УДК 004.056.5:004.421.4

Концепція виявлення потенційно фішингових доменів на основі моніторингу DNS-активності

Андрій Горбатюк[0009-0003-6478-3163]
Роман Киричок[0000-0002-9919-9691]

Київський столичний університет імені Бориса Грінченка, Україна

Анотація. У публікації представлено підхід до виявлення фішингових доменів на основі аналізу структурних та поведінкових характеристик DNS-трафіку у користувацькому середовищі браузера. Розроблено математичну модель інтегральної оцінки ризику, що поєднує метрики швидкості запитів, ентропії, репутації та поведінкових патернів із механізмом адаптивного калібрування вагових коефіцієнтів. Запропоновано архітектуру браузерного розширення для моніторингу DNS-запитів у реальному часі, яка забезпечує локальне виявлення фішингових доменів без залучення серверної інфраструктури.
Ключові слова: DNS-трафік, DNS-моніторинг, аналіз трафіку, поведінковий аналіз, браузерне розширення.

The Concept of Detecting Potentially Phishing Domains based on Monitoring DNS Activity

Andriy Horbatiuk [0009-0003-6478-3163]
Roman Kyrychok [0000-0002-9919-9691]

Borys Grinchenko Kyiv Metropolitan University, Ukraine

Abstract. The publication presents an approach to detecting phishing domains based on the analysis of structural and behavioral characteristics of DNS traffic in the user's browser environment. A mathematical model for integrated risk assessment has been developed, combining metrics of query speed, entropy, reputation, and behavioral patterns with a mechanism for adaptive calibration of weight coefficients. The architecture of a browser extension for real-time monitoring of DNS requests is proposed, which provides local detection of phishing domains without involving server infrastructure.
Keywords: DNS traffic, DNS monitoring, traffic analysis, behavioral analysis, browser extension.

Вступ

Система доменних імен (Domain Name System, DNS) є критичною складовою інфраструктурою Інтернету, яка забезпечує перетворення доменних імен на IP-адреси. Проте DNS-трафік дедалі частіше використовується зловмисниками для реалізації різноманітних кіберзагроз, зокрема фішингових атак, розповсюдження шкідливого програмного забезпечення, створення ботнетів та ексфільтрації даних. Традиційні методи захисту на мережевому рівні дозволяють нейтралізувати лише частину таких загроз і не забезпечують належного рівня захисту кінцевих користувачів, особливо в умовах використання зашифрованих протоколів обміну (DNS over HTTPS, DNS over TLS).

Більшість сучасних підходів до моніторингу DNS-трафіку орієнтовані на серверний або мережевий рівень аналізу, що обмежує можливості виявлення аномалій безпосередньо в користувацькому середовищі. У зв'язку з цим набуває особливої актуальності розроблення інструментів моніторингу на рівні браузера, які здатні в реальному часі інформувати користувача про потенційно небезпечну DNS-активність.

Метою даного дослідження є підвищення ефективності виявлення фішингових доменів завдяки своєчасному детектуванню аномальних доменних запитів на основі аналізу структурних і поведінкових характеристик DNS-трафіку у користувацькому середовищі.

Наукова новизна дослідження полягає у розробці нового концептуального підходу до виявлення фішингових доменів у режимі реального часу, реалізованого на рівні браузера. Зокрема основними науковими внесками є:

- розроблення математичної моделі для ризик-орієнтованого оцінювання DNS-запитів;
- формалізація процесу вагового агрегування метрик ризику із застосуванням нормалізації часткових показників та збалансованого розподілу їх впливу в інтегральній оцінці;
- запровадження механізму адаптивного калібрування системи оцінки ризику із градієнтним оновленням вагових параметрів, що забезпечує динамічне самонавчання та підвищення чутливості моделі до реальних патернів DNS-активності;
- уточнення критеріїв класифікації рівнів ризику та побудова гнучкої шкали оцінювання, яка дозволяє здійснювати реалістичну інтерпретацію рівнів загрози у режимі реального часу.

