіцієнти можуть визначатися статистично, експертно або шляхом рішення оптимізаційної задачі.

Структурна складова визначається на основі коефіцієнта Жаккара [11]:

$$R_V(e_i, e_j) = \frac{\left|V(e_i) \bigcap V(e_j)\right|}{\left|V(e_i) \bigcup V(e_j)\right|}, \tag{13}$$

де $V(e)$ – множини сутностей (user, host, IP, process тощо).

Якщо $R_V \to 0$ події не пов'язані структурно, та навпаки – якщо $R_V \to 1$ події мають однаковий набір сутностей.

Часова складова задається експоненційною функцією:

$$R_T(e_i, e_j) = \exp\left(-\frac{\left|t(e_i) - t\left(e_j\right)\right|}{\tau} \right), \tag{14}$$

де $\tau > 0$ – параметр часової чутливості, $t(e)$ – час події.

Якщо події близькі у часі, то значення R_T близьке до 1, та навпаки – якщо події далекі у часі, то значення R_T близьке до 0.

Складова узгодженості інцидентності визначається як:

$$R_\mu(e_i, e_j) = 1 - \left|\mu_E(e_i) - \mu_E(e_j)\right|, \tag{15}$$

де $\mu_E(e) \in [0,1]$ – інцидентність події безпеки.

Якщо події мають однакову інцидентність, то $R_\mu = 1$, та навпаки – якщо події мають інцидентність, яка сильно відрізняється, то $R_\mu = 0$. Це, з одного боку дозволяє зв'язувати однорідні за ризиком події і напроти, не відсікає низькоінцидентні події (вони можуть бути зв'язками).

Таким чином, функція $R(e_i, e_j)$ інтегрує структурні, часові та інцидентні характеристики подій і використовується для визначення їх зв'язності.

Дві події e_i, e_j вважаються зв'язаними $e_i \sim e_j$, якщо:

$$e_i \sim e_j \Leftrightarrow R(e_i, e_j) \geq \theta, \tag{16}$$

де θ – поріг зв'язності.

Отримане відношення $\sim$ використовується для подальшого аналізу структури подій безпеки.

Формально відношення зв'язності визначається як бінарне відношення на множині подій:

$$\sim \subseteq E \times E, \tag{17}$$

яке задається на основі узагальненої міри зв'язності подій (12).

З урахуванням введеної узагальненої функції зв'язності подій $R(e_i, e_j)$, доцільно визначити вагу події з виразу (6) на основі її структурної значущості у межах зв'язної компоненти C. Нехай $C \subseteq E$ – зв'язна компонента нечіткого гіперграфа. Тоді для кожної події $e_i \in C$ її вагу визначимо як:

$$\omega_C(e_i) = 1 + \sum_{\substack{e_j \in C, \\ j \neq i, \\ R(e_i, e_j) \geq \theta}} R(e_i, e_j), \tag{18}$$

де $\omega_C(e_i)$ – структурна вага події, що визначається в межах зв'язної компоненти, $R(e_i, e_j)$ – узагальнена функція зв'язності подій, θ – порогове значення зв'язності.

Величина $\omega_C(e_i)$ характеризує структурну значущість події у межах компоненти C: чим із більшою кількістю подій компоненти, та чим сильніше пов'язана подія e_i, тим більшим є її внесок у формування агрегованої інцидентності компоненти.

З урахуванням вагової функції агрегована інцидентність зв'язної компоненти C визначається як:

$$I(C) = \frac{\sum_{e \in C} \mu_E(e) \omega_C(e)}{\sum_{e \in C} \omega_C(e)}, \tag{19}$$

Такий підхід дозволяє враховувати не лише локальний рівень інцидентності окремих подій, а й їх структурну роль у формуванні кіберінциденту.

На наступному кроці виділяються зв'язні компоненти нечіткого гіперграфу подій безпеки. Для цього на основі відношення зв'язності (17) здійснюється групування подій у підмножини, в межах яких будь-які дві події пов'язані між собою безпосередньо або через ланцюг проміжних подій.

Кожна така підмножина утворює зв'язну компоненту гіперграфа:

$$\zeta = \{C_1, C_2, \ldots, C_k\}, \tag{20}$$

Отримані компоненти розглядаються як структурно цілісні групи взаємопов'язаних подій.

