УДК 004.056.5

Козюк Ю.Ю.

Грицак А.В.

АЛГОРИТМ ВИЯВЛЕННЯ АНОМАЛІЙ У ПОВІДОМЛЕННЯХ КОРПОРАТИВНОГО МЕСЕНДЖЕРА

Вінницький національний технічний університет

Анотація. У сучасному корпоративному середовищі важливо оперативно виявляти повідомлення, що можуть свідчити про порушення політик безпеки або потенційні загрози. У цій роботі представлено алгоритм виявлення аномалій у повідомленнях месенджера, що поєднує контентний і контекстний аналіз. Система базується на статистичних ознаках, поведінкових шаблонах і семантичних правилах, що дозволяє ідентифікувати підозрілі повідомлення з високою точністю. Такий підхід забезпечує ефективний моніторинг комунікацій без складних обчислювальних ресурсів.

Ключові слова: аномалія, корпоративний месенджер, безпека, контекстний аналіз, контентний аналіз, семантична інформація, виявлення відхилень, поведінкові шаблони.

Abstract. *In the modern corporate environment, it is important to quickly detect messages that may indicate a violation of security policies or potential threats. This paper presents an algorithm for detecting anomalies in messenger messages that combines content and contextual analysis. The system is based on statistical features, behavioural patterns, and semantic rules, which allows identifying suspicious messages with high accuracy. This approach ensures effective monitoring of communications without complex computing resources.*

Keywords: *anomaly, corporate messenger, security, contextual analysis, content analysis, semantic information, anomaly detection.*

Вступ

У сучасних умовах цифрової трансформації особливої актуальності набуває проблема своєчасного виявлення відхилень у цифровій комунікації. Зі зростанням обсягів обміну інформацією в корпоративних месенджерах виникає потреба у впровадженні ефективних механізмів аналізу змісту повідомлень. Комбіноване застосування контентного та контекстного аналізу дозволяє ідентифікувати аномальні повідомлення, що можуть свідчити про потенційні загрози або нетипову поведінку користувача. Запропонований підхід забезпечує адаптивність системи до специфіки корпоративного середовища, зберігаючи гнучкість, масштабованість та високу ефективність в умовах динамічних комунікаційних процесів.

Результати дослідження

У корпоративному середовищі надзвичайно важливо забезпечити контроль за несанкціонованими або підозрілими діями всередині комунікаційної системи. З цією метою в месенджер інтегровано систему виявлення аномалій, що поєднує методи контентного та контекстного аналізу повідомлень. В основі рішення — модель, яка класифікує кожне повідомлення як нормальне (0) або аномальне (1) на основі попередньо визначених ознак і поведінкових шаблонів [1].

Наступним чином можна побачити схематичне представлення алгоритму виявлення аномальних повідомлень у корпоративному месенджері. На рисунку 1 показано основні етапи роботи системи виявлення аномалій. Вхідним параметром є поточне повідомлення користувача, яке передається до модуля контентного аналізу. На цьому етапі система виділяє статистичні ознаки, що можуть вказувати на аномалію, такі як довжина повідомлення, частота використання нетипових слів або символів, а також структурні особливості тексту. Паралельно повідомлення проходить через модуль контекстного аналізу, де оцінюється відповідність змісту повідомлення до історичного профілю поведінки користувача. Тут враховується час надсилання, тематика, тональність, а також відповідність очікуваним шаблонам взаємодії в межах конкретного каналу комунікації (наприклад, проєктна група, технічна підтримка тощо).

