УДК 004.8+004.056

Л. М. Куперштейн
О. В. Залепа

ДО ПРОБЛЕМИ ЗБОРУ ДОМЕННОЇ ІНФОРМАЦІЇ НА ОСНОВІ МУЛЬТИАГЕНТНОГО ПІДХОДУ ДЛЯ ЗАДАЧ КІБЕРБЕЗПЕКИ

Вінницький національний технічний університет

Анотація

У роботі розглянуто доменну інформацію (WHOIS, DNS, SSL/TLS, дані з відкритих джерел) як набір артефактів, що дозволяють виявляти ознаки фішингу, шкідливої інфраструктури та аномалій конфігурації. Обґрунтовано доцільність мультиагентних систем у кібербезпеці для збору різнорідних dns-даних та їх подальшого аналізу. Проведено огляд поширених сервісів і інструментів доменного аналізу та визначено типові обмеження фрагментованих рішень, які актуалізують інтегрований мультиагентний підхід.

Ключові слова: доменна інформація, збір даних, DNS, OSINT, мультиагентні системи, штучний інтелект.

Abstract

The paper examines domain-related information (WHOIS, DNS, SSL/TLS, and open-source intelligence data) as a set of artifacts that enable the identification of phishing indicators, malicious infrastructure, and configuration anomalies. The relevance of multi-agent systems in cybersecurity is substantiated, particularly for collecting heterogeneous DNS-related data and performing their subsequent analysis. A review of widely used domain analysis services and tools is presented, and typical limitations of fragmented solutions are identified, highlighting the need for an integrated multi-agent approach..

Keywords: domain information, data collection, DNS, OSINT, multi-agent systems, artificial intelligence.

Вступ

У сучасних умовах доменні імена є не лише частиною інтернет-інфраструктури, але й ключовими артефактами кіберзагроз: вони масово застосовуються у фішингових кампаніях, C2-інфраструктурі, розповсюдженні шкідливого ПЗ, а також у схемах швидкого “перекочування” інфраструктури між хостингами/провайдерами. Практична проблема полягає у тому, що ручний доменний OSINT є фрагментованим (кілька сервісів/вкладок/утиліт), залежить від квот API та потребує істотної експертної інтерпретації результатів. У роботі вирішується задача автоматизації комплексного аналізу домену шляхом інтеграції різнорідних джерел у єдине рішення з автоматизованою інтерпретацією.

Результати дослідження

Доменна інформація трактується як сукупність реєстраційних, мережевих і криптографічних метаданих, що описують доменне ім’я та пов’язану інфраструктуру (делегування, DNS, IP, SSL/TLS, репутаційні ознаки) [1]. Її цінність у кібербезпеці полягає в тому, що вона дозволяє швидко сформувати технічний профіль ресурсу (вік домену, характер хостингу/провайдера, захист пошти, валідність TLS) і виконувати pivoting-кореляцію доменів за спільними атрибутами (NS/MX/IP/сертифікати/ASN тощо). Це критично як для CTI (встановлення зв’язків атакуючої інфраструктури), так і для DFIR (фіксація цифрових слідів та повторного використання елементів інфраструктури).

Комплексний аналіз домену потребує кореляції кількох класів даних:

- WHOIS – технічні реєстраційні метадані (реєстратор, дати, статуси, NS), які корисні для оцінки ризику, зокрема при аналізі “свіжих” доменів або нетипових статусів;
- DNS (A/AAAA/NS/MX/CNAME/TXT/CAA) – відображає маршрутизацію трафіку, поштову інфраструктуру та політики автентифікації листів;
- SSL/TLS – криптографічні параметри та перелік доменних імен у SAN;
- CT-журнали / crt.sh – пасивний пошук субдоменів за історією сертифікатів;
- Shodan/Censys – інфраструктурні ознаки (порти, банери, сервіси), але з типовими “сліпими зонами” для доменів за CDN/хмарою;
- VirusTotal – репутаційні індикатори, категорії та сигнали вендорів;

— контент вебсторінки – поведінкові/семантичні ознаки (ключові слова, структура, метадані), що особливо важливо для фішингу навіть за наявності валідного TLS.

