Methods of information systems protection

УДК 004.056:004.49 **doi: 10.20998/2522-9052.2019.4.13**

С. М. Лисенко[1], К. Ю. Бобровнікова[1], В. С. Харченко[2]

[1] Хмельницький національний університет, Хмельницький, Україна
[2] Національний аерокосмічний університет імені М. Є. Жуковського «ХАІ», Харків, Україна

МЕТОДИ ВИЯВЛЕННЯ БОТ-МЕРЕЖ В КОМП'ЮТЕРНИХ СИСТЕМАХ

Анотація. Об'єкт. Процес виявлення бот-мереж у корпоративних мережах на основі аналізу мережевого трафіка та поведінки програмного забезпечеення комп'ютерних системах. **Предмет.** Методи виявлення бот-мереж в комп'ютерних системах. **Мета.** Підвищення достовірності виявлення бот-мереж шляхом розроблення методів їх виявлення бот-мереж в корпоративних мережах. **Результати.** Запропоновано новий підхід до виявлення бот-мереж в корпоративних мережах на основі аналізу поведінки ботів. Виявлення бот-мереж здійснюється шляхом застосування розроблених двох методів: за допомогою аналізу на мережному рівні та на хостовому рівні. Перший метод дозволяє аналізувати поведінку програмного забезпечення на хості, що може вказувати на можливу присутність бота безпосередньо на хості і виявлення шкідливого програмного забезпечення, тоді як другий метод включає в себе моніторинг і аналіз DNS-трафіка, що також дозволяє зробити висновок про інфікування мережних хостів ботами бот-мережі. На основі запропонованих методів було розроблено ефективний інструмент виявлення бот-мереж - BotGRABBER. Він здатний виявляти боти, які використовують такі методи ухилення від виявлення як періодична зміна IP-відображення (cycling of IP mapping), «потік доменів» (domain flux), «швидкозмінні» мережі (fast flux) та DNS-тунелювання (DNS tunneling). **Висновки**. Використання розробленої системи дозволяє виявляти інфіковані ботами бот-мереж хости і локалізувати шкідливі програми з високою ефективністю – до 96%, а також демонструє низькі помилкові спрацьовування – на рівні 3-5%. Особливість запропонованого підходу полягає в тому, що виявлення бот-мереж є «невидимим» для власників бот-мереж.

Ключові слова: бот; бот-мережа; DNS-трафік; методи ухилення від виявлення бот-мереж; шкідливе програмне забезпечення; штучні імунні системи; детектори; позитивний алгоритм відбору

Вступ

На сьогоднішній день бот-мережі є джерелом таких кіберзлочинів, як DDoS-атаки, банківські шахрайства та кібер-шпигунства. Такі злочинні дії завдають значної шкоди економіці. Власники бот-мереж використовують різні технології для розробки, контролю, підтримки і приховування їх C&C-інфраструктури. Основними труднощами виявлення бот-мереж є те, що вони використовують такі технології, як P2P [1] та методи ухилення від виявлення, зокрема такі як періодична зміна IP-відображення, «потік доменів», «швидкозмінні» мережі та DNS-тунелювання. Ці технології підвищують резильєнтність бот-мереж стосовно дій, спрямованих на їх знешкодження. Для керування і контролю над інфікованими хостами переважна більшість бот-мереж використовують DNS. Крім того, деякі бот-мережі використовують методи шифрування корисного навантаження комунікацій по протоколу DNS для того, щоб запобігти їх виявленню [2].

Аналіз останніх досліджень та публікацій. Спільнота дослідників приділяє пильну увагу проблемі виявлення бот-мереж, і відома велика кількість підходів, присвячених її вирішенню.

Одним із способів вивчення поведінки бот-мереж є використання мереж-приманок. Наприклад, підходи [3,4] представляють детальний аналіз різних атак проти мереж-приманок на базі Linux. Підходи показують, як аналізувати атаки залежно від їх типів, таких як тривалість сесії, країни та регіональний Інтернет-реєстру (RIR) походження атаки. Крім того, мережі-приманки представлені як найефективніший інструмент виявлення та аналізу загроз з Інтернету. Автори показують останні результати для мереж-приманок на базі Dionaea (емуляція служб Windows), Kippo (емуляція служб Linux) та Glastopf (емуляція служб веб-сайтів). Недоліком підходу є те, що розмежування мереж-приманок за їх IP-адресами порівняно грубе.