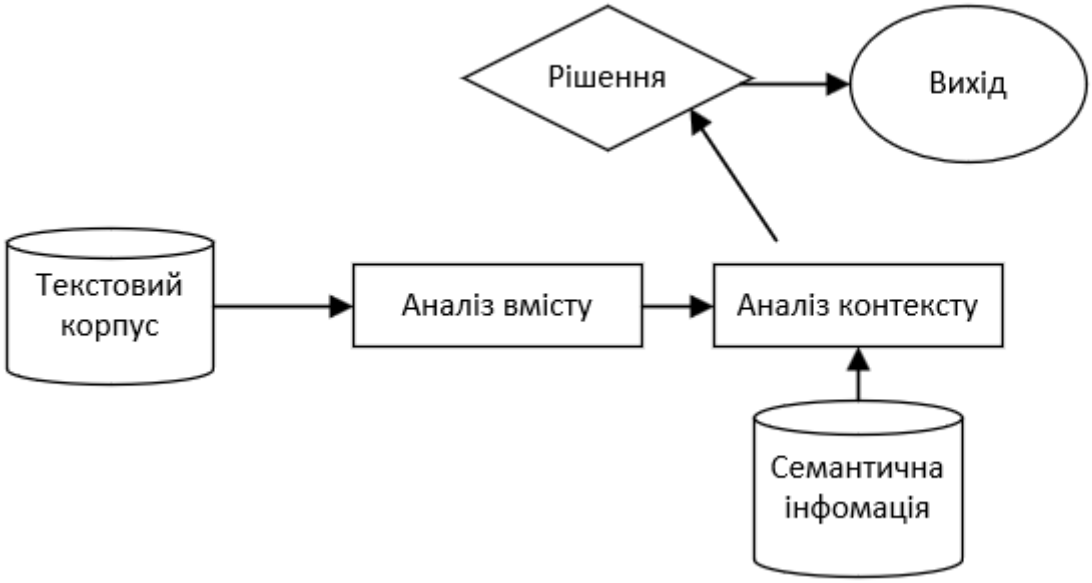

Рисунок 1– Схематичне представлення алгоритму виявлення аномальних повідомлень у корпоративному месенджері

Відповідно до представленої блок-схеми, функціонування системи передбачає проходження таких етапів алгоритмічної обробки, як формування текстового корпусу, контентний аналіз, виділення семантичної інформації, контекстний аналіз, прийняття рішення та виведення результатів.

Першим етапом функціонування системи виявлення аномалій є формування текстового корпусу, який містить сукупності попередньо зібраних повідомлень користувачів месенджера, що слугують базовими даними для навчання та калібрування алгоритмів. Окрім текстового вмісту, ці повідомлення містять також метаінформацію, зокрема часові мітки, ідентифікатори користувачів та частотні характеристики взаємодій. На основі цього корпусу будується модель типової поведінки користувачів, яка дозволяє системі виявляти відхилення від звичних патернів спілкування. Формування якісного та репрезентативного корпусу є критично важливою умовою забезпечення високої точності виявлення аномалій.

Другим етапом роботи системи виявлення аномалій є контентний аналіз, що передбачає обробку текстових повідомлень на основі їхніх статистичних характеристик. На цьому етапі кожне повідомлення проходить попередню обробку, у межах якої виділяються ознаки (features), потенційно релевантні для виявлення аномалій. До таких ознак належать: середня довжина повідомлення, частота використання специфічних слів або символів (наприклад, «ffffuuhigggg», «!!!»), частка великих літер, а також кількість повідомлень, надісланих за короткий часовий проміжок. Отримані характеристики перетворюються у числовий вектор, який подається на вхід моделі машинного навчання (наприклад, Isolation Forest, One-Class SVM або Autoencoder). Результатом роботи моделі є бінарна класифікація: 0 — нормальне повідомлення, 1 — виявлена аномалія. Таким чином, на цьому етапі система фіксує аномалії, що виявляються на основі статистичних відхилень від норми.

Третім етапом роботи системи виявлення аномалій є обробка семантичної інформації, яка виконує функцію збереження довідкових даних, необхідних для інтерпретації змісту повідомлень. До складу цього етапу входять заздалегідь сформовані словники ризикованих слів або фраз (наприклад, «пароль», «виведи гроші», «система зламана»), векторні подання слів (ембеддинги), отримані за допомогою моделей типу Word2Vec, FastText або BERT, а також інформація про допустимі тематики обговорення в межах конкретного каналу комунікації (наприклад, службові проєктні дискусії). Отримані семантичні характеристики є основою для подальшого контекстного аналізу й дозволяють доповнити оцінку повідомлення шляхом зіставлення його змісту з очікуваною поведінкою користувача.

