УДК 004.056.57

ПОРІВНЯННЯ МЕТОДІВ ЗАХИСТУ ВІД АТАК WEBSITE FINGERPRINTING

Л. В. Мєтєлєва[1, a]

[1] *Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Фізико-технічний інститут*

Анотація

Атаки за цифровими відбитками сайтів (Website Fingerprinting, WFP) дозволяють локальному пасивному зловмиснику ідентифікувати, який саме веб-ресурс відвідує жертва, базуючись на характеристиках мережевого трафіку. У даній роботі представлений огляд актуальних WFP атак. Зроблено аналіз підходів до захисту від таких атак. Створено прототип нового методу захисту від атак даного типу.

Ключові слова: website fingerprinting, атака за цифровими відбитками сайтів, анонімність, інформаційна безпека

Вступ

Сьогодні важко уявити життя без глобальної мережі Інтернет. З появою Інтернету також з'явилася і проблема захисту анонімності у мережі. Ми розглядаємо анонімність не лише як неможливість сайтів та сервісів визначити реальну IP-адресу користувача, а і як неможливість локального прослуховувача (напр. Інтернет-провайдер, IPS) дізнатися, яку саме веб-сторінку переглядає користувач.

Ранніми вирішеннями цих проблем було використання шифрованого тунелювання (SSL, SSH, VPN). Згодом з'явилися спеціалізовані проксі-системи, такі як JAP (Java Anon Proxy) та Tor (The Onion Router). Останній на сьогодні є найбільш популярним. Tor забезпечує анонімність у мережі шляхом створення випадковим чином маршрутів (circuit), що складаються з трьох вузлів, таким чином для серверу джерелом запиту є вихідний вузол, а вихідний вузол не знає, який з вузлів є вхідним і де саме знаходиться користувач. Додатково трафік користувача шифрується ключами всіх вузлів у пакетах фіксованого розміру.

Користувачами Tor є люди, що хочуть захистити свою приватність та анонімність в Інтернеті. У деяких випадках відсутність цих властивостей є критичною. Наприклад, громадські активісти, журналісти та інформатори використовують Tor для відправлення та доступу до інформації з територій, де трафік може прослуховуватися чи фільтруватися (напр. The Golden Shield Project). Військові та спеціальні служби використовують Tor під час розслідуваннь, розвідки, збору інформації та для комунікації. Tor використовують і інші користувачі, які бажають захистити свою приватність в Інтернеті.

Між користувачем та Tor-мережею безпека Інтернет-трафіку базується на шифруванні. Кількість користувачів величезна, розробники імплементують стійкі алгоритми шифрування, тому логічним способом компрометації анонімності є атака аналізу трафіку. Саме атакою цього типу є WFP.

У випадку WFP атаки зловмисник може знаходитися між жертвою та вхідним вузлом Tor-мережі або він може контролювати скомпрометований вхідний вузол. Атакувальник виокремлює з шифрованого трафіку певні характеристики, наприклад, напрям та кількість пакетів, інтервал між ними тощо. Ці характеристики є унікальними для кожного сайту і аналізуються за допомогою алгоритмів класифікації машинного навчання. Таким чином, зловмисник може дізнатися, яку саме веб-сторінку завантажив браузер жертви.

Отже, використання одного лише Tor чи шифрованого тунелювання недостатньо для забезпечення анонімності, для цього треба вживати додаткових заходів безпеки. Для того щоб створити прототип нового методу захисту, необхідно розглянути найбільш актуальні WFP-атаки, принципи та точність їх роботи, а також проаналізувати відомі методи протидії таким атакам та їх недоліки, зрештою, запропонувати новий метод захисту та описати принцип його роботи.

1. Атаки

Атаки за цифровими відбитками сайтів зводяться до завдання класифікації машинного навчання. Існує два підходи до тестування ефективності атак: закритий та відкритий світ (множини) вебсайтів. Закритий світ складається з множини веб-сторінок, відомих атакувальнику. Жертва може переглядати лише сторінки з цієї множини. Таким чином, метою зловмисника є ідентифікувати, яку саме сторінку з даної множини сторінок переглядає користувач. Завдання перших атак оцінювалися саме в умовах закритого світу. Наприклад, перша атака проти Tor базувалась на наївному баєсовому класифікаторі, тестувалась на множині з 775 сайтів та правильно ідентифікува-

[a] l.meteleva@gmail.com

ла лише 3 % сторінок [1]. Останні атаки проти Tor досягають більше 90 % точності [2, 3, 4, 5].