В роботі [5] описаний підхід механізму захисту рухомої цілі, який захищає автентифікованих клієнтів від DDoS-атак Інтернет-сервісів. Запропонований підхід включає групу динамічних прихованих проксі-серверів для передачі трафіку між автентифікованими клієнтами та серверами. З метою захисту клієнтів вони відокремлюються від зловмисних кібервторгнень за допомогою серій переборів та постійної заміни атакованих проксі-серверів резервними проксі-серверами та перепризначенням атакованих клієнтів на нові проксі. З метою реалізації відокремлення зловмисних вторгнень автори розробили ефективний жадібний алгоритм. Крім того, з метою оцінки ресурсів, необхідних для захисту від DDoS-атак та задоволення визначених рівнів QoS (Quality of Service) під час різних атак, вивчається та оцінюється можливість карантину вторгнень з використаням запропонованого жадібного алгоритму.

У дослідженні [6] пропонується процедура вилучення інформації та виявлення бот-мереж через сліди інфікованої системи з бот-мережею з метою реконструкції атаки бот-мережі та підготовки пакету цифрових доказів, який підтверджує шкідливі дії та шкідливі наслідки цієї атаки в суді.

В [7] представлена таксономія поведінкових ознак бот-мереж, методи виявлення та захисту. Цей

загальний огляд підкреслює можливості захисту мережі шляхом виявлення недоліків у існуючих підходах. Крім того, продемонстрована корисність класифікації за розмірами для оцінки методів виявлення бот-мереж за допомогою різних показників. Підхід демонструє особливості поведінки бот-мереж та вплив використання таксономії на точність виявлення бот-мереж.

Підхід, представлений у [8], описує принципи функціонування бот-мереж в різних аспектах: вибір кандидата у боти, побудова мережі, механізми / протоколи зв'язку з C&C та підходи до пом'якшення наслідків. Дослідження надає математичний аналіз двох підходів до усунення P2P бот-мереж: захист від атак отруєння індексу та Sybil, а також методи пасивного моніторингу, засновані на інфільтрації мережприманок чи захопленні ботів.

У статті [9] представлений підхід для виявлення як низькошвидкісних, так і високошвидкісних DDoS-атак. З цією метою автори використовують та емпірично оцінюють такі інформаційні показники, як ентропія Хартлі, Шеннона, Рені, узагальнена ентропія, розходження Кульбака-Лейблера та узагальнена міра інформаційної відстані. Дослідження включає відповідну метрику, яка полегшує побудову ефективної моделі для виявлення низькошвидкісних та високошвидкісних DDoS-атак.

У роботі [10] з метою забезпечення безпеки мережі наводиться таксономія інструментів для здійснення атак. Автори також представляють всебічний та структурований аналіз існуючих інструментів та систем, які можуть бути корисними як для зловмисників, так і для захисників мережі. У такому контексті обговорюються як переваги, так і недоліки представлених систем. Представлені рішення часто використовуються в мережній інфраструктурі для забезпечення безпеки, але вони неефективні для атак нульового дня.

У [11] проаналізовано мережу підприємства, яка використовує хмарні технології та програмно-конфігуровану мережу. Автори статті вивчали вплив заходів безпеки на механізми захисту від атак DDoS у мережі підприємства. Основна ідея статті полягає в тому, щоб показати, що використання архітектури для пом’якшення DDoS-атак може допомогти підприємствам захищатись від таких атак. Описана архітектура інтегрує високопрограмований мережний моніторинг, що дозволяє виявляти атаку, та дієву структуру управління для забезпечення швидкої та визначеної реакцію на атаку.

