TECHNICAL SCIENCES AND TECHNOLOGIES

DOI: https://doi.org/10.25140/2411-5363-2026-2(44)-248-259
УДК 004.891:004.056.5:656.07

Алла Григорівна Гребенник[1], Олексій Ігорович Трунов [2]

[1]старший викладач кафедри кібербезпеки та математичного моделювання
Національний університет «Чернігівська політехніка» (Чернігів, Україна)
E–mail: grebennik.alla@gmail.com. **ORCID**: https://orcid.org/0000-0002-7464-1412
Scopus Author ID: 57219054928

[2]аспірант, викладач кафедри інформаційних технологій та програмної інженерії
Національний університет «Чернігівська політехніка» (Чернігів, Україна)
E-mail: alexeytrunov1995@gmail.com. **ORCID**: https://orcid.org/0009-0002-0321-2669
Scopus Author ID: 59337119500

АЛГОРИТМ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ МЕРЕЖЕВОГО КОНТЕНТУ ДЛЯ ЗАБЕЗПЕЧЕННЯ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ ТРАНСПОРТНО-ЛОГІСТИЧНОГО ЦЕНТРУ

Досліджено проблему захисту транспортно-логістичних центрів від гібридних загроз, де традиційний моніторинг ключових слів є неефективним проти соціальної інженерії. Розроблено метод інтелектуального аналізу мережевого контенту, що поєднує кількісно-якісний контент-аналіз та якісний конверсаційний аналіз. Запропоновано математичний апарат на основі матриць порядку M_j, методу головних компонент (PCA) та нейро-нечіткої мережі для розрахунку інтегрального показника ризику R. Моделювання сценаріїв підтвердило високу чутливість методу до структурних аномалій діалогу навіть за умови використання персоналом легітимної лексики. Результати дозволяють автоматизувати виявлення прихованих загроз людського фактору та забезпечити проактивне реагування в системах підтримки ухвалення рішень.

***Ключові слова**: інформаційна безпека; транспортно-логістичний центр; інтелектуальний аналіз контенту; конверсаційний аналіз; матриця дистанцій; нейронечітка мережа; людський фактор; соціальна інженерія; система підтримки ухвалення рішень.*

Рис.: 1. Табл.: 3. Бібл.: 20.

Актуальність теми дослідження. В умовах геополітичної нестабільності цілісність ланцюгів постачання є фундаментом національної безпеки, а їхніми ключовими вузлами виступають транспортно-логістичні центри (ТЛЦ). Як багатофункціональні хаби, ТЛЦ інтегрують технологічні, транспортні та інформаційні потоки, забезпечуючи життєдіяльність держави.

Сучасні гібридні атаки на логістичний сектор спрямовані на стратегічне ушкодження цих критичних потоків. Концентрація великих масивів чутливих даних робить ТЛЦ пріоритетною мішенню, де традиційні технічні засоби захисту часто виявляються неефективними проти інсайдерських загроз, соціальної інженерії та маніпуляцій свідомістю персоналу.

Особливої гостроти проблема набуває в умовах повномасштабної війни, коли ТЛЦ стають ключовими вузлами розподілу військової та гуманітарної допомоги. Ворожі диверсії, дезінформаційні кампанії та психологічний тиск на персонал вимагають впровадження механізмів інтелектуального аналізу мережевого контенту. Це є життєво необхідною умовою для виявлення прихованих загроз у комунікаціях та забезпечення стійкості транспортно-логістичної галузі.

Постановка проблеми. Ефективне забезпечення інформаційної безпеки ТЛЦ на сучасному етапі суттєво обмежується невідповідністю традиційних засобів моніторингу новітнім загрозам, які реалізуються через текстові комунікації з високим рівнем семантичної невизначеності. Одним із ключових аспектів цієї проблеми є відсутність методології для агрегації різнорідних даних, що надходять від розрізнених засобів захисту (SIEM, IDS, антивірусне ПЗ), у єдиний інтегральний показник захищеності. Це призводить до розмиття загальної картини безпеки та, як наслідок, запізнілого виявлення інцидентів.

Водночас критично низька ефективність існуючих методів аналізу тексту, що базуються виключно на пошуку за ключовими словами без урахування структурної послідовності та динаміки діалогової взаємодії, унеможливлює автоматизоване спрацювання

тригерів у системах підтримки прийняття рішень систем захисту. Такий спрощений підхід не дозволяє ідентифікувати складні сценарії соціальної інженерії або інсайдерської діяльності, де загроза прихована у контексті, а не в конкретних термінах.

З огляду на це, актуальним науковим завданням є розробка алгоритму інтелектуального аналізу мережевого контенту, який би базувався на синергетичному поєднанні кількісно-якісного контент-аналізу та якісного конверсаційного аналізу. Таке поєднання дозволить не лише враховувати частотні характеристики загроз, а й аналізувати структурно-семіотичні параметри комунікацій для виявлення прихованих закономірностей та аномалій у системі ІБ ТЛЦ.

Аналіз останніх досліджень і публікацій. Теоретичне підґрунтя для моделювання та аналізу безпеки складних об'єктів закладено у фундаментальних працях В. Литвинова та ін. [1], де систематизовано методи захисту розподілених інформаційних систем. Подальший розвиток ці ідеї знайшли у дослідженні захисту корпоративних мереж від атак з використанням контент-аналізу інформаційних потоків [2], що заклало наукову базу для переходу до систем динамічного оцінювання ризиків [3] та побудови архітектур нульової довіри [4]. Такі підходи дозволяють будувати гнучкі та адаптивні системи захисту, здатні ефективно протидіяти динамічним змінам ландшафту загроз.