Етап 3. Оцінювання інцидентності зв'язної компоненти. Для кожної зв'язної компоненти $C_k \in \zeta$ визначається узагальнена інцидентність $I(C_k)$ на основі агрегування індивідуальних інцидентностей подій, що входять до компоненти, з урахуванням їх взаємозв'язків та узгодженості, шляхом застосування виразу (19). Саме значення агрегованої інцидентності характеризує ступінь відповідності даної компоненти ознакам кіберінциденту.

Зв'язна компонента C_k інтерпретується як кіберінцидент, якщо її інцидентність задовольняє умову (7).

У результаті формується множина виявлених кіберінцидентів INC.

Компоненти, для яких умова не виконується, розглядаються як фонові або потенційно незначущі подієві структури.

Результатом застосування методу є множина виявлених кіберінцидентів, кожен з яких представлений у вигляді зв'язної компоненти нечіткого гіперграфа подій безпеки з відповідним значенням інцидентності.

Порогове значення I_0 може визначатися: експертно – на основі практики SOC; статистично – шляхом аналізу розподілу $I(C)$; оптимізаційно – шляхом вирішення оптимізаційної задачі:

$$I_0 = \arg\max_{t\in[0,1]} F_1(t), \tag{21}$$

де $F_1(t)$ – значення F_1 – міри для порогу t.

Розглянемо ілюстративний приклад застосування розробленого методу для виявлення кіберінциденту у SIEM на основі трьох взаємопов'язаних подій журналів Wazuh, що відповідають типовому сценарію кібератаки типу brute force з подальшою компрометацією доступу (див. рис. 2).

Нехай у журналі подій зафіксовано три події безпеки. Перша подія e_1 відповідає серії невдалих спроб автентифікації за протоколом SSH з однієї IP-адреси:

timestamp = 2026-03-01 10:00:01
rule.id = 5710
description = 3 failed SSH logins
srcip = 192.168.1.10
dstip = 192.168.1.50
user = root
host = server-1

Друга подія e_2 відповідає успішній автентифікації за тим самим обліковим записом з тієї самої IP-адреси:

timestamp = 2026-03-01 10:00:15
rule.id = 5715
description = SSH login success
srcip = 192.168.1.10
dstip = 192.168.1.50
user = root
host = server-1

Третя подія e_3 фіксує зміну контрольної суми критичного файлу системи, що здійснюється в рамках відкритої сесії зловмисника, вказуючи на можливе закріплення в системі:

timestamp = 2026-03-01 10:00:20
rule.id = 550
description = Integrity checksum changed
user = root
host = server-1
file = /etc/shadow

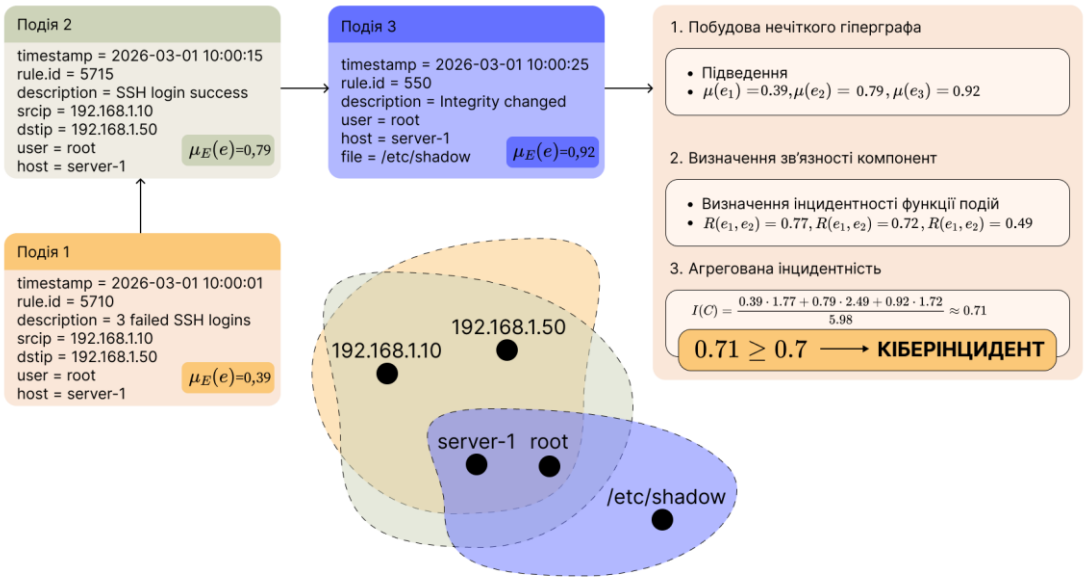

Рис. 2. Приклад виявлення кіберінциденту на основі трьох зв'язаних подій журналів SIEM Wazuh, що відповідають сценарію brute force з подальшою компрометацією доступу

Отже маємо три події: e_1 (серія невдалих спроб входу), e_2 (успішний вхід), e_3 (модифікація критичного файлу). Загальна множина сутностей: $V = \{192.168.1.10,\ 192.168.1.50,\ \text{root},\ \text{server-1},\ /\text{etc/shadow}\}$.