В [12] представлені принципи забезпечення безпеки бездротової мережі. Викладено чітке дослідження основних аспектів безпеки в самоорганізованих мережах та інших мережах, які використовують бездротові технології для зв'язку. Розглянуто основні аспекти безпеки та часто використовувані терміни. Після вивчення критичних проблем безпеки в множині бездротових мереж запропоновано конкретні рішення для загроз безпеки. Основним недоліком запропонованого підходу є те, що він не здатний реагувати на невідомі кібератаки, які виконуються бот-мережами.

У роботі [13] представлено підхід до виявлення аномальних шаблонів мережних з'єднань за допомогою штучних нейронних мереж, імунної системи, нейронечітких класифікаторів та їх комбінацій. У статті описана архітектура системи виявлення вторгнень на основі запропонованого методу. Розроблена система виявлення вторгнень є багаторівневою: в першу чергу проводиться сигнатурний аналіз, потім використовується комбінація адаптивних детекторів. Проведені експерименти демонструють ефективність методів з точки зору хибних спрацювань, виявлень та правильність класифікації.

У [14, 15] описано метод виявлення мережних атак та шкідливого коду. Метод грунтується на основних принципах штучної імунної системи, де імунні детектори мають структуру штучних нейронних мереж. Основна мета запропонованого підходу – виявлення раніше невідомих кібератак. Запропонована інтелектуальна система кіберзахисту може підвищити надійність виявлення вторгнень в комп'ютерних системах і, як результат, може зменшити фінансові втрати компаній від кібератак.

Загальним недоліком вищевказаних та інших відомих підходів є те, що вони не в змозі вирішити проблем нульового дня щодо виявлення бот-мереж.

1. Попередня робота

В роботі [16] був запропонований метод виявлення бот-мереж на основі DNS в корпоративних мережах. Метод включав два способи виявлення бот-мереж: пасивний і активний моніторинг DNS. Це надавало можливість виявляти бот-мережі, які використовують такі технології ухилення як періодична зміна IP-відображення, «потік доменів», «швидкозмінні» мережі та DNS-тунелювання [17-25, 27]. Така система – BotGRABBER - була заснована на кластерному аналізі ознак, отриманих з корисного навантаження DNS-повідомлень, та які можуть вказувати на використання технологій ухилення від виявлення бот-мереж. Результатом кластеризації був ступінь приналежності векторів ознак до одного з чотирьох кластерів, де приналежність вектора ознак до кластера вказувала на виконання запитів, що пов’язані з використанням технологій ухилення бот-мереж на основі DNS [16].

З метою усунення можливої невизначеності результатів кластеризації, метод застосовував додаткові ознаки, отримані шляхом активного DNS-зондування.

Основний недолік запропонованого підходу полягає в тому, що активне DNS-зондування може бути помітним для власника бот-мережі. Крім того, застосування кластерного аналізу в багатьох випадках має обмежені можливості для опису поведінки ботів, які використовують технології ухилення. Це призводить до неможливості виявлення бот-мереж, якщо їх поведінка може бути віднесена до «зловмисних» і «нормальних» кластерів одночасно. Крім того, цей метод може виявити вузли, які інфіковані бот-мережами, але не в змозі безпосередньо локалізувати шкідливе програмне забезпечення на інфікованому хості.

2. Технологія виявлення кіберзагроз в компютерних системах

В цій статті представлено новий підхід до виявлення бот-мереж на основі аналізу поведінки бот-мереж в корпоративній мережі. Технологія розроблена з метою усунення недоліків описаного вище підходу, а також розширення можливостей системи виявлення бот-мереж BotGRABBER.

З цією метою запропоновано здійснювати виявлення бот-мереж шляхом об'єднання двох етапів: за допомогою аналізу на мережному рівні і аналізу на хостовому рівні. Вони покликані покращити і уточнити результати один одного. З одного боку, цей підхід дозволяє аналізувати поведінку програмного забезпечення на хості, що може вказувати на можливу присутність бота безпосередньо на хості і надати можливість здійснити виявлення шкідливого програмного забезпечення. З іншого боку, він включає в себе моніторинг і аналіз DNS-трафіка, що дозволяє здійснити висновок щодо інфікування хостів мережі ботами бот-мережі. Таким чином, нова технологія виявлення бот-мереж дозволяє підвищити ефективність виявлення бот-мереж в корпоративних мережах.