Питання специфіки захисту саме ТЛЦ як критичної ланки ланцюгів постачання детально висвітлено у роботі [5], де проведено систематизацію підходів до оцінки ризиків ІБ. Важливим кроком на шляху до автоматизації моніторингу контенту став розроблений алгоритм визначення агрегованої динамічної оцінки стану безпеки мережевого контенту [6]. Запропонований підхід корелює з концепціями інтенційно-орієнтованих мереж [7] та сучасними методами детекції аномалій у гетерогенних системах [8]. Водночас, як зазначено в роботі [9], ефективність функціонування таких систем у складних IoT-середовищах ТЛЦ критично залежить від семантичної сумісності даних.

Сучасні виклики 2024–2026 років, зокрема шифрування TLS 1.3 [10] та використання ШІ-агентів для реалізації складних сценаріїв соціальної інженерії [11, 12], вимагають еволюції існуючих алгоритмів моніторингу. Попри наявність ґрунтовних напрацювань щодо агрегованої оцінки контенту, залишається нагальна потреба у поглибленні семантичного аналізу через формалізацію структурно-семіотичних характеристик діалогів.

Виділення недосліджених частин загальної проблеми. Незважаючи на значний прогрес у розробці інтелектуальних систем захисту, залишається невирішеною проблема цілісної оцінки стану безпеки мережевого контенту транспортно-логістичних центрів. Сучасні засоби моніторингу генерують значну кількість різнорідних показників, проте наразі відсутній формалізований підхід для їх агрегації з результатами семантичного аналізу комунікацій персоналу в єдиний індикатор. Крім того, адаптивний характер загроз, пов'язаних із соціальною інженерією, вимагає впровадження механізмів динамічного аналізу контексту, здатних виявляти аномалії в режимі реального часу на основі структурно-семіотичних властивостей діалогів, а не лише статистичних параметрів ключових слів.

Метою статті є розробка методу інтелектуального аналізу мережевого контенту ТЛЦ на основі поєднання методів контент-аналізу, конверсаційного аналізу та алгоритмів векторизації з використанням матриць дистанцій. Запропоноване рішення спрямоване на автоматизацію виявлення прихованих загроз людського фактора та формування оцінки стану інформаційної безпеки, що слугуватиме обґрунтованою основою для активації механізмів адаптивного реагування в системах підтримки прийняття рішень.

Об'єктом дослідження обрано процеси забезпечення інформаційної безпеки ТЛЦ в умовах активного використання неструктурованих мережевих комунікацій. Предметом виступають методи та моделі аналізу контенту на основі структурно-семіотичних харак-

теристик взаємодії суб'єктів. Для досягнення поставленої мети у роботі послідовно реалізовано низку наукових завдань, що включають класифікацію джерел контенту та специфічних загроз ІБ, пов’язаних із людським фактором, обґрунтування інтеграції контент- та конверсаційного аналізу для виявлення контекстних аномалій, а також розробку методу векторизації неструктурованого контенту на основі матриць дистанцій для формалізації динаміки діалогу.

Виклад основного матеріалу. В умовах цифровізації логістичних процесів забезпечення цілісності ланцюгів постачання залежить не лише від стійкості програмно-технічних засобів, а й від надійності людського фактору. На відміну від технічних вразливостей, які піддаються патчингу, людський фактор залишається найбільш непередбачуваним джерелом ризиків. Недбалість, помилки персоналу, а також цілеспрямована діяльність інсайдерів та атаки із застосуванням соціальної інженерії створюють унікальні виклики для ІБ ТЛЦ.

Для проактивного виявлення загроз, пов'язаних із людським фактором, критично важливим є моніторинг та інтелектуальний аналіз мережевого контенту. У контексті ТЛЦ цей контент класифікується за трьома основними напрямами:

1) внутрішній тип контенту, який охоплює Email, корпоративні месенджери (Teams, Slack), записи VoIP та внутрішні звіти. Ключовими загрозами є інсайдерські витоки, змови, саботаж, а також передача паролів;

2) зовнішній контент включає соціальні мережі, професійні форуми та моніторинг Dark Web. Основні ризики тут пов'язані з підготовкою до фішингу та збором розвідувальних даних про вантажі;

3) технічний контент містить логи систем, SIEM-події та журнали доступу, які аналізуються для виявлення аномальної активності (UEBA) та спроб несанкціонованого доступу [13].

Для всебічного захисту ТЛЦ пропонується синергетичне поєднання двох взаємодоповнюючих підходів контент-аналізу (кількісного-якісного) [2] та конверсаційного аналізу (якісного), що становить суму процедур емпіричного дослідження діалогічного усного розмовного мовлення [14].

Для обґрунтування вибору найбільш ефективних методів аналізу мережевого контенту в складних гетерогенних системах, якими є сучасні ТЛЦ, необхідно чітко розмежувати можливості та обмеження кількісного та якісного підходів. Порівняльна характеристика контент-аналізу та конверсаційного аналізу, що включає сутність підходів, інструментарій та приклади застосування для специфіки логістичних комунікацій, наведена в табл. 1.