На відміну від ізольованого розгляду окремих подій, у запропонованому методі аналізується їх структурний зв'язок, що дає змогу інтерпретувати таку сукупність як можливий кіберінцидент.

Етап 1. Побудова нечіткого гіперграфа подій безпеки. На початку кожна подія відображається у множину сутностей: $V(e_1)=\{192.168.1.10, 192.168.1.50, root, server\_1\}$, $V(e_2)=\{192.168.1.10, 192.168.1.50, root, server\_1\}$, $V(e_3)=\{root, server\_1, /etc/shadow\}$. Тоді множина вершин гіперграфа визначається як: $V=V(e_1)\cup V(e_2)\cup V(e_3)=\{192.168.1.10, 192.168.1.50, root, server\_1, /etc/shadow\}$, а множина гіперребер: $E_H=\{e_1,e_2,e_3\}$.

На основі лінійної моделі інцидентності для цих подій та початкових даних (див. табл. 1) маємо:

Таблиця 1

Значення параметрів для подій

Параметр	Зміст	Вага	Значення		
			e_1	e_2	e_3
A_e	аномальність	0.10	0.30	0.80	0.90
R_e	критичність правила	0.20	0.30	0.70	0.90
C_e	контекст	0.30	0.30	0.80	0.90
T_e	семантика (MITRE)	0.30	0.60	0.90	1.00
H_e	історія	0.10	0.30	0.60	0.80

$$\mu_E(e_1)=0.1\cdot0.3+0.2\cdot0.3+0.3\cdot0.3+0.3\cdot0.6+0.1\cdot0.3=0.39,$$
$$\mu_E(e_2)=0.1\cdot0.8+0.2\cdot0.7+0.3\cdot0.8+0.3\cdot0.9+0.1\cdot0.6=0.79,$$
$$\mu_E(e_3)=0.1\cdot0.9+0.2\cdot0.9+0.3\cdot0.9+0.3\cdot1.0+0.1\cdot0.8=0.92.$$

Таким чином, нечіткий гіперграф $H=(V,E_H,\mu_E)$ сформовано.

Етап 2. Визначення зв'язності між подіями. Для оцінювання зв'язності подій e_1 та e_2, e_2 та e_3, e_1 та e_3 використовується узагальнена функція (12):

$$R(e_1,e_2)=\alpha R_V(e_1,e_2)+\beta R_T(e_1,e_2)+\gamma R_\mu(e_1,e_2),$$
$$R(e_2,e_3)=\alpha R_V(e_2,e_3)+\beta R_T(e_2,e_3)+\gamma R_\mu(e_2,e_3),$$
$$R(e_1,e_3)=\alpha R_V(e_1,e_3)+\beta R_T(e_1,e_3)+\gamma R_\mu(e_1,e_3),$$

де для прикладу приймемо $\alpha=\beta=\gamma=1/3$.

Структурна складова, визначена за коефіцієнтом Жаккара, дорівнює:

$$R_V(e_1,e_2)=\frac{|V(e_1)\bigcap V(e_2)|}{|V(e_1)\bigcup V(e_2)|}=\frac{4}{4}=1.0,$$
$$R_V(e_2,e_3)=\frac{|V(e_2)\bigcap V(e_3)|}{|V(e_2)\bigcup V(e_3)|}=\frac{2}{5}=0.4,$$
$$R_V(e_1,e_3)=\frac{|V(e_1)\bigcap V(e_3)|}{|V(e_1)\bigcup V(e_3)|}=\frac{2}{5}=0.4,$$

Оскільки різниця між часовими мітками подій становить $\Delta t_{1,2}=14\,\text{sec}$, $\Delta t_{2,3}=5\,\text{sec}$, $\Delta t_{1,3}=19\,\text{sec}$, а параметр часової чутливості прийнято $\tau=40$, маємо:

$$R_T(e_1,e_2)=\exp(-\frac{14}{40})\approx 0.71$$
$$R_T(e_2,e_3)=\exp(-\frac{5}{40})\approx 0.88$$
$$R_T(e_1,e_3)=\exp(-\frac{19}{40})\approx 0.62$$