На основі розробленої математичної моделі запропоновано архітектуру браузерного розширення, що визначає можливі принципи інтеграції розроблених механізмів оцінювання ймовірності фішингової активності домену та слугуватиме базою для подальших експериментальних досліджень.

Методи та моделі

Концептуальна модель ризик-орієнтованого оцінювання DNS-запитів

Запропонована модель оцінювання ризику спрямована на визначення ймовірності фішингової активності домену на основі комплексного аналізу характеристик DNS-трафіку у користувацькому середовищі.

Вона поєднує декілька незалежних показників, кожен з яких відображає окремий аспект потенційно небезпечної поведінки домену, що дозволяє підвищити точність класифікації у порівнянні з однофакторними підходами.

Модель базується на чотирьох групах часткових метрик:

- метрика швидкості запитів, M_{rate} — характеризує інтенсивність звернень до домену за певний часовий інтервал;
- метрика ентропії доменного імені, $M_{entropy}$ — оцінює ступінь випадковості символів і виявляє DGA-домени;
- метрика репутації домену, $M_{reputation}$ — відображає зовнішній рівень довіри за даними чорних списків, віком домену та статусом SSL/TLS-сертифіката;
- метрика поведінкових патернів, $M_{behavior}$ — аналізує динаміку звернень користувача до домену, виявляючи часові та частотні аномалії.

Кожна з цих метрик є нормалізованою в діапазоні [0,1], що забезпечує можливість подальшої інтеграції їх у єдину функцію ризику.

Комбінування метрик здійснюється за допомогою зваженого агрегування, де вагові коефіцієнти відображають відносну значущість кожної складової.

Результатом моделі є числове значення $Risk(d,t)$, яке описує рівень ризику для домену d у момент часу t та використовується для класифікації рівня загрози. Концептуально модель можна подати у вигляді системи:

$$\mu = \{M_{rate}, M_{entropy}, M_{reputation}, M_{behavior}\} \rightarrow Risk(d,t),$$

де μ — множина часткових метрик, а $Risk(d,t)$ — результат їх агрегування у підсумкову оцінку.

Таким чином, модель забезпечує багатофакторний аналіз DNS-активності, який поєднує структурні, поведінкові та репутаційні характеристики доменів і слугує основою для подальшого формального опису часткових метрик та інтегральної математичної моделі у наступних підпунктах.

Метрика швидкості DNS-запитів

Перша метрика оцінює інтенсивність DNS-запитів до конкретного домену за визначений часовий інтервал. Аномально висока швидкість може вказувати на DGA (Domain Generation Algorithm) активність або DNS-тунелювання.

Формула розрахунку метрики швидкості:

$$M_{rate}(d,t) = \frac{N_{req}(d,t)}{\Delta t} \cdot k_{norm} \quad (1)$$

$M_{rate}(d,t)$ — значення метрики швидкості для домену у момент часу; $N_{req}(d,t)$ — кількість запитів до домену за часовий інтервал; Δt — часове вікно спостереження (наприклад, 60 секунд); k_{norm} — нормалізаційний коефіцієнт для приведення значення до діапазону [0,1].

Нормалізація метрики:

$$k_{norm} = \frac{1}{R_{max}} \quad (2)$$

де R_{max} — максимально допустима швидкість запитів (порогове значення, наприклад, 100 запитів/хвилину).

Нормалізоване значення метрики:

$$M_{rate\_norm}(\mathrm{d},\mathrm{t}) = min(1, \frac{N_{req}(d,t)}{\Delta t \cdot R_{max}}) \quad (3)$$

Метрика ентропії доменного імені

Ентропія доменного імені використовується для виявлення доменів, згенерованих алгоритмічно (DGA). Такі домени зазвичай мають вищу ентропію порівняно з легітимними доменами.