Таблиця 1 – Порівняльна характеристика методів аналізу мережевого контенту для цілей ІБ

Параметр порівняння	**Контент-аналіз**	**Конверсаційний аналіз**
1	2	3
Сутність та підхід	Кількісний скринінг загроз, базується на квантифікаційній обробці тексту (перетворенні тексту в числа/метрики).	Якісна семантична верифікація взаємодії. Досліджує структуру, логіку діалогу та динаміку взаємодії.
Масштаб та швидкість	Дозволяє автоматизувати моніторинг великих масивів даних у реальному часі. Має високу швидкість обробки.	Більш трудомісткий метод, складніший для повної автоматизації на великих обсягах. Глибокий аналіз кожного комунікативного діалогу.
На чому фокусується	Фокусується на формі (слова, фрази, емоційна тональність окремих повідомлень).	Фокусується на контексті та сенсі (порядок ходів, приховані наміри, логіка розмови).

Закінчення табл. 1

1	2	3
Приклади застосування/ інструменти	Сентимент-аналіз: оцінка емоційної напруженості для виявлення деструктивних станів персоналу (стрес, агресія). Частотний аналіз: детекція аномального використання маркерних слів («військовий вантаж», «маршрут», «код доступу»).	Аналіз мовленнєвих ходів: виявлення тиску, маніпуляції або нетипових пауз (наприклад, у розмовах диспетчерів). Ідентифікація соціальної інженерії: розпізнавання претекстингу та вішингу через аналіз відхилень від стандартних протоколів. Виявлення прихованих інтенцій: фіксація підготовки до шкідливих дій на етапі обговорення.
Обмеження та недоліки	Часто ігнорує підтекст та складні маніпулятивні сценарії.	Потребує складніших алгоритмів для автоматизації, вищий ризик хибних спрацювань без контексту.

Джерело: розроблено авторами.

Доцільність інтеграції цих методів у підсистеми підтримки ухвалення рішень систем захисту ІБ ТЛЦ полягає у створенні багаторівневого фільтра безпеки, де контент-аналіз виступає як засіб первинного масового скринінгу, що виділяє підозрілі сесії за формальними ознаками, а конверсаційний аналіз проводить поглиблене дослідження виявлених аномалій, використовуючи алгоритми векторизації та матриці порядку для оцінки структурної цілісності діалогу.

Такий підхід дозволяє перейти системам захисту безпеки від пасивного контролю персоналу до активного моніторингу впливів людського фактору, де будь-яка відхилена від норми комунікація стає тригером для превентивного реагування.

Математична формалізація діалогової взаємодії. Для реалізації інтелектуального аналізу мережевого контенту необхідно перейти від синтаксичного представлення тексту до аналізу його структурно-семіотичних характеристик. Розглянемо комунікаційну подію (діалог) як впорядковану множину текстових одиниць (повідомлень, реплік), що передаються в мережевому середовищі ТЛЦ.

Нехай $D = \{T_1, T_2,\ldots,T_m\}$ – множина діалогів, де кожен окремий діалог $T_j \in D$ (де $j \in \{1,\ldots, m\}$) є результатом конкатенації послідовних текстових одиниць:

$$T_j = t_{j1} \odot t_{j2} \odot \ldots \odot t_{jn},$$

де t_{ji} – i-та текстова одиниця (фраза, речення, повідомлення) у j-му діалозі,

$\odot$ – символ позначає операцію послідовного поєднання з урахуванням часової мітки.

Для виявлення прихованих загроз, пов'язаних із маніпуляціями або інсайдерською діяльністю, недостатньо аналізувати лише частоту вживання слів (класичний TF-IDF) [15]. Необхідно враховувати структурну логіку, порядок та дистанцію між ключовими семантичними концептами.

Введемо множину референтних концептів $C = \{c_1, c_2,\ldots,c_k\}$, що є релевантними для забезпечення інформаційної безпеки та логістичних процесів ТЛЦ (на кшталт: «вантаж», «маршрут», «доступ», «пароль»).

Для кожного діалогу T_j будується матриця дистанцій M_j розмірності $k \times k$ (де k – кількість виділених референтних концептів):

$$M_j = \begin{pmatrix} d_{11} & d_{12} & \ldots & d_{1k} \\ d_{21} & d_{22} & \ldots & d_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ d_{k1} & d_{k2} & \ldots & d_{kk} \end{pmatrix}.$$

Кожен елемент матриці d_{ab} (де $a, b \in \{1,\ldots, k\}$) відображає відносну відстань між позиціями появи концепту c_a та концепту c_b у текстовій структурі діалогу T_j. Розрахунок виконується як кількість текстових одиниць t, що знаходяться між зазначеними концептами. Якщо концепти в діалозі відсутні або слідують у логічно розірваній послідовності, значення d_{ab} прямує до максимального порогу невизначеності D_{max} ($d_{ab} \rightarrow D_{max}$).

Такий підхід дозволяє формалізувати «сценарій» розмови. Наприклад, у легітимній комунікації диспетчера ТЛЦ послідовність референтних концептів $c_{\text{вантаж}} \rightarrow c_{\text{маршрут}} \rightarrow c_{\text{підтвердження}}$ має стабільні та низькі дистанції у матриці M_j. Аномальне відхилення цих дистанцій або поява концептів $c_{\text{доступ}}$ або $c_{\text{пароль}}$ у нетиповому контексті різко змінює профіль матриці M_j, вказуючи на потенційну загрозу.

Отримані матриці M_j можуть бути розрідженими та мати велику розмірність. Для їх подальшої ефективної обробки використовується метод головних компонент (PCA) з метою зниження розмірності [16].

Для цього матриця M_j спочатку розгортається у вектор $vec(M_j)$ довжиною k^2, після чого застосовується PCA-трансформація:

$$V_j = PCA(flatten(M_j)).$$

Зазначені вектори ознак V_j використовуються як вхідні дані для реалізації повноцінного циклу аналізу контенту в гетерогенному середовищі ТЛЦ.