Складова узгодженості інцидентності дорівнює:

$$R_\mu(e_1,e_2)=1-|\mu(e_1)-\mu(e_2)|=1-|0.39-0.79|=0.60$$
$$R_\mu(e_2,e_3)=1-|\mu(e_2)-\mu(e_3)|=1-|0.79-0.92|=0.87$$
$$R_\mu(e_1,e_3)=1-|\mu(e_1)-\mu(e_3)|=1-|0.39-0.92|=0.47$$

Тоді узагальнена функція зв'язності становить:

$$R(e_1,e_2)=0.33\cdot(1.0+0.71+0.60)\approx 0.77$$
$$R(e_2,e_3)=0.33\cdot(0.40+0.88+0.87)\approx 0.72$$
$$R(e_1,e_3)=0.33\cdot(0.4+0.62+0.47)\approx 0.49$$

Нехай поріг зв'язності $\theta=0.7$. Безпосередній зв'язок між першою та третьою подією є слабким $R(e_1,e_3)=0.49<0.7$. Проте, відповідно до Означення 1, послідовність подій формує зв'язну компоненту, оскільки існує ланцюг $e_1\to e_2\to e_3$, де $R(e_1,e_2)=0.77\geq 0.7$ та $R(e_2,e_3)=0.72\geq 0.7$. Таким чином, події утворюють єдину зв'язну компоненту $C=\{e_1,e_2,e_3\}$.

Етап 3. Оцінювання інцидентності зв'язної компоненти. Для кожної події в межах компоненти визначається структурна вага (18). Для першої події (e_1) враховується лише зв'язок з e_2, оскільки $R(e_1, e_3) = 0.49 < 0.7$:

$$\omega_C(e_1) = 1 + R(e_1, e_2) = 1 + 0.77 = 1.77$$

Для другої події (e_2) враховуються обидва зв'язки, оскільки вони перевищують поріг:

$$\omega_C(e_2) = 1 + R(e_2, e_1) + R(e_2, e_3) = 1 + 0.77 + 0.72 = 2.49$$

Для третьої події (e_3) враховується лише прямий зв'язок з e_2:

$$\omega_C(e_3) = 1 + R(e_3, e_2) = 1 + 0.72 = 1.72$$

Сума структурних ваг усіх подій компоненти становить $\sum \omega_C = 1.77 + 2.49 + 1.72 = 5.98$.

Агрегована інцидентність компоненти визначається як зважена сума локальних інцидентностей подій:

$$I(C) = \frac{\mu_E(e_1)\omega_C(e_1) + \mu_E(e_2)\omega_C(e_2) + \mu_E(e_3)\omega_C(e_3)}{\sum \omega_C} = \frac{0.39 \cdot 1.77 + 0.79 \cdot 2.49 + 0.92 \cdot 1.72}{5.98} \approx 0.71$$

При умові, що порогове значення для прийняття рішення: $I_0 = 0.7$, маємо: $I(C) = 0.71 \geq 0.7$. Відповідно до цього, приймаємо рішення, що зв'язна компонента $C = \{e_1, e_2, e_3\}$ інтерпретується як кіберінцидент.

Розглянутий приклад показує, що жодна з наведених подій окремо не є достатньою підставою для однозначного висновку про наявність кіберінциденту. Проте їх сукупний аналіз з урахуванням спільних сутностей, часової близькості та високої узгодженості інцидентності дозволяє виявити структурно пов'язаний ланцюг подій, характерний для атаки типу brute force з подальшим успішним входом до системи.

У традиційних SIEM події e_1 та e_2 ймовірно були б проігноровані як фоновий шум або розглядалися б як ізольовані неуспішні/успішні спроби входу. Однак застосування запропонованого методу дозволило об'єднати їх із критичною подією e_3 в єдину структуру завдяки наявності проміжних зв'язків. При встановленні порогового значення для прийняття рішення на рівні $I_0 = 0.7$, агрегована інцидентність зв'язної компоненти $I(C) = 0.71$ перевищує заданий поріг. Відповідно, система приймає обґрунтоване рішення щодо класифікації даної сукупності подій як цілісного кіберінциденту.

Таким чином, запропонований метод забезпечує виявлення кіберінциденту не на рівні окремої події, а на рівні зв'язної компоненти нечіткого гіперграфа подій безпеки.