Два методи працюють паралельно таким чином:

1). Якщо виявлено підозріле програмне забезпечення на хості корпоративної мережі, то воно блокується. Після цього підхід передбачає запит результатів аналізу мережного рівня стосовно підозрілих DNS-запитів від потенційно інфікованого хоста і від інших хостів мережі для подальшого аналізу.

2). Якщо за допомогою аналізу мережного рівня виявляються підозрілі DNS-запити, то формується список інфікованих хостів мережі. На основі знань щодо підозрілої поведінки програмного забезпечення здійснюється локалізація бота на інфікованому хості, а потім його блокування.

Розроблена технологія включає в себе два методи виявлення бот-мереж. Розглянемо кожен з них.

2.1. Метод виявлення бот-мереж на основі аналізу поведінки ботів на хості. З метою виявлення підозрілого програмного забезпечення з ознаками ботів бот-мереж, був розроблений новий метод, заснований на їх поведінці на хості. Він включає в себе наступні кроки:

1) побудова множини поведінок ботів на різних етапах їх життєвого циклу у вигляді шаблонів (представлені у вигляді бітових рядків);

2) моніторинг системних викликів програмного забезпечення хоста;

3) представлення спостережуваної поведінки програмного забезпечення в якості бітового рядка;

4) порівняння побудованого бітового рядка, що описує поведінку програмного забезпечення, з множиною бітових рядків, які представлють шаблони поведінок бот-мереж на різних етапах життєвого циклу ботів;

5) блокування функціонування програмного забезпечення, що оцінюється як шкідливе;

6) запити про результати, отримані за допомогою аналізу на мережному рівні про підозрілі DNS-запити від потенційно інфікованого хоста і від інших хостів мережі для подальшого аналізу.

Поведінкова модель бота. З метою виявлення ботів бот-мереж на хостах ми повинні дослідити їх ознаки і побудувати їх поведінкову модель. Розроблена поведінкова модель враховує онаки бота і формалізує його процес функціонування на хості протягом життєвого циклу, який включає в себе п'ять етапів: (1) інфікування; (2) первинна реєстрація або з'єднання з C&C сервером; (3) виконання шкідливої активності; (4) обслуговування; (5) припинення функціонування бота.

На основі знань щодо поведінок і основних ознак ботів бот-мереж ми можемо побудувати множину поведінок ботів на різних етапах їх життєвого циклу $DB = \{\Phi_{l_j}\}, j = \overline{1,5}$, де $L = \{l_j\}_{j=1}^{5}$ – набір етапів життєвого циклу бота. Представимо відомий поведінковий шаблон і поведінку спостережуваного програмного забезпечення у вигляді бітових рядків $\Phi = y_1 y_2 ... y_n$ і $T = t_1 t_2 ... t_m$ відповідно, де $y_i \in Y, t_i \in Y$. Визначимо $\Omega(\Phi, T)$ – булеву функцію порівняння рядків відомого поведінкового шаблону і поведінки спостережуваного програмного забезпечення, яка вказує на співпадіння або неспівпадіння.

Порівняння побудованого бітового рядка, що описує поведінку програмного забезпечення, з множиною бітових рядків, які описують поведінкові шаблони бот-мереж. На наступному етапі ми відслідковуємо поведінку програмного забезпечення на хості і представляємо її у вигляді бітового рядка з метою подальшого порівняня з раніше відомими шаблонами поведінок для ботів.

Всі групи поведінок представлені у вигляді шаблонів (а саме у вигляді бітових рядків). Таким чином, завдання ідентифікації підозрілої поведінки будь-якого програмного забезпечення полягає у знаходженні збігів рядка для спостережуваної поведінки з рядком, представленим в базі поведінок бот-мереж. Для вирішення цієї задачі був використаний алгоритм приблизного порівняння рядків (рішення задачі k-відмінності). Нехай задані два рядки, послідовність $T = t_1 t_2 ... t_m$ і шаблон $\Phi = y_1 y_2 ... y_n$ в деякому алфавіті Σ, а також ціле число k, тоді алгоритм дозволяє знайти всі підрядки Φ' для T з відстанню редагування щонайбільше k для Φ. Відстань редагування визначає мінімальну кількість операцій для редагування (відмінність), які необхідні для перетворення Φ' до Φ. Під операцією редагування мається на увазі вставка, видалення або зміна символу. Проблема k відмінності є окремим випадком зі зміною в одну операцію редагування [26]. Обробка шаблону потребує часу $O(mn)$. Табл. 1 демонструє експериментальні результати приблизного порівняння рядків для різних значень довжини R і параметра k. Таким чином, якщо $k = 0$, то це означає точний збіг, і немає ніякого рішення. Проте, коли значення $k = 2$, $k = 3$, то число рішень є досить низьким. Збільшення k призводить до зростання числа