Концептуальна модель інтелектуального аналізу контенту ТЛЦ. Для реалізації проактивного захисту гетерогенного середовища ТЛЦ від загроз людського фактора нами було розроблено інтегрований підхід до аналізу неструктурованого мережевого контенту, що базується на поєднанні методів контент-аналізу та конверсаційного аналізу. Функціонування розробленої моделі підпорядковано загальній структурно-логічній схемі алгоритму, що ілюструє повний цикл обробки даних – від збору з джерел TMS/WMS [17], [18] до формування агрегованих результатів для системи підтримки прийняття рішень (рис. 1).

Рис. 1. Алгоритм інтегрованого підходу до інтелектуального аналізу контенту ТЛЦ

Алгоритм інтегрованого підходу до інтелектуального аналізу мережевого контенту ТЛЦ. Алгоритм складається з послідовної реалізації чотирьох ключових етапів:

Етап 1. Збір даних. Забезпечує централізоване накопичення вихідної інформації у єдиному аналітичному сховищі шляхом автоматизованої агрегації розрізнених потоків даних для подальшої комплексної обробки.

Етап 2. Інтегрований метод інтелектуального аналізу. Центральний етап алгоритму, на якому неструктурований контент трансформується у формалізовані вектори ознак безпеки. Процес обробки включає:

– очищення тексту, токенізацію та нормалізацію даних;

– контент-аналіз для вилучення кількісних ознак (ключові слова, сентимент) та конверсаційний аналіз для ідентифікації структурно-семіотичних аномалій діалогу;

– об'єднання отриманих ознак у єдиний вектор V_j для подальшої оцінки поведінкових профілів (UEBA);

– розрахунок інтегрального показника ризику $R \in [0: 1]$ нейронечіткою мережею на основі бази лінгвістичних правил (сформульованими експертами з ІБ) та класифікація поточної сесії за рівнем загрози [19].

Процес логічного висновку базується на обробці вхідних змінних, що відображають як емоційний стан персоналу, так і відповідність його дій логістичним протоколам.

Прикладом формалізованого нечіткого правила для ідентифікації прихованої загрози (наприклад, логістичної змови або інсайдерського витоку) є така логічна конструкція:

IF (Сентимент є «стрес/терміновість»).

AND (Ключові слова TMS є «нетипові/аномальні»).

AND (Семантична дистанція d_{ab} «вантаж–підтвердження» $\rightarrow D_{\max}$).

THEN (Інтегральний ризик R є «високий/тривога»).

Така конструкція демонструє синергію джерел даних: сентимент-аналіз фіксує психологічну напруженість співробітника, аналіз логів TMS виявляє технічні аномалії, а зростання дистанції d_{ab} у матриці дистанцій M_j свідчить про структурний розрив у стандартному діалоговому сценарії «запит–підтвердження» та дозволяє визначити поріг ризику.

Етап 3. Превентивне реагування. У разі перевищення порогу ризику ($R > 0{,}7$) автоматично виробляється сигнал для активації превентивних протоколів безпеки та механізмів адаптивного захисту підозрілого мережевого контенту ТЛЦ.

Етап 4. Формування результатів. Результатом роботи алгоритму є пакет аналітичних даних, що містить матриці семантичних дистанцій та хмари тегів (маркерів ІБ), інтегровані ознаки, верифіковані гіпотези ризику (підтвердження або заперечення виявленої загрози), оцінка механізмів реагування. Ці дані зберігаються в єдиному аналітичному сховищі для вдосконалення подальшого моніторингу загроз мережевому контенту.

Для деталізації методологічних засад, що лежать в основі описаного інтегрованого підходу, розроблено систему критеріїв, які використовуються для аналізу мережевого контенту в гетерогенному середовищі ТЛЦ (табл. 2).

Таблиця 2 – Критерії для аналізу мережевого контенту щодо забезпечення інформаційної безпеки

Критерій / Вимір	**Об'єкт аналізу (контекст застосування)**	**Опис та ключові індикатори / Маркери**
Кількісні та семантичні метрики	Внутрішній контент (email/чат); Зовнішній контент; Технічний контент.	Частотний аналіз, що включає детекцію аномальної частоти ключових слів (паролі, маршрути, коди, вантажі). Сентимент-аналіз визначає емоційну напруженість спілкування (стрес, агресія, терміновість) перед атакою.
Структурно-комунікативні патерни	Внутрішній контент (email/чат); Зовнішній контент (соціальна інженерія).	Конверсаційний аналіз, що досліджує порядок мовленнєвих ходів, нетипових пауз, перебивань. Ідентифікація мовленнєвих актів з метою виявлення ознак тиску, маніпуляції, прихованих інструкцій або ухилення від відповіді.
Логістична семантика та контекст	Внутрішній контент; Зовнішній контент; Технічний контент (TMS/WMS логи).	Інтеграція з TMS/WMS забезпечує зіставлення комунікацій про вантажі з реальними даними логістичних систем. Маркери логістики/ІБ дозволяють виявляти специфічні терміни (термінал, TMS, код, загроза, фішинг) у нетиповому контексті.
Поведінкові та ризикові профілі	Внутрішній контент (співробітники); Технічний контент (UEBA).	Фіксація відхилень комунікативної поведінки співробітника від норми. Віднесення сесії до специфічних гіпотез (інсайдер, фішинг, змова, доступ тощо).

Джерело: розроблено авторами.

