УДК 004.77+0004.56

А. В. Борусевич
Л. М. Куперштейн

АНАЛІЗ ФАКТОРІВ У ЗАДАЧАХ ВИЗНАЧЕННЯ ТИПУ ОПЕРАЦІЙНОЇ СИСТЕМИ

Вінницький національний технічний університет

Анотація
У даній роботі розглянуто ознаки операційних систем, що можуть бути використані для визначення типу операційної системи віддаленого вузла. Проведено експеримент на основі методів машинного навчання по визначенню рівня важливості факторів.
Ключові слова: операційна система, безпека, мережа, фактори, машинне навчання.

Abstract
This paper discribes the features of operating systems that can be used to determine the type of operating system of a remote node. An experiment was conducted on the basis of machine learning methods to determine the level of importance of factors.
Keywords: operating system, security, network, factors, machine learning.

Вступ

Вирішуючи задачу визначення типу операційної системи (ОС), незалежно від використовуваного методу (сигнатури, рішення за допомогою штучного інтелекту тощо), необхідно визначити характеристики, які безпосередньо можуть вказувати на це [1]. Також потрібно враховувати різні фактори, що впливають на процес визначення типу операційної системи [2]. Від коректно сформованого та достатнього набору ознак може суттєво залежати здатність засобів класифікувати тип ОС [3].

Заголовки мережевих протоколів як ознаки операційної системи

Визначення операційної системи має здійснюватися шляхом аналізу параметрів відповідних полів у різних заголовках мережевих протоколів. Сучасні ОС підтримують велику кількість мережевих протоколів. Реалізація в більшості випадків стандартна, але значення деяких параметрів може суттєво відрізнятися в різних ОС [4]. Далі наведено деякі параметричні характеристики протоколів IP, ICMP і TCP в різних операційних системах.

Протокол IP є основним протоколом зв'язку в наборі протоколів Інтернету. Його завдання — передавати дейтаграми між мережами. Завдання протоколу полягає в доставці пакетів від вихідного вузла до вузла призначення, використовуючи тільки IP-адресу, зазначену в заголовку пакета. Параметри IP, які можна використовувати для визначення операційної системи хоста:

1. TTL - Час життя пакета (максимальна кількість переходів між маршрутизаторами). Значення цього параметра встановлюється по-різному в кожній операційній системі. UNIX-подібні системи зазвичай мають значення 64, операційні системи серії Windows - 128, Solaris/AIX - 254 [5].

2. Біт DF (не фрагментувати) – біт (значення 1), який вказує на те, що пакет не може бути фрагментований. Деякі операційні системи встановлюють цей біт лише в пакетах з прапором SYN, тоді як інші перевіряють, чи встановлено принаймні прапор SYN. Цей параметр слід розглядати в поєднанні зі значеннями інших параметрів.

3. ID - Поле ідентифікації дейтаграми. Певні операційні системи мають різні підходи до вирішення проблем, які додають цінності галузі [6]. Сімейство операційних систем OpenBSD використовує псевдовипадкове значення як приріст, тоді як Windows і Solaris використовують значення «1» як приріст.

ICMP – це протокол, який використовується мережевими пристроями для надсилання повідомлень про помилки та сервісних повідомлень. Параметри протоколу ICMP, які можна використовувати для визначення операційної системи хоста:

1. Код ICMP - код пакета ICMP. Це поле разом із полем Тип описує характер повідомлення. Хоча цей параметр має бути встановлений на 0 під час генерації пакета ехо-відповіді, деякі операційні сис-

теми встановлюють для нього щось інше. MacOS і Linux встановлюють для цього поля ненульове значення.

2. Ідентифікатори та порядкові номери - для ідентифікації запитів до різних джерел і для ідентифікації запитів у цільовому вузлі. Операційні системи Windows використовують константне значення поля Identifier, на відміну від Linux та MacOS.

3. Echo message - дані пакетів ICMP Echo. В операційних системах сімейства Windows у кожному запиті міститься 32 байти даних, ідентичних для усіх пакетів. Тоді як Unix подібні системи надсилають 56 байт даних, з яких перших 8 це мітка часу, а решта 48 незмінні в усіх пакетах [7].