Формула обчислення ентропії Шеннона:

$$H(d) = -\sum_{i=1}^{n} p_i \log_2(p_i) \quad (4)$$

де $H(d)$ — ентропія доменного імені d ; n — кількість унікальних символів у доменному імені; p_i — ймовірність появи і-го символу в доменному імені.

Розрахунок ймовірності символу:

$$p_i = \frac{f_i}{L} \quad (5)$$

де f_i — частота появи і-го символу в доменному імені; L — загальна довжина доменного імені (без урахування TLD)

Нормалізована метрика ентропії:

$$M_{entropy}(d) = \frac{H(d)}{H_{max}} \quad (6)$$

де — максимально можлива ентропія для алфавіту Σ (для латинського алфавіту та цифр |Σ| = 36, тому $H_{max} \approx 5.17$.

Метрика репутації домену

Репутаційна метрика базується на даних із зовнішніх джерел - списків відомих шкідливих доменів, фішингових баз даних, репутаційних сервісів.

Формула репутаційної метрики:

$$M_{reputation}(d) = w_1 \cdot I_{blacklist}(d) + w_2 \cdot S_{age}(d) + w_3 \cdot S_{cert}(d) \quad (7)$$

Де $I_{blacklist}(d)$ — індикатор присутності домену в чорних списках (1 — присутній, 0 — відсутній); $S_{age}(d)$ — оцінка віку домену (нормалізована в діапазоні [0,1], де 1 — дуже молодий домен); $S_{cert}(d)$ — оцінка сертифіката (1 — відсутній/недійсний, 0 — дійсний); $w_1, w_2, w_3,$ — вагові коефіцієнти ($w_1 + w_2 + w_3 = 1$). При цьому для вагових коефіцієнтів рекомендуються наступні значення:

- $w_1 = 0.6$ (найбільша вага для присутності в чорних списках);
- $w_2 = 0.25$ (середня вага для віку домену);
- $w_3 = 0.15$(менша вага для статусу сертифіката)

Оцінка віку домену:

$$S_{age}(\mathrm{d}) = e^{-\lambda \cdot age(d)} \quad (8)$$

де $age(d)$ — вік домену в днях; λ — параметр швидкості зниження підозрілості (рекомендоване значення $\lambda = 0.01$).

Ця експоненційна функція забезпечує, що домени віком менше 30 днів отримують високі оцінки підозрілості (близько 0.74 для домену віком 30 днів), тоді як домени старше 6 місяців отримують низькі оцінки (близько 0.05).

Метрика поведінкових патернів

Поведінкова метрика аналізує патерни доступу до доменів, виявляючи аномальну активність. Формула поведінкової метрики:

$$M_{behavior} = \alpha \cdot D_{temporal}(d,t) + \beta \cdot D_{frequency}(d,t) + \gamma \cdot D_{pattern}(d,t) \quad (9)$$

де $D_{temporal}(d,t)$ — відхилення часового патерну (аномалії у часі доступу); $D_{frequency}(d,t)$ — відхилення частотного патерну (аномалії у періодичності); $D_{pattern}(d,t)$ (аномалії у послідовності запитів); $\alpha + \beta + \gamma = 1$ — вагові коефіцієнти.

$$D_{temporal}(d,t) = \frac{|t - \mu_t(d)|}{\sigma_t(d)} \quad (10)$$

де $\mu_t(d)$ — середній час доступу до домена $\sigma_t(d)$ — стандартне відхилення часу доступу; t — поточний час доступу.

Розрахунок частотного відхилення:

$$D_{frequency}(d,t) = \left|\frac{f_{current}(d,t) - f_{avg}(d)}{f_{avg}(d)}\right| \quad (11)$$

де $f_{current}(d,t)$ — поточна частота запитів до домену d; $f_{avg}(d)$ — середня частота запитів за історичний період.