рішень, причому час пошуку збігів також зростає. Дослідження показують, що вирішення задачі виявлення підозрілої поведінки можливе, коли параметр, $k = 4$. Представляючи $\Omega(\Phi_{l_j}, T)$ як факт ідентифікації підозрілої поведінки будь-якого програмного забезпечення, яка схожа на певний етап життєвого циклу ботів l_j, процедура моніторингу може бути описана у вигляді алгоритму 1.

Таблиця 1 – **Експериментальні результати приблизного порівняння рядків для різних значень довжини *R* і параметра** k

	Алфавіт Q	**Довжина послідовності R**	**Параметр k-відмінності**	**Кількість знайдених рядків**
Φ_1	430	35	0	0
	430	35	2	0
	430	35	3	0
	430	35	4	1
	430	35	5	2
Φ_2	430	78	0	0
	430	78	2	0
	430	78	3	1
	430	78	4	3
	430	78	5	8
Φ_3	430	93	0	0
	430	93	2	1
	430	93	3	4
	430	93	4	17
	430	93	5	24

```
if ((Ω(Φ_{l_1}, T) and Ω(Φ_{l_2}, T)) or Ω(Φ_{l_2}, T) is true) then
    mark software as suspicious, querying the results of the network −
    level analysis about anomaly in the DNS − traffic
    and possible infection of the host;
    if (the results of the network − level analysis indicates, that
    the host is infected) then
        block_the_software
    end
end
if (Ω(Φ_{l_3}, T) or Ω(Φ_{l_4}, T) is true) then
    block_the_software.
end
```

Алгоритм 1. Функціонування методу виявлення бот-мереж на основі аналізу поведінки ботів на хості на стадії моніторингу

2.2 Метод виявлення бот-мереж на основі DNS, який ґрунтується на використанні штучних імунних систем. З метою виявлення бот-мереж в мережі був розроблений новий метод на основі DNS, який ґрунтується на використанні штучних імунних систем (ШІС). Метод призначений для виявлення аномалій в DNS-трафіку та включає в себе наступні кроки:

1) побудова множини поведінкових шаблонів ботів в DNS-трафіку (представлених у вигляді бітових рядків);

2) збір вхідного DNS-трафіка мережі;

3) аналіз полів TTL вхідних DNS-повідомлень щодо певного доменного імені;

4) побудова вектора ознак на основі ознак, вилучених з вхідних DNS-повідомлень щодо певного доменного імені;

5) виявлення аномалій в DNS-трафіку, засноване на використанні штучної імунної системи;

6) одержання списку заражених хостів в мережі, виконання аналізу програмного забезпечення на інфікованому хості на хостовому рівні, локалізація і блокування програмного забезпечення.

Розглянемо кроки методу більш детально.

Побудова множини поведінкових шаблонів ботів в DNS-трафіку. З метою виявлення DNS-туннелювання аналізуються такі ознаки, отримані з корисного навантаження DNS-повідомлень [23-25]:

1) l_N – довжина запитуваного доменного імені, $l_N \in [75, 255]$;

2) n_U – кількість унікальних символів в доменному імені, $n_U \in (27, 37]$;

3) e_N – ентропія доменного імені, $e_N \geq f_{Eb32}$, де f_{Eb32} – функція залежності ентропії поля DNS-повідомлення від його довжини, n – основа кодування;

4) f_{UR} – використання рідко вживаних типів DNS-записів, які зазвичай не використовується типовим клієнтом (наприклад, TXT, що найбільш часто використовуються для тунелювання (за винятком поштових серверів), KEY або NULL);

5) e_R – ентропія DNS-записів, які містяться в DNS-повідомленнях (CNAME, TXT, NS, MX, KEY, NULL тощо), $e_R \geq f_{Eb64}, e_R \geq f_{Eb256}$;

6) l_P – максимальний розмір DNS-повідомлень, $l_P > 300$.

