УДК 004.056:004.8

Я. В. Кордонський
О. П. Войтович

ГІБРИДНИЙ ПІДХІД ДО АВТОМАТИЗОВАНОГО ВИЯВЛЕННЯ КОНФІДЕНЦІЙНОЇ ІНФОРМАЦІЇ У ФАЙЛАХ

Вінницький національний технічний університет

Анотація

Запропоновано гібридний підхід до побудови та впровадження засобу автоматизованого аналізу файлових масивів щодо конфіденційної інформації, що дозволяє виявити приховані внутрішні загрози та запобігти критичним втратам даних, які виникають через неуважність або низьку цифрову гігієну працівників.

Ключові слова: автоматизований аналіз, конфіденційна інформація, загрози, дані, цифрова гігієна.

Abstract

A hybrid approach is proposed for the development and implementation of an automated tool for analyzing file arrays for confidential information, which enables the detection of hidden insider threats and prevents critical data loss resulting from employee negligence or poor digital hygiene.

Keywords: automated analysis, confidential information, threats, data, digital hygiene.

Вступ

Сьогодні цифровізація корпоративних процесів призводить до накопичення величезних масивів даних, що значно підвищує ризики несанкціонованого доступу до комерційних таємниць та персональних даних [1]. Основним джерелом витоків часто є не зовнішні кібератаки, а внутрішні загрози, спричинені недотриманням персоналом правил цифрової гігієни та випадковим збереженням або пересиланням конфіденційної інформації у відкритих сховищах [2]. Метою дослідження є вибір актуальної методики автоматизованого аналізу вмісту файлів щодо наявності конфіденційної інформації для запобігання її витоку та посилення проактивного захисту підприємства.

Результати дослідження

Сучасна автоматизація пошуку конфіденційної інформації в неструктурованих файлових масивах базується на стратегічному поєднанні детермінованих алгоритмів із ймовірнісними моделями машинного навчання, що знаменує перехід від жорстких правил до гнучкого семантичного аналізу. Регулярні вирази залишаються стандартом для ідентифікації даних із фіксованою структурою, таких як номери банківських карток, ідентифікаційні коди або стандартні формати документів. Перевагою є надзвичайна швидкість обробки, яка може бути у десятки тисяч разів вищою за LLM [3]. Проте регулярні вирази виявляються неефективними у "забруднених" текстах де наприклад, номер телефону написаний словами "п'ять п'ять п'ять" або розділений нестандартними символами, звичайний шаблон не спрацює, крім того, вони позбавлені контекстуального розуміння. Наприклад якщо два слова поруч написані з великої літери, він додасть їх до групи імен, а це буде не ім’я, а назва компанії, то це буде колізія, яких може бути багато [4].

Впровадження трансформерних моделей та архітектур типу BERT [5] докорінно змінило правила гри завдяки механізмам контекстної уваги, які дозволяють виконувати розпізнавання іменованих сутностей на основі лінгвістичного оточення. Сучасні дослідження 2025 року [6-8] вказують на те, що оптимальним рішенням є побудова багаторівневих конвеєрів "pipelines". Прикладом такого підходу є фреймворк RECAP (Regex and Context-Aware Prompting), який поєднує детерміновану точність правил із когнітивними можливостями LLM через трифазну архітектуру, яка включає базове гібридне виявлення, вирішення багатозначності для спірних фрагментів тексту та фінальну контекстну фільтрацію помилкових спрацювань [8]. Паралельно з цим, алгоритм CHPDA (Context-Aware Hybrid Pattern Detection Algorithm) інтегрує швидкодіючі рушії типу RE2 з оцінкою близькості ключових слів, що дозволяє досягти значення успішного розпізнавання на рівні 0.91 при збереженні високої масштабованості [9].

Технічне порівняння Regex та LLM, проведене у 2025 році, виявило, що для стандартизованих даних, а саме медичних кодів BI-RADS, регулярні вирази працюють у 28120 разів швидше при майже ідентичній точності [1]. Проте для нечітких описів LLM є незамінним інструментом та підтверджує доцільність використання гібридних архітектур, що корелює із концепцією Zero Trust [10], де доступ до файлу надається не лише на основі ідентичності користувача, а й на основі автоматизованої оцінки чутливості самого вмісту [2].