Інтегральна модель оцінки ризику

Фінальна оцінка ризикованості домену базується на зваженій комбінації всіх метрик:

(12)

$$Risk(d,t) = w_r \cdot M_{rate\_norm}(d,t) + w_e \cdot M_{entropy}(d) + w_{rep} \cdot M_{reputation}(d) + w_b \cdot M_{behavior}(d,t)$$

де $Risk(d,t)$ — інтегральна оцінка ризику для домену d у момент часу t (діапазон [0,1]); w_r, w_e, w_{rep}, w_b — вагові коефіцієнти для відповідних метрик. Умова нормалізації: $w_r + w_e + w_{rep} + w_b = 1$.

Рекомендована конфігурація вагових коефіцієнтів:

$$\begin{cases} w_r = 0.15 (\text{швидкість запитів}) \\ w_e = 0.25 (\text{ентропія}) \\ w_{rep} = 0.40 (\text{репутація}) \\ w_b = 0.20 (\text{поведінка}) \end{cases} \quad (13)$$

Найбільша вага надається репутаційним даним, оскільки вони базуються на перевірених джерелах загроз. Ентропія домену має другу за важливістю вагу, оскільки ефективно виявляє DGA-домени.

Класифікація рівнів ризику:

$$Level(d,t) = \begin{cases} LOW, \text{ якщо } Risk(d,t) < 0.3 \\ MEDIUM, \text{ якщо } 0.3 \le Risk(d,t) < 0.6 \\ HIGH, \text{ якщо } 0.6 \le Risk(d,t) < 0.8 \\ CRITICAL, \text{ якщо } Risk(d,t) > 0.8 \end{cases} \quad (14)$$

Таке розмежування рівнів ризику дозволяє реалізувати гнучкий механізм реагування на потенційні загрози.

Наприклад, домени з високим ризиком можуть блокуватись або супроводжуватись попередженнями користувачу, тоді як низькоризикові запити лише логуються для подальшого аналізу.

Адаптивне калібрування системи

Для підвищення точності системи реалізовано механізм адаптивного налаштування вагових коефіцієнтів на основі зворотного зв'язку від користувача:

$$w_i^{(n+1)} = w_i^{(n)} + \eta \cdot \delta_i \quad (15)$$

де $w_i^{(n+1)}$ — поточне значення і-го вагового коефіцієнта на ітерації n; η — швидкість навчання (рекомендоване значення $\eta = 0.1$; δ_i — градієнт помилки для і-ї метрики.

Розрахунок градієнта помилки:

$$\delta_i = \frac{\partial E}{\partial w_i} = (Risk_{predicted} - Risk_{actual}) \cdot M_i \quad (16)$$

де E — функція помилки; $Risk_{predicted}$ — передбачений рівень ризику; $Risk_{actual}$ — фактичний рівень ризику (на основі дій користувача); M_i — значення і-ї метрики.

Таким чином, розроблена математична модель поєднує чотири взаємодоповнюючі метрики у єдину модель інтегральної оцінки ризику, доповнену механізмом адаптивного калібрування вагових коефіцієнтів.

Результати

Архітектура браузерного розширення

Архітектура розширення DNS Sentinel базується на модульному підході та включає наступні основні компоненти:

1. Background Service Worker — основний компонент, який виконується у фоновому режимі та координує роботу всієї системи.
2. Content Script — скрипт, що впроваджується на веб-сторінки для збору даних про мережеву активність
3. Storage Layer — рівень зберігання даних, що забезпечує персистентність інформації та оптимізацію продуктивності. Включає наступні підкомпоненти:
   a. Domain Statistics (DS) — сховище статистичних даних про домени: кількість запитів, часові мітки звернень, історія обчислених метрик. Використовується для розрахунку метрик швидкості та поведінки.
   b. Configuration Store (CF) — сховище налаштувань системи: вагові коефіцієнти метрик ($w_r, w_e, w_{rep} w_b$,), порогові значення класифікації ризику, параметри адаптивного навчання. Дозволяє користувачу налаштовувати чутливість системи.
   c. History Log (HL) — журнал подій із записами всіх виявлених загроз, включаючи повну інформацію про метрики, час виявлення та дії користувача. Використовується для адаптивного калібрування системи (формули 15-16) та аудиту безпеки.
4. Analysis engine — модуль аналізу, що реалізує математичні моделі оцінки ризиків
5. UI Components — інтерфейсні компоненти для відображення інформації користувачу