Протокол TCP – основний протокол обміну даними. Забезпечує надійну та впорядковану передачу даних між застосунками у мережі. Параметри протоколу TCP, які можуть використовуватись для визначення операційної системи вузла:

1. Sequence Number - при створенні з'єднання вузол надсилає ISN (Initial Sequence Number - початковий номер послідовності). Для сімейства Windows це значення збільшується періодично на певну величину, тоді як для FreeBSD та Linux це значення є псевдовипадковим. Враховуючи, що більшість систем використовують випадкові значення, параметр матиме невеликий вплив на достовірність визначення [8].

2. Window size - використовується для визначення того, скільки ще байт може отримати вузол. Для сімейства Windows це значення 65535 або 8192, тоді як для Linux – 5840.

Протокол HTTP – основний протокол, призначенням якого є передача веб сторінок. Параметром протоколу HTTP, який може використовуватись для визначення операційної системи вузла, є User Agent. Цей параметр використовується для ідентифікації клієнтської частини. Досить часто містить інформацію про тип операційної системи, оскільки від цього параметра залежить вигляд сторінок, які переглядає користувач вузла.

Фактори, що впливають на виявлення ОС

Для визначення переліку факторів та їх рівнів впливу на виявлення ОС, було проведено експеримент, в ході якого на досліджуваних операційних системах виконувались наступні дії:

- надсилання 20 ICMP пакетів з полем Type 8 Echo Request з операційної системи, що вивчається;
- перегляд відео на веб-ресурсі YouTube (тривалістю не менше ніж 30 секунд) з операційної системи, що вивчається;
- перегляд різних веб-сторінок з операційної системи, що вивчається;

завантаження зображень з мережі Інтернет з операційної системи, що вивчається.

В результаті було визначено наступний перелік заголовків протоколів, які можна використовувати для визначення типу ОС:

1. IP – version, hdr_len, dsfield, dsfield_dscp, dsfield_ecn, len, id, flags, flags_rb, flags_df, flags_mf, frag_offset, ttl, proto, checksum, checksum_status.
2. ICMP – type, code, checksum, checksum_status, ident, seq, seq_le, data, data_data, data_len.
3. TCP – hdr_len, flags, flags_res, flags_ns, flags_cwr, flags_ecn, flags_urg, flags_ack, flags_push, flags_reset, flags_syn, flags_fin, flags_str, window_size_value, window_size, window_size_scalefactor, checksum, checksum_status.
4. DNS – id, flags, flags_response, flags_opcode, flags_truncated, flags_recdesired, flags_z, flags_checkdisable, count_queries, count_answers, count_auth_rr, count_add_rr, qry_name_len, count_labels, qry_type, qry_class.
5. HTTP – user_agent.

Отриманий перелік було оброблено методом машинного навчання, використовуючи метод рекурсивного зменшення ознак (RFE, recursive feature elimination) на основі моделі типу «дерево рішень». Після обробки, було отримано наступні результати важливостей факторів у відсотковому співвідношенні (рис. 1). Важливості ознак за спаданням:

1. icmp_data_len (x9) - 36,87%;
2. tcp_window_size_scalefactor (x13) - 29,5184%.
3. ip_ttl (x3)- 24,8846%;
4. icmp_ident (x5) - 8,71%;
5. icmp_seq_le (x7) - 0,0089%;
6. tcp_window_size_value (x11) - 0,0081%.

