УДК 004.89
Боднар Д. - ст. гр. СНнм-61
*Тернопільський національний технічний університет імені Івана Пулюя*

# СИНЕРГІЯ МЕТОДІВ ОБРОБКИ ПРИРОДНОЇ МОВИ ТА ПОВЕДІНКОВОЇ АНАЛІТИКИ В ІНТЕЛЕКТУАЛЬНИХ СИСТЕМАХ МОНІТОРИНГУ

Науковий керівник: к. соц. ком Липак Г.І.

Bodnar D.
*Ternopil Ivan Puluj National Technical University*

# THE SYNERGY OF NATURAL LANGUAGE PROCESSING AND BEHAVIOURAL ANALYTICS IN INTELLIGENT MONITORING SYSTEMS

Supervisor: Ph.D Lypak H.I.

Сучасні системи безпеки більше не можуть покладатися лише на аналіз цифрових слідів (час входу, обсяг трафіку), оскільки цього недостатньо для виявлення складних аномалій. Ключовим напрямком розвитку стає поєднання обробки природної мови (NLP) з поведінковою аналітикою (UBA). Такий підхід дозволяє не просто фіксувати дії користувача, а інтерпретувати його наміри, аналізуючи цифровий слід у комунікаціях [1].

Впровадження семантичного аналізу ґрунтується на використанні сучасних архітектур типу Transformer (наприклад, BERT), які, на відміну від застарілих методів (TF-IDF), здатні розпізнавати тонкі нюанси професійної лексики та зміни в тональності повідомлень [2]. Схема інтеграції цих методів у єдиний контур моніторингу наведена на рисунку 1.


Рисунок 1. Схема архітектури комбінованої системи виявлення внутрішніх загроз на основі NLP та UBA

Важливим елементом є перетворення неструктурованих текстів у графові структури. Використання методів навчання на графах дозволяє візуалізувати не лише

факт обміну повідомленнями, а й сутності (теми), про які йдеться [3]. Це допомагає виявити аномальні інформаційні потоки, наприклад, коли співробітник починає обговорювати нетипові для нього теми або формує нові нехарактерні зв'язки у соціальному графі організації. Схема інтеграції цих методів у єдиний контур моніторингу наведена на рисунку 1.

Порівняльний аналіз демонструє, що такий інтегрований підхід суттєво знижує рівень хибнопозитивних спрацювань. Семантичний контекст слугує підтвердженням або спростуванням аномалії: технічний сплеск активності, який супроводжується використанням специфічних інструментів, описаних у документації розробників, класифікується системою як норма, а не як загроза [4].

Запропонована синергія методів NLP та UBA відкриває новий рівень інтелектуального моніторингу внутрішніх загроз. Поєднання глибокого семантичного розуміння тексту з графовим представленням комунікацій дозволяє не лише виявляти аномалії, а й інтерпретувати їхній контекст, значно зменшуючи кількість хибних тривог. Подальші дослідження доцільно спрямувати на емпіричну верифікацію ефективності системи на реальних корпоративних даних, інтеграцію з іншими джерелами (лог-файли, аудити доступу) та розробку адаптивних моделей, що враховують специфіку різних галузей. Таким чином, комбінований підхід може стати основою для створення проактивних систем кібербезпеки наступного покоління.

## Списки використаних джерел

1. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding // Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL). – 2019. – P. 4171–4186.
2. Naseer, H., et al. User and Entity Behavior Analytics for Insider Threat Detection: A Review // Computers & Security. – 2024. – Vol. 136. – P. 103–541.
3. Hamilton, W. L. Graph Representation Learning // Synthesis Lectures on Artificial Intelligence and Machine Learning. – Morgan & Claypool Publishers, 2020. – 159 p.
4. Zhu, Y., & Yan, J. Semantic-aware Behavioral Modeling for Enterprise Security // IEEE Transactions on Information Forensics and Security. – 2025. – Vol. 20. – P. 442–457.