Алгоритм функціонування розширення

Алгоритм функціонування розширення передбачає такі етапи:

1. Перехоплення веб-запиту та виділення домену з URL. Перехоплює веб-запит, виділяє доменне ім'я з URL та фіксує часову мітку. Ініціалізуються структури даних для збереження метрик та оцінки ризику.
2. Обчислення метрик. Паралельно розраховуються чотири метрики: швидкість запитів (частота звернень), ентропія (складність доменного імені), репутація (перевірка чорних списків та SSL) та поведінка (аналіз історичних відхилень).
3. Розрахунок інтегрального ризику.Отримані значення метрик комбінуються у фінальну оцінку ризику з використанням зважених коефіцієнтів відповідно до (12). Виконується класифікація рівня загрози згідно з пороговими значеннями (14).

4. Класифікація та реагування. Залежно від рівня ризику (CRITICAL, HIGH, MEDIUM, LOW) система блокує домен, попереджає користувача або просто логує подію. Критичні загрози (≥0.8) блокуються негайно.
5. Оновлення статистики. Система оновлює статистику домену та глобальні показники, при наявності зворотного зв'язку адаптує вагові коефіцієнти. Весь цикл обробки займає 5-50 мілісекунд.Обговорення

Обговорення

Запропонований підхід реалізує аналіз DNS-трафіку безпосередньо в браузері, що усуває залежність від серверних систем і забезпечує оперативне реагування на шкідливі запити.

Основні метрики виявлення фішингу:

- Швидкість запитів — виявляє сплески активності (легітимні користувачі: 10-15 запитів/хв, фішинг: 50-100+).
- Ентропія — ідентифікує DGA-домени (фішингові "paypa1-secure-verify-2x8k9m.com" мають H ≈ 3.8, легітимні "paypal.com" — H ≈ 2.4).
- Репутація — перевіряє anti-phishing бази, вік домену (70% фішингових молодші 30 днів) та SSL-сертифікати.
- Поведінка — аналізує аномальний час доступу та нетипові паттерни переходів.

Інтегральна оцінка використовує зважені коефіцієнти з пріоритетом репутаційної складової для мінімізації хибнопозитивних спрацьовувань.

Переваги:

- Моніторинг працює навіть при зашифрованих протоколах (DoH/DoT).
- Незалежність від мережевої інфраструктури.
- Багатофакторний аналіз підвищує стійкість.
- Адаптивне навчання персоналізує систему.

Обмеження:

- Складність виявлення високоякісного цільового фішингу з легітимними хостингами.
- Проблеми з гомогліфними атаками та IDN-доменами.
- Хибнопозитивні спрацьовування на легітимних короткотермінових піддоменах.
- Залежність від актуальності зовнішніх баз (нові домени діють 2-6 годин до реєстрації).
- API-обмеження (Google Safe Browsing: 10,000 запитів/день).

Поєднання формалізованих метрик і адаптивного калібрування створює гнучку основу для клієнтських систем захисту від фішингу.

Висновки

У роботі запропоновано концептуальну модель оцінювання ймовірності фішингової активності доменів на основі аналізу структурних і поведінкових характеристик DNS-трафіку у користувацькому середовищі.

Розроблено математичну модель інтегральної оцінки ризику, що поєднує чотири метрики: швидкість запитів, ентропію, репутацію та поведінкові патерни у єдину зважену функцію ризику. Передбачено механізм адаптивного калібрування вагових коефіцієнтів, який забезпечує самонавчання системи на основі користувацького зворотного зв’язку.