Для ілюстрації ефективності запропонованого підходу та верифікації здатності матриць дистанцій M_j виявляти приховані загрози, було проведено імітаційне моделювання комунікаційних сесій персоналу ТЛЦ. Для моделювання було обрано ключові референтні концепти ІБ ТЛЦ: «вантаж» (c_1), «маршрут» (c_2), «пароль/код» (c_3) та «доступ» (c_4).

Розглянемо два контрастні сценарії діалогової взаємодії, що ілюструють відмінності між легітимною та аномальною комунікацією.

Сценарій 1. Процес відправки вантажу (легітимна комунікація).

Діалог $T_{\text{legit}} = \{t_1, t_2, t_3, t_4, t_5\}$ відображає стандартну робочу взаємодію:

t_1 – диспетчер А: «Прибув вантаж (c_1) на термінал 2.»

t_2 – оператор Б: «Зрозумів. Потрібно затвердити маршрут (c_2) для водія.»

t_3 – диспетчер А: «Так, я перевіряю документи.»

t_4 – диспетчер А: «Все добре. Маршрут (c_2) підтверджено, можна відправляти вантаж (c_1).»

t_5 – оператор Б: «Дякую, відправляю.»

У цьому діалозі концепти c_3 («пароль/код») та c_4 («доступ») відсутні. Відповідна матриця дистанцій M_{legit} (де $D_{\max} = 5$) має наступний вигляд:

$$M_{legit} = \begin{pmatrix} 0 & 1 & 5 & 5 \\ 3 & 0 & 5 & 5 \\ 5 & 5 & 0 & 5 \\ 5 & 5 & 5 & 0 \end{pmatrix}.$$

Сценарій 2. Ознаки інсайдерської змови (аномальна комунікація).

Діалог $T_{\text{anom}} = \{t_1, t_2, t_3, t_4, t_5\}$ імітує спробу інсайдерського витоку інформації:

t_1 – диспетчер А: «Я бачу, що вантаж (c_1) вже на терміналі.»

t_2 – оператор Б: «Так, але не можу зайти в TMS. Дай мені свій пароль ($c3$), щоб я перевірив.»

t_3 – диспетчер А: «Ти ж знаєш правила, це заборонено. Я просто дам тобі доступ (c_4).»
t_4 – оператор Б: «Добре. Мені просто треба знати, який маршрут (c_2).»

У цьому випадку концепти c_3 та c_4 з'являються у нетиповому контексті, а дистанції між ними та логістичними концептами різко змінюються. Матриця дистанцій M_{anom} має такий вигляд:

$$M_{anom} = \begin{pmatrix} 0 & 3 & 1 & 2 \\ 5 & 0 & 5 & 5 \\ 5 & 2 & 0 & 1 \\ 5 & 1 & 5 & 0 \end{pmatrix}.$$

Отримані вектори ознак $V_j = PCA(flatten(M_j))$ для обох сценаріїв було подано на вхід нейронечіткої мережі. Результати розрахунку інтегрального показника ризику R представлено в табл. 3.

Таблиця 3 – Результати розрахунку інтегрального показника ризику R для різних сценаріїв

Сценарій комунікації	Характер взаємодії	Показник ризику R	Стан безпеки	Рекомендована дія СППР
Сценарій 1	Легітимний діалог диспетчера	0,12	Безпечно	Продовження моніторингу
Сценарій 2	Ознаки логістичної змови	0,58	Підозріло	Поглиблена перевірка сесії
Сценарій 3	Спроба витягування пароля	0,84	Високий ризик	Блокування/попередження
Сценарій 4	Запит на несанкціонований доступ	0,9	Критична загроза	Негайне реагування

Джерело: розроблено авторами.

Для легітимних комунікацій показник ризику R стабільно знаходиться у межах [0; 0,3]. Це обумовлено відповідністю структурних дистанцій d_{ab} у матриці M_j стандартним протоколам взаємодії ТЛЦ. Для аномальних сесій, де зафіксовано порушення семантичного порядку або появу критичних концептів ІБ («пароль», «доступ»), показник R різко зростає, переходячи в підозрілу зону (ризик $R \in (0{,}3; 0{,}7]$) або зону високого ризику ($R > 0{,}7$).

Проведений аналіз результатів моделювання підтверджує ефективність запропонованого інтегрованого підходу. Одержані результати, демонструють, що перехід від простого контент-аналізу ключових слів до структурно-семіотичного аналізу діалогів за допомогою матриць дистанцій M_j дозволяє виявляти аномалії, які ігноруються традиційними системами захисту.

Порівняння матриць M_{legit} та M_{anom} показує, що поява нетипових для логістичних комунікацій концептів (таких як «пароль» та «доступ») різко змінює профіль матриці, дистанції d_{ab} між ними та легітимними концептами («вантаж», «маршрут») наближаються до мінімальних значень. Це свідчить про їх тісний контекстний зв'язок, що є ознакою аномального сценарію розмови (маніпуляції, тиску, змови).

Запропонований метод є логічним продовженням ідей адаптивного управління безпекою критичної інфраструктури до якої відносяться ТЛЦ.

Висновки. У роботі вирішено актуальне наукове завдання щодо розробки методу інтелектуального аналізу мережевого контенту для забезпечення ІБ ТЛЦ в умовах активного використання неструктурованих мережевих комунікацій.

Основними науковими та практичними результатами є формалізована концептуальна модель алгоритму інтелектуального аналізу неструктурованого контенту, яка, на відміну від існуючих, базується на синергії методів кількісного-якісного контент-аналізу та якісного конверсаційного аналізу. Обґрунтовано застосування методу векторизації на

основі матриць дистанцій референтних концептів, що дозволяє формалізувати та аналізувати структурно-семіотичні характеристики діалогів для проактивного виявлення прихованих загроз.