Четвертим етапом функціонування системи виявлення аномалій є контекстний аналіз, який передбачає багатовимірне порівняння поточного повідомлення з історичним профілем поведінки користувача. На цьому етапі алгоритм оцінює наявність стилістичних характеристиках (зміни у словниковому запасі, тональності чи синтаксичних структурах), а також тематичному спрямуванні комунікації (наприклад, перехід від робочих обговорень до тем, пов’язаних із особистими фінансами). Крім того, здійснюється перевірка відповідності змісту повідомлення встановленим корпоративним стандартам і політикам. Виявлені на цьому етапі контекстні аномалії слугують підґрунтям для формування остаточного рішення щодо класифікації повідомлення.

П’ятим етапом роботи системи виявлення аномалій є етап прийняття рішення, який полягає в інтеграції результатів контентного та контекстного аналізу. Для цього застосовується логічне правило або ансамблева модель, що враховує сигнали з двох джерел: статистичних ознак повідомлення, отриманих у процесі контентного аналізу, та поведінкових і семантичних характеристик, виявлених під час контекстного аналізу. У базовій конфігурації повідомлення класифікується як підозріле, якщо хоча б один з аналізаторів фіксує аномалію (тобто результат дорівнює «1»). У випадку потенційного хибного спрацювання передбачена можливість додаткової перевірки повідомлення за допомогою спеціального модуля повторної валідації, який застосовує альтернативні евристичні правила та більш глибокий семантичний аналіз.

Шостим етапом функціонування системи виявлення аномалій є етап виведення результату, який реалізує дії на основі прийнятого рішення. Якщо повідомлення класифіковано як безпечне (результат 0), воно безперешкодно доставляється до відповідного чату без додаткових обмежень. У випадку виявлення аномалії (результат 1), повідомлення автоматично блокується або передається адміністратору для подальшої перевірки. Крім того, система може фіксувати подію у журналі безпеки для подальшого аудиту. Такий підхід забезпечує оперативне реагування на потенційні загрози, зокрема шахрайство, фішинг або витік конфіденційної інформації.

Тож описана система виявлення аномалій у текстових повідомленнях є багаторівневою архітектурою, яка поєднує як формальні ознаки повідомлень, так і глибше розуміння їхнього змісту та контексту. Цей підхід дозволяє виявляти не тільки очевидно підозрілі повідомлення, а й реагувати на складніші, приховані загрози.

З метою оцінки ефективності розробленого підходу до виявлення аномалій у текстових повідомленнях доцільним є проведення порівняльного аналізу з існуючими рішеннями, що вже застосовуються для аналогічних завдань. Такий аналіз дозволяє об’єктивно визначити ключові переваги запропонованої системи, а також окреслити її потенційні обмеження в умовах практичного використання. Основні результати порівняння представлені у (таблиці 1**)**, яка містить характеристику популярних алгоритмів виявлення аномалій та їх зіставлення із запропонованим рішенням за низкою критеріїв.

Таблиця 1– Порівняння розробленої системи з існуючими рішеннями

Критерій	SpamAssassin	Perspective API	OpenAI Moderation API	FastText + Logistic Regression	Розроблена система
Тип повідомлень	Email	Текстові повідомлення	Текстові повідомлення	Текстові повідомлення	Текстові повідомлення
Контентний аналіз	Так	Так	Так	Так	Так
Контекстний аналіз	Ні	Ні	Обмежено	Ні	Так
Урахування поведінки користувача	Ні	Ні	Ні	Ні	Так
Підтримка семантичних словників/тем	Обмежено	Обмежено	Частково	Ні	Так
Можливість кастомізації під домен	Обмежено	Ні	Ні	Так	Так
Використання ML/AI	Ні	Так	Так	Так	Так
Пояснюваність рішень	Частково	Обмежено	Ні	Так	Так
Автоматичне реагування на загрозу	Так	Ні	Ні	Ні	Так
Підтримка журналу/аудиту	Частково	Ні	Ні	Ні	Так
Середня точність на тестовому наборі	~65%	~78%	~85%	~70%	92–95%