Відкритий світ складається з множини сторінок, лише частина яких відома зловмиснику. Ця множина поділяється на дві підмножини. Перша – сторінки, які моніторяться зловмисником, наприклад, такі, що підлягають цензурі. Решта сайтів формує другу підмножину і лише частково відома зловмиснику. Даний тип оцінювання атаки є більш реалістичним. Перша атака, що була оцінена таким чином, базувалась на SVM класифікаторі зі 100 характеристиками, та розмір множини – 1000 сторінок. Результати тестування показали 97 % вірних та 0,005 % помилкових спрацьовувань [6].

У 2014 році Ван та ін. [5] опублікували атаку, що базувалась на методі найближчих k сусідів (k-nearest neighbors). Множина характеристик включала в себе майже 4000 ознак, 3000 з яких описували унікальні розміри пакетів і є безкорисними у випадку комунікації через Tor (тому що розміри пакетів там фіксовані). Решта характеристик описувала загальний розмір, час та напрямок передачі пакетів. Щоб покращити точність роботи класифікатора при такій кількості характеристик, автори проаналізували їх та надали їм різну вагу, відповідно до їх значимості. Щоб класифікувати новий зразок трафіку класифікатор враховує класи найближчих k сусідів. Якщо всі k класи є однаковими – робиться припущення, інакше алгоритм повертає порожню відповідь. Автори тестували атаку зі 100 сторінок, що моніторяться, та 5000 сторінок, що не моніторяться. Алгоритм показав 85 % вірних (true positive rate, TPR) та 0,6 % помилкових спрацьовувань (false positive rate, FPR). Головним недоліком цієї атаки була величезна кількість характеристик, що призводить до перенавантаження та вимагає значних ресурсів для проведення атаки.

Два роки по тому Панченко та ін. [4] створили нову SVM-атаку – CUMUL. Перед збором характеристик трафік попередньо фільтрувався від недопустимих значень. Для записів трафіку, що залишились, підраховувалась кумулятивна сума розмірів пакетів, де розміри вихідних пакетів додаються з від'ємним знаком, а вхідні – зі знаком «плюс». Решта характеристик описує розмір, напрямок та порядок пакетів. Характеристики збиралися на прикладному, транспортному рівнях та з Tor-пакетів. Загалом нараховувалося 104 характеристики. Експерименти з такими ж умовами, як у попередній роботі, показали, що точність нової атаки на вище, а ресурсоємність значно знизилась. Під час тестування робилось припущення, що у користувача відкрито більше одної вкладки одночасно, як і у більшості реальних користувачів. Результати тестування показали, що декілька відкритих вкладок у браузері користувача значно знижують точність класифікації. Так, точність розпізнавання основних сторінок сильно варіюється та падає від 80 % до 5 % при зростанні розміру можливих фонових вкладок з 1000 до 100000. Точність стабілізується на рівні близько 80 % до 110000 фонових сторінок. Відповідно до таких результатів, автори зробили висновок, що WFP атаки можуть давати далеко не такі хороші результати в реальному світі, як в експериментальних умовах, тому що вони не є масштабованими.

Табл. 1. Точність роботи класифікаторів

Класифікатор	Кількість характеристик	TPR	FPR
k-NN	3736	85 %	0,6 %
CUMUL	104	97 %	2 %
k-fingerprinting	≈ 600	88 %	0,5 %

Остання атака k-fingerprinting була опублікована Хаєсом та ін.[3] на основі алгоритму Random forest. Їх «випадковий ліс» складався з k дерев, кожне з який мало власний набір із 150-200 характеристик. Відповіді всіх дерев формували вектор. Кінцеве рішення приймалось на основі підрахунку відстані Хеммінга між отриманим вектором та векторами з навчальної бази. Для перевірки точності роботи алгоритму вони використали той же набір сайтів, що і Ван та ін. [5]. При $k = 3$ алгоритм мав 88 % вірних та 0,5 % помилкових спрацьовувань. Додатково вони провели тести на значно більшому наборі сайтів – понад 100000 з різними значеннями параметру k. Дані експерименти показали, що на точність роботи алгоритму більше впливає параметр k, ніж розмір відкритого світу. Тобто дана атака є масштабованою і може бути застосована в реальному світі.