6. Висновки та перспективи подальших досліджень

У роботі вперше розроблено метод виявлення кіберінцидентів у SIEM, який, на відміну від існуючих, базується на нечіткому гіперграфовому поданні подій безпеки та забезпечує виявлення кіберінцидентів як зв'язних компонент множини подій, сформованих за відношенням зв'язності, заданим на основі узагальненої функції зв'язності.

Метод відрізняється використанням узагальненої функції зв'язності подій безпеки, що визначається як зважена комбінація структурної подібності, часової близькості та узгодженості інцидентності, а також формалізацією критерію прийняття рішення у вигляді порогової умови на агреговану інцидентність зв'язної компоненти, обчислену на основі лінійної моделі інцидентності події.

Практичне значення одержаного результату полягає у можливості підвищення ефективності виявлення кіберінцидентів у SIEM-системах за рахунок переходу від аналізу окремих подій безпеки до виявлення їх структурно пов'язаних сукупностей, що відповідають розвитку кіберзагроз. Запропонований метод дозволяє враховувати багатосутнісну природу подій, часову динаміку їх виникнення та невизначеність інтерпретації, що забезпечує більш обґрунтоване прийняття рішень у процесах виявлення кіберінцидентів.

Результати роботи можуть бути використані при розробленні та вдосконаленні SIEM-рішень, систем виявлення вторгнень, а також аналітичних модулів центрів моніторингу безпеки (SOC) для автоматизованого виявлення складних багатокрокових атак.

Перспективами подальших досліджень є розроблення алгоритмів побудови нечіткого гіперграфа та виділення й аналізу в ньому зв'язних компонент.

Внесок авторів Ігор Субач – концептуалізація дослідження, постановка проблеми, формалізація постановки задачі, участь у формуванні висновків дослідження; Данило Копич – збір і аналіз джерел, підготовка огляду літератури, формалізація та розроблення методу виявлення кіберінцидентів.

Декларація про штучний інтелект

Інструменти штучного інтелекту використовувалися для мовно-стилістичного редагування тексту та не впливали на науковий зміст, результати та висновки дослідження.

Конфлікт інтересів

Автори заявляють про відсутність конфлікту інтересів та підтверджують, що під час підготовки цієї роботи не існувало жодних комерційних, фінансових чи інших взаємовідносин, які могли б бути розцінені як такі, що здатні вплинути на результати дослідження або їх інтерпретацію. Робота виконана відповідно до принципів академічної доброчесності, етичних норм проведення наукових досліджень та вимог редакційної політики щодо запобігання конфлікту інтересів.

Список використаної літератури

1. Субач, І., & Копич, Д. (2026). Модель виявлення кіберінцидентів на основі нечіткого гіперграфу подій безпеки SIEM. *Collection "Information Technology and Security"*, *14*(1). – прийнята до друку.
2. Akoglu, L., Tong, H., & Koutra, D. (2014). Graph-based anomaly detection and description: A survey. *Data Mining and Knowledge Discovery*, *29*(3), 626–688. https://doi.org/10.1007/s10618-014-0365-y
3. Borgohain, R. (2012). FuGeIDS: Fuzzy genetic paradigms in intrusion detection systems. *arXiv*. https://arxiv.org/abs/1204.6416
4. Субач, І., & Фесьоха, В. (2017). Модель виявлення кібернетичних атак на інформаційно-телекомунікаційні системи на основі описання аномалій їх роботи зваженими нечіткими правилами. *Collection "Information Technology and Security"*, *5*(2), 145–152. https://doi.org/10.20535/2411-1031.2017.5.2.136984
5. Subach, I., & Fesokha, V. (2017). Model of detection of anomalies in information and telecommunication networks of military management bodies on the basis of fuzzy sets and fuzzy logic output. *Collection of scientific works of VITI, (3),* 158-164.
6. Kay, B., Aksoy, S. G., Baird, M., Best, D. M., Jenne, H., Joslyn, C., ... & Purvine, E. (2023). Hypergraph topological features for autoencoder-based intrusion detection for cybersecurity data. *arXiv preprint* arXiv:2312.00023. https://doi.org/10.48550/arXiv.2312.00023
7. Raman, M. R. G., Somu, N., Kirthivasan, K., Sriram, V. S., & Liscano, R. (2017). An efficient intrusion detection system based on hypergraph – genetic algorithm for parameter optimization and feature selection in support vector machine. *Expert Systems with Applications*, *83*, 211–225. https://doi.org/10.1016/j.eswa.2017.04.019
8. Scarfone, K., & Mell, P. (2007). Guide to intrusion detection and prevention systems (IDPS). *NIST Special Publication*, *800*(2007), 94. https://doi.org/10.6028/NIST.SP.800-94
9. Shyu, M. L., Chen, S. C., Sarinnapakorn, K., & Chang, L. (2003). A novel anomaly detection scheme based on principal component classifier. In *Proceedings of the IEEE Foundations and New Directions of Data Mining Workshop* (pp. 172–179).
10. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, *41*(3), 1–58. https://doi.org/10.1145/1541880.1541882