Тож зіставлення характеристик у Таблиці 1 свідчить про те, що запропонована система має розширений функціонал, який охоплює не лише контентний, а й контекстуальний аналіз, а також урахування поведінкових особливостей користувачів. Завдяки такій інтегрованій структурі система демонструє вищу ефективність виявлення аномальних повідомлень порівняно з існуючими аналогами, забезпечуючи гнучкість та відповідність вимогам конкретного корпоративного середовища.

Висновки

Отже, у роботі проаналізовано ключові етапи функціонування системи виявлення аномалій у корпоративному месенджері та запропоновано комплексний підхід, що поєднує контентний і контекстний аналіз. Встановлено, що інтеграція статистичних ознак повідомлень із поведінковими та семантичними характеристиками дозволяє підвищити точність виявлення потенційно небезпечного контенту. Контентний аналіз забезпечує обробку поверхневих характеристик тексту, тоді як контекстний дозволяє врахувати відхилення від звичної моделі комунікації користувача. Така багаторівнева архітектура сприяє своєчасному виявленню складних загроз, підвищує надійність системи безпеки месенджера та відповідає сучасним вимогам до захисту корпоративного інформаційного середовища.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Kidney and Liver Function Indices of Prosopis Africana Seed Extract on Testosterone and Estradiol-Induced Benign Prostatic Hyperplasia in Adult Male Rats [Електронний ресурс] / R. O. Adeyemi, O. A. Ogunlesi, O. T. Olusola, O. S. Olamide. – Режим доступу до ресурсу: https://www.researchgate.net/publication/330026215_Kidney_And_Liver_Function_Indices_Of_Prosopis_Africana_Seed_Extract_On_Testosterone_And_Estradiol_Induced_Benign_Prostatic_Hyperplasia_In_Adult_Male_Rats. (Дата звернення 27.05.2025).

2. A Comprehensive Review on Security and Privacy Issues in Internet of Things (IoT) [Електронний ресурс] / A. K. Sharma, N. A. S. C. Rani, K. T. Z. Zaheer. – Режим доступу до ресурсу: https://www.sciencedirect.com/science/article/pii/S2352914822000146. (Дата звернення 27.05.2025).

3. Oloid: Advanced Biometric Access Control Solutions [Електронний ресурс]. – Режим доступу до ресурсу: https://www.oloid.com/. (Дата звернення 27.05.2025).

4. Anomaly Detection with Machine Learning Techniques and Applications [Електронний ресурс]. – Режим доступу до ресурсу: https://www.doit.com/anomaly-detection-with-machine-learning-techniques-and-applications/. (Дата звернення 27.05.2025).

Козюк Юлія Юріївна – студентка групи 2КІТС-21б, факультет менеджменту та інформаційної безпеки, Вінницький національний технічний університет, Вінниця, e-mail: juliakozyk999@gmail.com

Науковий керівник: **Грицак Анатолій Васильович** – кандидат технічних наук спеціальністю 125«Кібербезпека», доцент кафедри менеджменту та безпеки інформаційних систем, Вінницький національний технічний університет, Вінниця, e-mail: grytsak.a.v@gmail.com

Koziuk Yuliia Y. – student of the 2KITS-21b group, Faculty of Management and Information Security, Vinnytsia National Technical University, Vinnytsia, e-mail: juliakozyk999@gmail.com

Supervisor: **Hrytsak Anatoliy V.** - Candidate of Technical Sciences in 125 "Cybersecurity", Associate Professor of the Department of Management and Security of Information Systems, Vinnytsia National Technical University, Vinnytsia, e-mail