У таблиці 1 підсумовано інформацію про всі розглянути класифікатори.

2. Захист

У 2006 Ліберatоре і Левін [7] запропонували доповнювати всі пакети зайвими байтами до фіксованого розміру пакету. Даний підхід має значні перевантаження пропускної здатності (145 %) та маскує лише одну з характеристик трафіку – розмір пакетів, яка і так є фіксованою в Tor.

Цю ідею пізніше у 2009 Райт та ін. [8] вдосконалили і доповнювали байти лише до пакетів з унікальними розмірами таким чином, щоб даний зразок трафіку виглядав для зловмисника як зразок завантаження іншої сторінки. Ці зміни значно знизили перевантаження пропускної здатності (bandwidth overhead, 39 %), але цей підхід так само маскує лише одну характеристику, більше того, він вимагає попереднього створення бази зі зразками трафіку завантаження сторінок.

Разом з першою атакою на Tor Панченко та ін. [6] запропонували і захист на програмному рівні. Кожен раз, коли користувач завантажував сторінку, додаток до веб-браузера паралельно завантажував випадкову веб-сторінку. Такий захист значно знизив точність класифікатора з даної роботи з 54 % до 3 %, але, в той же час, і зменшилась пропускна здатність на близько 85 %.

У тому ж році один з розробників Tor, Майк Перрі [9] додав до Tor Browser Bundle можливість завантажувати у випадковому порядку запити, що

відсилаються за допомогою HTTP Pipeline (відправлення декількох запитів до сервера одночасно, не очікуючи відповідь на попередній запит від сервера). Головними недолікам такого підходу є те, що HTTP Pipeline не підтримується всіма серверами і те, що нові класифікатори використовують такий набір характеристик, що робить даний захист безкорисним.

Ло і ін. [10] запропонували зміни до HTTP і TCP заголовків, щоб перетворювати трафік. За допомогою заголовків HTTP Range, HTTP Pipeline, а також максимального розміру сегмента і розміру вікна прийому у TCP автори змінювали розміри пакетів та час між їх відправленням/отриманням. Цей підхід має недоліки попереднього, а також незначне перевантаження пропускної здатності (5 %).

Даєр та ін. [11] у 2012 опублікували BuFLO (Buffered Fixed-Length Obfuscator) для SSH тунелювання. BuFLO намагається видалити всю інформацію шляхом передачі пакетів фіксованого розміру у фіксованому часовому інтервалі та фіксованого розміру послідовностей пакетів. Якщо дані для передачі менше зазначеного буферу, BuFLO заповнює його зайвими байтами. Якщо завантаження веб-сторінки ще не закінчилося, тоді BuFLO буде продовжувати посилати пакети фіксованого розміру із випадковими байтами протягом ще 10 секунд. Якщо кількість переданих пакетів менше фіксованого значення - відправлялися додаткові пакети. Такий захист знижував точність тодішніх класифікаторів (з $97,5$ % до $27,3$ % для [6]), але майже не впливав на роботу новіших атак, більше того, залежно від обраних фіксованих параметрів, перевантаження пропускної здатності може бути 94 %-419 %. Також визначення, коли саме закінчується завантаження сторінки, є не тривіальною задачею.

Наступні чотири методи створені з метою зменшення значного зниження пропускної здатності BuFLO. Деякі для цього доповнювали дані не до фіксованого числа, а до найближчої степені двійки [12]. У іншій роботі автори [13] доповнювали дані до числа, кратного певному параметру. Останні ідеї знизили перевикористання пропускної здатності до $180 - 200$ %. Ще два підходи [5, 14] мали схожий принцип роботи. Вони попередньо аналізували зразки трафіку та поділяли їх на класи анонімності, відповідно до розміру та кількості пакетів. Під час завантаження користувачем сторінки вони доповнювали трафік до найближчого класу, адже так зловмисник міг побачити лише до якого класу належить сторінка. Це знизило розмір додаткового навантаження на пропускну здатність (біля 60 %), але з'явилась нова проблема визначення розміру та параметрів класу анонімності, що потребує попередніх обрахунків та оновлення.