На основі запропонованої моделі сформовано архітектуру браузерного розширення, що дозволяє здійснювати моніторинг DNS-запитів у режимі реального часу та оперативно попереджати користувача про потенційно небезпечні ресурси.

Запропонований підхід може бути використаний як основа для створення нових поколінь клієнтських систем аналізу DNS-трафіку, орієнтованих на підвищення рівня захисту від фішингових загроз.

Подальші дослідження доцільно спрямувати на:

- інтеграцію методів машинного навчання для автоматичного визначення оптимальних вагових коефіцієнтів;
- розробку метрики візуальної подібності символів для виявлення гомогліфних атак;
- впровадження принципів federated learning для розподіленого обміну даними між користувачами без розкриття персональної інформації;
- адаптацію моделі для багатомовних доменів і регіональних фішингових кампаній;
- експериментальну перевірку моделі на реальних наборах DNS-трафіку з подальшим оцінюванням її точності та продуктивності.

Джерела

1. Романюк М. І., Власюк Г. Г. (2018). Основи теорії інформації та кодування: лабораторний практикум. Київ: КПІ ім. Ігоря Сікорського. https://ela.kpi.ua/items/1cc27217-226a-484b-b8ac-9497fffe37b5

2. Гуськова В. Г., Бідюк П. І., Гасанов А. С. (2022). Ймовірнісно-статистичні методи моделювання і прогнозування. К.: Видавництво НПУ ім. М.П. Драгоманова. https://enpuir.udu.edu.ua/entities/publication/2a02994e-ce04-4178-9637-6a191525cccd

3. Булдигін В. В., Алєксєєва І. В., Гайдей В. О., Диховичний О. О., Коновалова Н. Р., Федорова Л. Б. (2011). Лінійна алгебра та аналітична геометрія: Навчальний посібник. Київ: ТВіМС. 224 с. https://matan.kpi.ua/public/files/Posibnyk LA+AG.pdf

4. Дубовик В. П., Юрик І. І. (2005). Вища математика: навчальний посібник. Київ: А.С.К. 648 с. https://grigorieva-n-a.at.ua/Liter/1.pdf

5. Вітлінський В. В., Терещенко Т. О., Савіна С. С. (2016). Економіко-математичні методи та моделі: оптимізація: навчальний посібник. Київ: КНЕУ. 303 с. https://kneu.edu.ua/get_file/7762/%D0%95%D0%BA%D0%BE%D0%BD%D0%BE%D0%BC%D1%96%D0%BA%D0%BE-%D0%BC%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D1%87%D0%BD%D1%96%20%D0%BC%D0%B5%D1%82%D0%BE%D0%B4%D0%B8%20%D1%96%20%D0%BC%D0%BE%D0%B4%D0%B5%D0%BB%D1%96%20%D0%BE%D0%BF%D1%82%D0%B8%D0%BC%D1%96%D0%B7%D0%B0%D1%86%D1%96%D1%8F.pdf

6. Шевченко, С., Жданова, Ю., Спасітєлєва, С., Мазур, Н., Складанний, П., & Негоденко, В. (2024). Математичні методи в кібербезпеці: кластерний аналіз та його застосування в інформаційній та кібернетичній безпеці. Кібербезпека: освіта, наука, техніка, 3(23), 258–273. https://doi.org/10.28925/2663-4023.2024.23.258273

7. Костюк, Ю. В., Складанний, П. М., Гулак, Г. М., Бебешко, Б. Т., Хорольська, К. В., & Рзаєва, С. Л. (2025). Системи захисту інформації. Київ: Київський столичний університет імені Бориса Грінченка.

8. Гулак, Г. М., Жильцов, О. Б., Киричок, Р. В., Коршун, Н. В., & Складанний, П. М. (2023). Інформаційна та кібернетична безпека підприємства (підручник). Київ: Київський столичний університет імені Бориса Грінченка.