Деталізовано математичний апарат методу, який включає розрахунок матриць дистанцій M_j та застосування PCA-трансформації для отримання векторів ознак меншої розмірності. Ці вектори використовуються як вхідні дані для нейронечіткої мережі, що на основі бази лінгвістичних правил розраховує інтегральний показник ризику для поточної комунікаційної сесії.

Проведено моделювання та аналіз результатів, які підтвердили здатність запропонованого методу виявляти аномалії в самій логіці та структурі діалогу, навіть якщо використовуються легітимні слова. Порівняльний аналіз матриць порядку для легітимних та аномальних діалогів показав чіткі відмінності в профілях дистанцій d_{ab}, що забезпечує коректну класифікацію загроз людського фактора в процесі захисту мережевого контенту.

Практичне значення отриманих результатів полягає в обґрунтуванні інтеграції розробленого методу в аналітичні модулі підтримки прийняття рішень в системах захисту, що реалізують автоматизований моніторинг та оцінку комунікацій персоналу.

Перспективи подальших досліджень полягають у вивченні впливу використання ШІ-агентів на динаміку діалогової взаємодії.

Заява про використання генеративного ШІ та технологій на основі ШІ в процесі написання тексту статті.

Для підвищення якості рукопису автори застосували інструменти на основі штучного інтелекту (зокрема, Chat GPT) з метою усунення стилістичних та граматичних помилок. Весь згенерований або модифікований контент був ретельно перевірений, відредагований та схвалений авторами, які несуть повну відповідальність за остаточний зміст публікації.

Список використаних джерел

1. Литвинов, В. В., та ін. (2017). *Методи аналізу та моделювання безпеки розподілених інформаційних систем*. ЧНТУ.

2. Lytvynov, V., Stoianov, N., Skiter, I., Trunova, O. V., & Hrebennyk, A. (2018). Corporate networks protection against attacks using content-analysis of global information space. *Technical Sciences and Technologies, (1),* 115–130. https://doi.org/10.25140/2411-5363-2018-1(11)-115-130.

3. Cheimonidis, P., & Rantos, K. (2023). Dynamic risk assessment in cybersecurity: A systematic literature review. *Future Internet, 15(10*), 324. https://doi.org/10.3390/fi15100324.

4. Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture* (NIST Special Publication 800-207). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-207.

5. Трунов, О. В., & Дорош, М. С. (2025). Систематизація підходів до оцінки ризиків інформаційної безпеки транспортно-логістичних центрів. *Технічні науки та технології, (2 (40))*, 207–220. https://doi.org/10.25140/2411-5363-2025-2(40)-207-220.

6. Гребенник, А. Г., & Трунов, О. В. (2025). Алгоритм визначення агрегованої динамічної оцінки стану безпеки мережевого контенту. *Технічні науки та технології, (3 (41))*, 158–168. https://doi.org/10.25140/2411-5363-2025-3(41)-158-168.

7. Бешлей, М. І. (2021). *Синтез та реалізація інтенційно-орієнтованих інфокомунікаційних мереж для адаптивного надання сервісів* (Дисертація доктора технічних наук, Національний університет «Львівська політехніка»). https://lpnu.ua/sites/default/files/2021/dissertation/11222/disbeshleyminova.pdf.

8. Lutsiv, N., et al. (2021). Deep semisupervised learning-based network anomaly detection in heterogeneous information systems. *Computers, Materials & Continua, 70*(1), 413–431. https://doi.org/10.32604/cmc.2022.018773.

9. Ganzha, M., et al. (2021). Semantic interoperability. In *Interoperability of heterogeneous IoT platforms* (pp. 91–112). Springer. https://doi.org/10.1007/978-3-030-82446-4_5.

10. National Institute of Standards and Technology. (2025). *Addressing visibility challenges with TLS 1.3 within the enterprise* (2nd preliminary draft, NIST Special Publication 1800-37). https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1800-37.pdf.

11. Li, Y., Joshi, K., Wang, X., & Wong, E. (2025). MAVUL: Multi-agent vulnerability detection via contextual reasoning and interactive refinement. *arXiv*. https://arxiv.org/abs/2510.00317.

12. Yao, R., et al. (2025). EcoSafeRAG: Efficient Security through Context Analysis in Retrieval-Augmented Generation. arXiv preprint *arXiv:2505.13506.* https://doi.org/10.48550/arXiv.2505.13506.

13. Ямнич, А., & Коробейнікова, Т. (2025). Формалізований опис процесної моделі динамічного аналізу та прогнозування ризиків інформаційної безпеки для персоналу. *Herald of Khmelnytskyi National University. Technical Sciences, (359 (6.2)),* 74–82. https://doi.org/10.31891/2307-5732-2025-359-80.

14. Іліаді, О. І. (2023). *Проблеми текстології та дискурсології: Конспект лекцій із силабусом*. ПНПУ ім. К. Д. Ушинського.

15. Das, M., Selvakumar, K., & Alphonse, P. J. A. (2021). A comparative study on TF-IDF feature weighting method and its analysis using unstructured dataset. *arXiv*. https://doi.org/10.48550/arXiv.2308.04037.

16. Greenacre, M., Groenen, P. J. F., Hastie, T., D'Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, *2*(1). https://doi.org/10.1038/s43586-022-00184-w.