11. Maosa, H., Ouazzane, K., & Ghanem, M. C. (2023). A hierarchical security events correlation model for real-time cyber threat detection and response. *arXiv preprint* arXiv:2312.01219. https://doi.org/10.3390/network4010004

12. MITRE ATT&CK. (2026). *MITRE ATT&CK framework*. https://attack.mitre.org

13. Альперт, С. І. (2019). Основні міри подібності та нові підходи до їх застосування при класифікуванні гіперспектральних космічних зображень. *Математичні машини і системи*, (1), 143-151.

14. Behl, A., Behl, K., & Behl, K. (2017). Cybersecurity and cyberwar: What everyone needs to know. *Oxford University Press.* https://doi.org/10.1093/wentk/9780199918096.001.0001

References

1. Subach, I., & Kopych, D. (2026). Cyber incident detection model based on fuzzy hypergraph of SIEM security events. *Collection "Information Technology and Security"*, *14*(1). – In press.

2. Akoglu, L., Tong, H., & Koutra, D. (2014). Graph-based anomaly detection and description: A survey. *Data Mining and Knowledge Discovery*, *29*(3), 626–688. https://doi.org/10.1007/s10618-014-0365-y

3. Borgohain, R. (2012). FuGeIDS: Fuzzy genetic paradigms in intrusion detection systems. *arXiv*. https://arxiv.org/abs/1204.6416

4. Subach, I., & Fesokha, V. (2017). A model for detecting cyber attacks on information and telecommunication systems based on describing anomalies in their operation with weighted fuzzy rules. *Collection "Information Technology and Security"*, *5*(2), 145–152. https://doi.org/10.20535/2411-1031.2017.5.2.136984

5. Subach, I., & Fesokha, V. (2017). Model of detection of anomalies in information and telecommunication networks of military management bodies on the basis of fuzzy sets and fuzzy logic output. *Collection of scientific works of VITI*, *(3),* 158-164.

6. Kay, B., Aksoy, S. G., Baird, M., Best, D. M., Jenne, H., Joslyn, C., ... & Purvine, E. (2023). Hypergraph topological features for autoencoder-based intrusion detection for cybersecurity data. *arXiv*. https://doi.org/10.48550/arXiv.2312.00023

7. Raman, M. R. G., Somu, N., Kirthivasan, K., Sriram, V. S., & Liscano, R. (2017). An efficient intrusion detection system based on hypergraph – genetic algorithm for parameter optimization and feature selection in support vector machine. *Expert Systems with Applications*, *83*, 211–225. https://doi.org/10.1016/j.eswa.2017.04.019

8. Scarfone, K., & Mell, P. (2007). Guide to intrusion detection and prevention systems (IDPS). *NIST Special Publication*, *800*(2007), 94. https://doi.org/10.6028/NIST.SP.800-94

9. Shyu, M. L., Chen, S. C., Sarinnapakorn, K., & Chang, L. (2003). A novel anomaly detection scheme based on principal component classifier. In *Proceedings of the IEEE Foundations and New Directions of Data Mining Workshop* (pp. 172–179).

10. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, *41*(3), 1–58. https://doi.org/10.1145/1541880.1541882

11. Maosa, H., Ouazzane, K., & Ghanem, M. C. (2023). A hierarchical security events correlation model for real-time cyber threat detection and response. *arXiv*. https://doi.org/10.3390/network4010004

12. MITRE ATT&CK. (2026). *MITRE ATT&CK framework*. https://attack.mitre.org

13. Alpert, S. I. (2019). Basic similarity measures and new approaches to their application in the classification of hyperspectral space images. *Mathematical Machines and Systems*, (1), 143–151.

14. Behl, A., Behl, K., & Behl, K. (2017). Cybersecurity and cyberwar: What everyone needs to know. *Oxford University Press.* https://doi.org/10.1093/wentk/9780199918096.001.0001

Надійшла до редакції: 06.04.26
Прийнята до друку:12.06.26
Опубліковано:30.06.26