У 2015 році Ван та ін. [15] опублікували ідею маскування трафіку шляхом переведення роботи браузера у напівдуплексний режим. Це означає, що браузер не буде відправляти нових запитів, поки сервер не відповість на всі попередні. Додатково вони доповнювали розмір пакетів та відправляли зайві пакети. Недоліками даного захисту є необхідність модифікувати Tor Browser та збільшення трафіку на 80 %.

Табл. 2. Порівняння методів захисту на стороні клієнта

Захист	Bandwidth overhead	Time overhead	Точність k-NN
[7]	61 %	0 %	60 %
[8]	50 %	0 %	94 %
[10]	$6,4$ %	0 %	96 %
[6]	110 %	25 %	18 %

У 2016 році Хуаресом та ін. [16] була запропонована ідея доповнювати корисні дані сміттям не випадковим чином, а аналізуючи гістограми розподілу пакетів. За допомогою кінцевого автомата вони зробили трафік випадковим для класифікатора, знижуючи його ефективність до простого вгадування. Вадами такого підходу є необхідність побудови відповідних гістограми та їх постійного оновлення, а також додаткове навантаження пропускної здатності на 54 %.

3. Новий підхід до захисту

Аналізуючи відомі підходи до реалізації захисту від атак за цифровими відбитками веб-сайтів, ми можемо побачити, що всі вони мають свої недоліки, що перешкоджають їх широкому використанню у реальних умовах. Одні вимагають змін на серверній стороні чи Tor Browser Bundle, для використання інших користувач повинен встановити та налаштовувати додаткове програмне забезпечення, оновлювати необхідні бази. Важко змусити всіх власників серверів додати захист від WFP атак та постійно оновлювати його. Таким чином, більш реальним є використання сучасної інфраструктури програмного захисту на стороні клієнта, що дозволяє з мінімальними змінами та зусиллями підвищити рівень конфіденційності та анонімності у мережі. У табл. 2 наведено порівняння методів захисту на стороні клієнта під час оцінювання k-NN атаки [5].

Найбільш оптимальною реалізацією захисту від вищезгаданих атак є розширення для браузера. Наприклад, створення розширення з використанням браузерного API WebExtensions. WebExtensions підтримується більшістю нових версій сучасних браузерів. Таким чином, він підходить не лише для Tor Browser Bundle, а і для тих, хто використовує Tor мережу через інші браузери.

Аналізуючи характеристики трафіку, що використовуються новими атаками, вищеопубліковані способи захисту, можна сформувати ідею нового підходу до захисту: додавання випадкового часу затримки перед відправленням запиту до сервера, розділення одного запиту на декілька підзапитів та одночасне відправлення зайвих запитів, що значно зменшить точність роботи класифікаторів в основі цих атак. На рис. 1 зображено діаграму запитів під час використання захисту. До початкового запиту додається

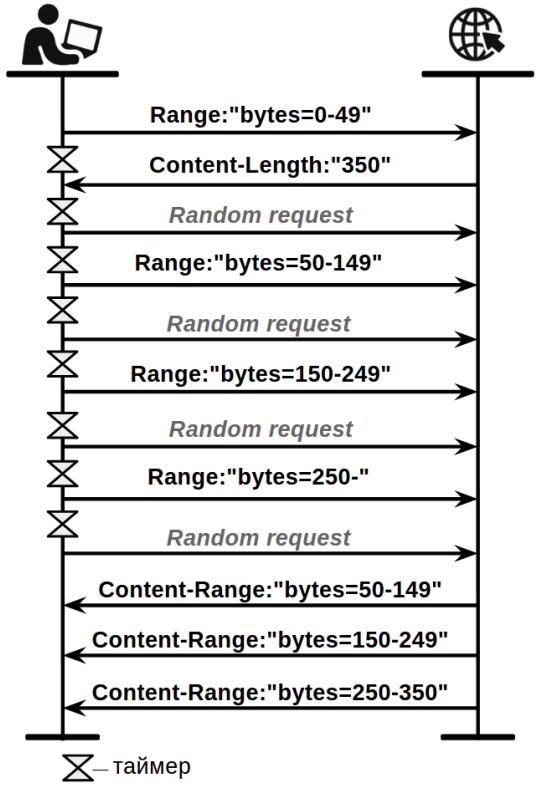

Рис. 1. Діаграма запитів із захистом

заголовок `Range: «bytes=0-49»`, що дає нам можливість дізнатися повний розмір контенту із заголовка відповіді `Content-Length` та отримати перші 50 байтів контенту. Далі, решта байтів умовно ділиться на випадкову кількість часткових запитів, між якими випадковим чином додаються чи не додаються зайві запити та затримки випадкової тривалості.