17. Нечипоренко, Т., & Крохмаль, Р. (2026). Цифрові інструменти логістичного менеджменту в управлінні бізнес-процесами підприємства. *Сталий розвиток економіки, (1 (58))*, 341–348. https://doi.org/10.32782/2308-1988/2026-58-45.

18. Федік, Л. Ю. (2025). Інтеграція регулюючих механізмів у системи управління складом (WMS/WCS): забезпечення ефективності роботизованих комплексів. *Грааль науки, (54)*, 473–482. https://doi.org/10.36074/grail-of-science.18.07.2025.054.

19. Лавров, В., Дудатьєв, А., & Гарнага, В. (2025). Нейро-нечітка система ANFIS для оцінювання ризику дезінформації в умовах інформаційної війни. *Кібербезпека: освіта, наука, техніка, (4 (28))*, 321–333. https://doi.org/10.28925/2663-4023.2025.28.805.

20. Коробейнікова, Т. І., & Ямнич, А. Б. (2024). Багатовимірна матриця класифікації інформації для оцінки ризиків інформаційної безпеки. *Інформаційні технології та комп'ютерна інженерія, (2)*, 91–106. https://doi.org/10.31649/1999-9941-2024-60-2-91-106.

References

1. Lytvynov, V. V., et al. (2017). *Metody analizu ta modeliuvannia bezpeky rozpodilenykh informatsiinykh system [Methods for analysis and modeling of distributed information systems security]*. ChNTU.

2. Lytvynov, V., Stoianov, N., Skiter, I., Trunova, O. V., & Hrebennyk, A. (2018). Corporate networks protection against attacks using content-analysis of global information space. *Technical Sciences and Technologies, (1), 115–130.* https://doi.org/10.25140/2411-5363-2018-1(11)-115-130.

3. Cheimonidis, P., & Rantos, K. (2023). Dynamic risk assessment in cybersecurity: A systematic literature review. Future Internet, 15(10), 324. https://doi.org/10.3390/fi15100324.

4. Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). Zero Trust Architecture (NIST Special Publication 800-207). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-207.

5. Trunov, O. V., & Dorosh, M. S. (2025). Systematyzatsiia pidkhodiv do otsinky ryzykiv informatsiinoi bezpeky transportno-lohistychnykh tsentriv [Systematization of approaches to information security risk assessment of transport and logistics centers]. *Tekhnichni nauky ta tekhnolohii – Technical Sciences and Technologies,* (2(40)), 207–220. https://doi.org/10.25140/2411-5363-2025-2(40)-207-220.

6. Hrebennyk, A. H., & Trunov, O. V. (2025). Alhorytm vyznachennia ahrehovanoi dynamichnoi otsinky stanu bezpeky merezhevoho kontentu [Algorithm for determining the aggregated dynamic assessment of the network content security state]. *Tekhnichni nauky ta tekhnolohii – Technical Sciences and Technologies,* (3(41)), 158–168. https://doi.org/10.25140/2411-5363-2025-3(41)-158-168.

7. Beshleyi, M. I. (2021). Syntez ta realizatsiia intentsiino-oriientovanykh infokomunikatsiinykh merezh dlia adaptyvnoho nadannia servisiv [Synthesis and implementation of intent-oriented infocommunication networks for adaptive service delivery] (Doctoral dissertation, Lviv Polytechnic National University). https://lpnu.ua/sites/default/files/2021/dissertation/11222/disbeshleyminova.pdf.

8. Lutsiv, N., et al. (2021). Deep semisupervised learning-based network anomaly detection in heterogeneous information systems. *Computers, Materials & Continua, 70*(1), 413–431. https://doi.org/10.32604/cmc.2022.018773.

9. Ganzha, M., et al. (2021). Semantic interoperability. In Interoperability of Heterogeneous IoT Platforms (pp. 91–112). Springer. https://doi.org/10.1007/978-3-030-82446-4_5.

10. National Institute of Standards and Technology. (2025). Addressing Visibility Challenges with TLS 1.3 within the Enterprise (*2nd Preliminary Draft; NIST Special Publication 1800-37).* https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1800-37.pdf.

11. Li, Y., Joshi, K., Wang, X., & Wong, E. (2025). MAVUL: Multi-Agent Vulnerability Detection via Contextual Reasoning and Interactive Refinement. arXiv. https://arxiv.org/abs/2510.00317.

12. Yao, R., et al. (2025). EcoSafeRAG: Efficient Security through Context Analysis in Retrieval-Augmented Generation. arXiv preprint. https://doi.org/10.48550/arXiv.2505.13506.

13. Yamnych, A., & Korobeinikova, T. (2025). Formalizovanyi opys protsesnoi modeli dynamichnoho analizu ta prohnozuvannia ryzykiv informatsiinoi bezpeky dlia personalu [Formalized description of the process model for dynamic analysis and forecasting of information security risks for personnel]. *Herald of Khmelnytskyi National University. Technical Sciences,* (359(6.2)), 74–82. https://doi.org/10.31891/2307-5732-2025-359-80.

14. Iliadi, O. I. (2023). Problemy tekstolohii ta dyskursolohii: Konspekt lektsii iz sylabusom [Problems of textology and discoursology: Lecture notes with syllabus] (Study guide). *PNPU im. K. D. Ushynskoho*.

15. Das, M., Selvakumar, K., & Alphonse, P. J. A. (2021). A Comparative Study on TF-IDF feature Weighting Method and its Analysis using Unstructured Dataset. arXiv. https://doi.org/10.48550/arXiv.2308.04037.