Оптимальна методика аналізу має працювати за принципом нарощування складності: від початкової індексації та мета-аналізу через швидке сканування регулярними виразами до залучення локальних мовних моделей для аналізу найбільш неоднозначних фрагментів. Їх робота завершується автоматичним присвоєнням міток чутливості та накладанням відповідних політик захисту від витоків [11].

Висновки

Гібридний підхід дозволяє перетворити статичне сховище файлів на динамічно кероване середовище, де рівень захисту кожного об'єкта автоматично адаптується до зміни його змісту та створює надійний фундамент для побудови відмовостійкої системи безпеки, яка здатна ефективно масштабуватися разом із розширенням інформаційної інфраструктури підприємства Найбільш ефективним рішенням для ідентифікації конфіденційної інформації є впровадження гібридних архітектур, які поєднують високу обчислювальну швидкість регулярних виразів із семантичним інтелектом мовних моделей.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Egho-Promise E. I., Sitti M. Big Data Security Management in Digital Environment. *American Journal of Multidisciplinary Research & Development (AJMRD)*. 2024. URL: https://www.researchgate.net/publication/385411123_Big_Data_Security_Management_in_Digital_Environment.
2. Sarah Mathew B., Lasekan O. A., Godoy Pena M. Behavioural Determinants of Cyber Hygiene: Rethinking the Role of Education in Digital Trust and Usability. *International Research Journal of Multidisciplinary Scope*. 2025. URL: https://doi.org/10.47857/irjms.2025.v06i04.07251.
3. A comparative performance analysis of regular expressions and a large language model-based approach to extract the BI-RADS score from radiological reports / F. Dennstädt et al. *JAMIA Open*. 2025. Vol. 8, no. 6. URL: https://doi.org/10.1093/jamiaopen/ooaf128.
4. Luo Q., Zhang G., Sun Y. A Novel Method for Named Entity Recognition in Long-Text Safety Accident Reports of Prefabricated Construction. *Buildings*. 2025. URL: https://doi.org/10.3390/buildings15173063.
5. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin et al. *arXiv*. 2019. URL: https://doi.org/10.48550/arXiv.1810.04805.
6. Melnyk S. Safe Observability: A Framework for Automated PII Redaction from LLM Prompts in OpenTelemetry Pipelines. *International Journal of Computer (IJC)*. 2025. URL: https://ijcjournal.org/InternationalJournalOfComputer/article/view/2458.
7. Large Language Models For Text Classification: Case Study And Comprehensive Review / A. Kostina et al. *arXiv*. 2025. URL: https://doi.org/10.48550/arXiv.2501.08457.
8. An Evaluation Study of Hybrid Methods for Multilingual PII Detection / H. Rajgarhia et al. *arXiv*. 2025. URL: https://doi.org/10.48550/arXiv.2510.07551.
9. Koli L., Kalra S., Singh K. Decoding Complexity: Intelligent Pattern Exploration with CHPDA (Context Aware Hybrid Pattern Detection Algorithm). *arXiv*. 2025. URL: https://doi.org/10.48550/arXiv.2502.07815.
10. A Survey on Zero Trust Architecture: Challenges and Future Trends / Y. He et al. *Wireless Communications and Mobile Computing*. 2022. Vol. 2022. P. 1–13. URL: https://doi.org/10.1155/2022/6476274.
11. Adaptive PII Mitigation Framework for Large Language Models / S. Asthana et al. *arXiv*. 2025. URL: https://doi.org/10.48550/arXiv.2501.12465.

Кордонський Ярослав Володимирович — студент 2БС-22б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, e-mail: 04-22-237.stud@vntu.vn.ua.

Науковий керівник: ***Войтович Олеся Петрівна*** — к. тех. н., доцент, кафедра захисту інформації, Вінницький національний технічний університет, м. Вінниця

Kordonskyy Yaroslav V. — student of group 2BS-22b, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: 04-22-237.stud@vntu.vn.ua.

Research supervisor: ***Voytovich Olesya P.*** — Cand. in Techical Sciences, Associate Professor, Department of Information Protection, Vinnytsia National Technical University, Vinnytsia.