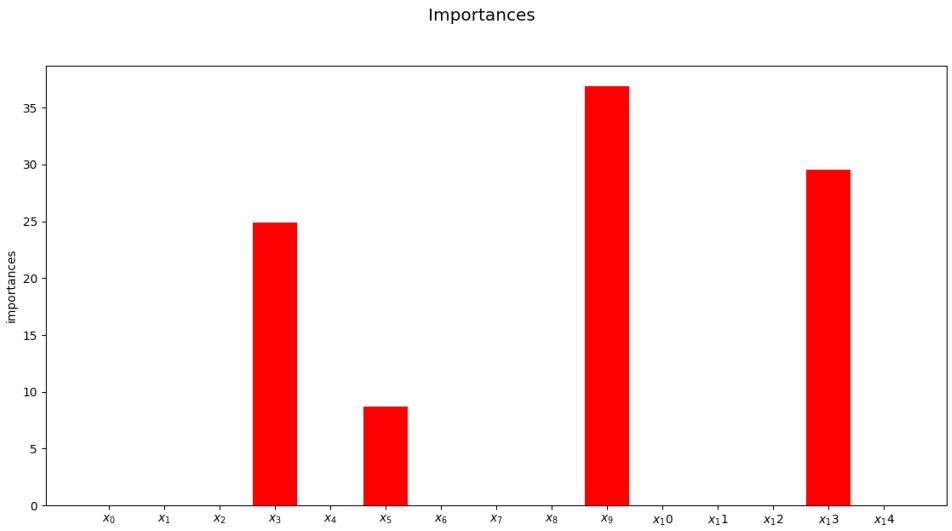

Рисунок 1 – Графік важливості ознак у відсотковому співвідношенні

Поле протоколу HTTP user_agent варто розглядати окремо, оскільки на відміну від інших факторів, його значення можна легко змінювати, перешкоджаючи процесу визначення ОС. На основі відібраних факторів отримано класифікатор, який із 99% точністю зміг визначити тип та версію ОС.

Висновки

Здійснено аналіз ознак ОС, які можуть бути використані для визначення її типу та версії. В результаті експериментів отримано перелік найбільш релевантних факторів. Використання їх для розробки моделі машинного навчання показало високоточні результати. У якості перспективи досліджень можна розширити спектр досліджуваних типів ОС.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Борусевич, А. В., Куперштейн, Л. М., «Машинне навчання у задачах виявлення типу операційної системи віддаленого хоста» в Матеріали конференції «Молодь в науці: дослідження, проблеми, перспективи (МН-2022)», Вінниця, 2022. [Електронний ресурс]. Режим доступу: https://conferences.vntu.edu.ua/index.php/mn/mn2022/paper/viewFile/14534/12285 Дата звернення: Трав. 2022

2. Борусевич, А. В., Куперштейн, Л. М., «Аналіз методів та засобів виявлення типу операційної системи віддаленого хоста» в Матеріали конференції «L Науково-технічна конференція підрозділів Вінницького національного технічного університету (2021)», Вінниця, 2021. [Електронний ресурс]. Режим доступу: https://conferences.vntu.edu.ua/index.php/allvntu/index/pages/view/zbirn2021 Дата звернення: Трав. 2022

3. L. Kupershtein, T. Martyniuk, O. Voitovych and A. Borusevych: Remote Host Operation System Type Detection Based on Machine Learning Approach. Proc. of Intelligent Solutions 2021 (Computational Intelligence & Decision Making Theory), Ukraine, September 28-30, 2021, CEUR-WS.org, online CEUR-WS.org/Vol-3106/Paper_7.pdf.

4. De Montigny-Lebouf A. A Multi-Packet Sugnature Approach to Passive Operating System Detection : монографія. Канада, 2005. 182 с.

5. Default TTL (Time To Live) Values of Different : веб-сайт. URL: https://subinsb.com/default-device-ttl-values/ (дата звернення 27.05.2022).

6. Passive OS Fingerprinting Update : веб-сайт. URL: https://isc.sans.edu/diary/Passive+OS+Fingerprinting+Update/18 (дата звернення 27.05.2022).

7. A. De Montigny-Leboeuf. 2004. A Multi-Packet Signature Approach to Passive Operating System Detection. DRDC OTTAWA TM 2005-018 / CRC-TN-2005-001. Communications Research Centre Canada.Defence R&D Canada - Ottawa

8. Определение операционной системы удаленного хоста : веб-сайт. URL: https://nmap.org/nmap-fingerprinting-article-ru.html (дата звернення 27.05.2022).

Борусевич Артур Вячеславович — студент групи 1БС-21м, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, e-mail: borusevych.av@gmail.com

Куперштейн Леонід Михайлович — кандидат технічних наук, доцент кафедри захисту інформації, Вінницький національний технічний університет, м. Вінниця

Borusevych Artur V. — Student of Information Technologies and Computer Engineering Department, Vinnytsia National Technical University, email : borusevych.av@gmail.com

Kupershtein Leonid M. — PhD, Associated Professor of Information Protection Chair, Vinnytsia National Technical University, Vinnytsia, email: kupershtein.lm@gmail.com