Оскільки джерела доменної розвідки є різнорідними та мають різні затримки й обмеження доступу, доцільною є централізовано-координована мультиагентна модель, в якій спеціалізовані агенти виконують незалежні запити паралельно, а координатор відповідає за синхронізацію, тайм-ліміти, уніфікацію форматів і агрегацію результатів. Така організація підтримує принципи єдиної відповідальності, слабкої зв'язаності та масштабування через паралельні запуски агентів при пакетному аналізі доменів[2].

Аналіз поширених інструментів і сервісів доменної розвідки (наприклад, Maltego, SecurityTrails, Censys, Shodan, VirusTotal) показує, що вони ефективні в межах окремих класів джерел (пасивні DNS-дані, інфраструктурні ознаки, сертифікати, репутаційні сигнали), однак практичне застосування часто супроводжується фрагментованістю процесу. Аналітику доводиться виконувати переходи між різними платформами, узгоджувати результати вручну, враховувати квоти безкоштовних планів і відмінності форматів даних. Крім того, більшість рішень не забезпечує єдиної узгодженої моделі результатів і не надає інтегральної оцінки ризику домену з автоматизованою класифікацією, залишаючи фінальну інтерпретацію на експерта. У таких умовах мультиагентний підхід є більш придатним для задач швидкого комплексного аналізу. Різні джерела опитуються паралельно, дані нормалізуються в єдину структуру та можуть бути використані для автоматичного скорингу, кореляції і формування підсумкового технічного звіту.

Висновки

Доменна інформація є ключовим джерелом технічних і репутаційних ознак для CTI та DFIR, оскільки дозволяє швидко формувати профіль ресурсу й виявляти приховані зв'язки інфраструктури через кореляцію реєстраційних, DNS-, криптографічних та інфраструктурних артефактів. Комплексний доменний OSINT потребує одночасного залучення кількох класів джерел (WHOIS, DNS, SSL/TLS, CT-журнали, інтернет-сканери, репутаційні платформи та контент-аналіз), оскільки кожне з них має власні обмеження та "сліпі зони", а достовірний висновок формується лише на основі їх поєднання.

Аналіз існуючих інструментів показує, що попри високу корисність окремих сервісів, практичний процес часто залишається фрагментованим і залежним від ручної інтерпретації, квот та неоднорідності форматів даних. Тому більш ефективним для задач швидкого й відтворюваного аналізу є інтегрований підхід із централізовано-координованою мультиагентною моделлю, де паралельне опитування джерел, нормалізація результатів і агрегація в єдину структуру даних створюють основу для автоматизованого скорингу, кореляції артефактів та формування узгодженого технічного звіту.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Domain Names: Understanding the Basics. URL: https://www.rdj.ie/insights/domain-names-understanding-the-basics (date of access: 08.09.2025).
2. Wang L., Ma C., Feng X. et al. A survey on large language model based autonomous agents. Frontiers of Computer Science, 2024, Vol. 18, Article 186345. DOI: 10.1007/s11704-024-40231-1 (дата звернення: 16.12.2025).

Куперштейн Леонід Михайлович — к. т. н., доцент кафедри захисту інформації, Вінницький національний технічний університет, м. Вінниця, email: kupershtein@vntu.edu.ua

Залепа Олександр Вячеславович — студент групи 1БС-20б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, email: sashasun2002@gmail.com

Kupershtein Leonid — PhD (eng), associated professor of information protection department, Vinnytsia National Technical University, Vinnytsia, email: kupershtein@vntu.edu.ua.

Oleksandr Zalepa — student of group 1БС-20б, Faculty of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, email: sashasun2002@gmail.com