16. Greenacre, M., Groenen, P. J. F., Hastie, T., et al. (2022). Principal component analysis. *Nature Reviews Methods Primers, 2*(1), Article 100. https://doi.org/10.1038/s43586-022-00184-w.

17. Nechyporenko, T., & Krokhmal, R. (2026). Tsyfrovi instrumenty lohistychnoho menedzhmentu v upravlinni biznes-protsesamy pidpryiemstva [Digital tools of logistics management in the management of business processes of the enterprise]. *Stalyi rozvytok ekonomiky – Sustainable Development of Economy,* (1(58)), 341–348. https://doi.org/10.32782/2308-1988/2026-58-45.

18. Fedik, L. Y. (2025). Intehratsiia rehuliuiuchykh mekhanizmiv u systemy upravlinnia skladom (WMS/WCS): Zabezpechennia efektyvnosti robotyzovanykh kompleksiv [Integration of regulating mechanisms into warehouse management systems (WMS/WCS): Ensuring the efficiency of robotic complexes]. *Grail of Science,* (54), 473–482. https://doi.org/10.36074/grail-of-science.18.07.2025.054

19. Lavrov, V., Dudatiev, A., & Harnaha, V. (2025). Neiro-nechitka systema ANFIS dlia otsiniuvannia ryzyku dezinformatsii v umovakh informatsiinoi viiny [Neuro-fuzzy ANFIS system for assessing the risk of disinformation in conditions of information warfare]. *Kiberbezpeka: osvita, nauka, tekhnika – Cybersecurity: Education, Science, Technique,* (4(28)), 321–333. https://doi.org/10.28925/2663-4023.2025.28.805.

20. Korobeinikova, T. I., & Yamnych, A. B. (2024). Bahatovymirna matrytsia klasyfikatsii informatsii dlia otsinky ryzykiv informatsiinoi bezpeky [A multidimensional information-classification matrix for information security risk assessment]. *Informatsiini tekhnolohii ta kompiuterna inzheneriia – Information Technologies and Computer Engineering,* (2), 91–106. https://doi.org/10.31649/1999-9941-2024-60-2-91-106.

Дата першого надходження статті до видання: 09.03.2026
Дата прийняття статті до друку після рецензування: 20.03.2026

TECHNICAL SCIENCES AND TECHNOLOGIES

UDC 004.891:004.056.5:656.07

Alla Hrebennyk[1], Oleksii Trunov [2]

[1]Senior Lecturer at the Department of Cybersecurity and Mathematical Modeling
Chernihiv Polytechnic National University (Chernihiv, Ukraine)
E-mail: grebennik.alla@gmail.com **ORCID:** https://orcid.org/0000-0002-7464-1412
Scopus Author ID: 57219054928

[2]PhD student, Lecturer at the Department of Information Technology and Software Engineering
Chernihiv Polytechnic National University (Chernihiv, Ukraine)
E-mail: alexeytrunov1995@gmail.com **ORCID:** https://orcid.org/0009-0002-0321-2669
Scopus Author ID: 59337119500

ALGORITHM FOR INTELLIGENT ANALYSIS OF NETWORK CONTENT FOR ENSURING THE INFORMATION SECURITY OF A TRANSPORT AND LOGISTICS CENTER

In the context of current geopolitical instability and full-scale war, transport and logistics centers (TLCs) are becoming strategically important hubs whose resilience directly impacts national security. However, the growing number of hybrid attacks targeting the human factor through social engineering and insider activity creates new challenges for information security systems. Effective protection of TLCs at the current stage is significantly limited by the inadequacy of traditional monitoring tools against emerging threats implemented through text communications with a high level of semantic uncertainty. Keyword-only searching, without accounting for the structural logic of a dialogue, fails to identify complex manipulation scenarios where the threat is hidden within the context rather than specific terms.

In view of this, the aim of the study is to develop a method for the intelligent analysis of TLC network content based on a combination of content analysis, conversation analysis, and vectorization algorithms using distance matrices to automate the detection of hidden threats. The synergy of these methods allows for both mass screening of messages and deep semantic verification of dialogue interaction. The mathematical framework of the method is based on the formalization of conversation "scenarios" through the construction of distance matrices M_j, where the elements represent relative distances between reference security concepts.

To reduce the dimensionality of the resulting data, Principal Component Analysis (PCA) is employed, allowing for the formation of feature vectors for subsequent processing by a neuro-fuzzy network. Based on a linguistic rule base, the network calculates an integrated risk indicator $R \in [0; 1]$, which serves as a trigger for activating preventive measures within security decision-support systems. Simulation results of legitimate and anomalous communication scenarios confirmed the high sensitivity of the method to structural deviations in dialogue logic, even when personnel utilize legitimate vocabulary. The analysis demonstrated a clear differentiation of risk levels for various types of threats, verifying the effectiveness of the developed mathematical framework. The practical significance of the results lies in the possibility of integrating the method into analytical protection modules for logistics systems, enabling security state assessment and proactive response to human-factor-related incidents. Prospects for further research involve studying the impact of AI agents on the dynamics of dialogue interaction.

Keywords: *information security; transport and logistics center; intelligent content analysis; conversation analysis; distance matrix; human factor; social engineering; decision support system.*

Fig.: 1. Table: 3. References: 20.

Гребенник А. Г., Трунов О. І. Алгоритм інтелектуального аналізу мережевого контенту для забезпечення інформаційної безпеки транспортно-логістичного центру. *Технічні науки та технології,* 2026. № 2(44). С. 248-259. DOI: https://doi.org/10.25140/2411-5363-2026-2(44)-248-259.