Висновки

У даній роботі було розглянуто та проаналізовано актуальні атаки за цифровими відбитками сайтів та методи протидії їм. На основі проведеного аналізу було запропоновано ідею нового способу підвищення рівня анонімності у мережі шляхом створення розширення до браузера, що змінює основні характеристики трафіку.

Перелік використаних джерел

1. Herrmann Dominik, Wendolsky Rolf, Federrath Hannes. Website Fingerprinting: Attacking Popular Privacy Enhancing Technologies with the Multinomial Naive-bayes Classifier // 2009 ACM Workshop on Cloud Computing Security. — 2009. — P. 31–42.
2. Touching from a Distance: Website Fingerprinting Attacks and Defenses / Xiang Cai, Xin Cheng Zhang, Brijesh Joshi, Rob Johnson // 2012 ACM Conference on Computer and Communications Security. — 2012. — P. 605–616.
3. Hayes Jamie, Danezis George. k-fingerprinting: A robust scalable website fingerprinting technique. — 2016.
4. Website fingerprinting at internet scale / Andriy Panchenko, Fabian Lanze, Andreas Zinnen et al. // Network & Distributed System Security Symposium. IEEE Computer Society. — 2016.
5. Effective Attacks and Provable Defenses for Website Fingerprinting / Tao Wang, Xiang Cai, Rishab Nithyanand et al. // 23rd USENIX Conference on Security Symposium. — San Diego, CA, 2014. — P. 143–157.
6. Website Fingerprinting in Onion Routing Based Anonymization Networks / Andriy Panchenko, Lukas Niessen, Andreas Zinnen, Thomas Engel // 10th Annual ACM Workshop on Privacy in the Electronic Society. — 2011. — P. 103–114.
7. Liberatore Marc, Levine Brian Neil. Inferring the Source of Encrypted HTTP Connections // 13th ACM Conference on Computer and Communications Security. — 2006. — P. 255–263.
8. Wright Charles V., Coull Scott E., Monrose Fabian. Traffic Morphing: An Efficient Defense Against Statistical Traffic Analysis // 16th Network and Distributed Security Symposium. — IEEE, 2009. — P. 237–250.
9. Perry Mike. Experimental Defense for Website Traffic Fingerprinting. Tor project Blog. — 2011. — Access mode: `https://blog.torproject.org/blog/experimental-defense-website-traffic-fingerprinting` (online; accessed: 2017-03-01).
10. HTTPOS: Sealing Information Leaks with Browser-side Obfuscation of Encrypted Flows / X. Luo, P. Zhou, E. Chan, W. Lee // 18th Annual Network and Distributed Systems Symposium. — 2011.
11. Peek-a-Boo, I Still See You: Why Efficient Traffic Analysis Countermeasures Fail / Kevin P. Dyer, Scott E. Coull, Thomas Ristenpart, Thomas Shrimpton // 2012 IEEE Symposium on Security and Privacy. — 2012. — P. 332–346.
12. Cai Xiang, Nithyanand Rishab, Johnson Rob. CS-BuFLO: A Congestion Sensitive Website Fingerprinting Defense // 13th Workshop on Privacy in the Electronic Society. — Scottsdale, Arizona, USA, 2014. — P. 121–130.
13. A Systematic Approach to Developing and Evaluating Website Fingerprinting Defenses / Xiang Cai, Rishab Nithyanand, Tao Wang et al. // SIGSAC Conference on Computer and Communications Security. — Scottsdale, Arizona, USA, 2014. — P. 227–238.
14. Nithyanand Rishab, Cai Xiang, Johnson Rob. Glove: A Bespoke Website Fingerprinting Defense // 13th Workshop on Privacy in the Electronic Society. — Scottsdale, Arizona, USA, 2014. — P. 131–134.
15. Wang Tao, Goldberg Ian. Walkie-Talkie: An Effective and Efficient Defense against Website Fingerprinting // 21st European Symposium on Research in Computer Security. — 2015.
16. Toward an Efficient Website Fingerprinting Defense / Marc Juarez, Mohsen Imani, Mike Perry et al. // 21st European Symposium on Research in Computer Security. — 2016. — P. 27–46.