З метою виявлення таких технологій ухилення, як «швидкозмінні» мережі, «потік доменів» та періодична зміна IP-відображення, використовуються наступні ознаки [17-22]:

1) n_{IP} – кількість IP-адрес, пов'язаних з доменним ім'ям, $n_{IP} \in (5, \infty)$;

2) s_{IP} – середня дистанція між IP-адресами, пов'язаними з доменним ім'ям, $s_{IP} \in (65535, \infty)$;

3) n_A – кількість A-записів, що відповідають доменному імені, у вхідному DNS-повідомленні, $n_A \in (5, \infty)$;

4) s_A – середня дистанція між IP-адресами в множині A-записів для доменного імені у вхідному DNS-повідомленні, $s_A \in (65535, \infty)$;

5) n_{UA} – кількість унікальних IP-адрес в множинах A-записів, що відповідають доменному імені, у вхідних DNS-повідомленнях, $n_{UA} \in (8, \infty)$;

6) s_{UA} – середня дистанція між унікальними IP-адресами в множинах A-записів, що відповідають доменному імені, у DNS-повідомленнях, $s_{UA} \in (65535, \infty)$;

7) n_D – кількість доменних імен, які спільно використовують IP-адресу, що відповідає доменному імені, $n_D \in [8, \infty]$;

8) $t_{mod}, t_{med}, t_{aver}$ – TTL-період, мода, $t_{mod} \in [0, 900]$, медіана, $t_{med} \in [0, 900]$, середнє арифметичне значення, $t_{aver} \in [0, 900]$;

9) f_S – ознака успішності DNS-запиту.

Крім того, метод використовує дві ознаки: групове очищення локальних кешів DNS f_f [28] і синхронізації DNS-запитів від хостів мережі f_q [16, 28].

Таким чином, ми можемо представити шаблон, який описує поведінку бота, як набір бітових рядків ($a...s$ – індексні номери бітів в закодованій послідовності шаблону):

$$p_i = \left\langle \begin{array}{c} l_{N_1} ... l_{N_a} n_{U_{a+1}} ... n_{U_{a+b-1}} e_{N_{a+b}} ... e_{N_c} t_{mod_{c+1}} .. t_{mod_{c+d-1}} \\ t_{med_{c+d}} ... t_{med_e} t_{aver_{e+1}} .. t_{aver_{e+f-1}} n_{A_{e+f}} ... n_{A_g} n_{IP_{g+1}} ... n_{IP_{g+h-1}} \\ s_{IP_{g+h}} ... s_{IP_i} s_{A_{i+1}} ... s_{A_{i+j-1}} n_{UA_{i+j}} ... n_{UA_k} s_{UA_{k+1}} ... s_{UA_{k+l-1}} \\ n_{D_{k+l}} ... n_{D_m} f_{UR_{m+1}} ... f_{UR_{m+n-1}} e_{R_{m+n}} ... e_{R_o} l_{P_{o+1}} ... l_{P_{o+l-1}} \\ f_{S_{o+p}} ... f_{S_q} f_{q_{q+1}} ... f_{q_{q+r-1}} f_{f_{q+r}} ... f_{f_s} \end{array} \right\rangle. \quad (2)$$

Збір вхідного DNS-трафіка мережі. Вхідний DNS-трафік збирається за допомогою множини мережних сніфферів, підключених до комутатора з функцією дзеркалювання портів.

ПРИМІТКА. Для відкидання легітимних DNS-запитів і виявлення відомих шкідливих доменних імен метод використовує «білі» і «чорні» списки доменних імен.

Аналіз полів TTL вхідних DNS-повідомлень щодо певного доменного імені. На основі значень полів TTL обробляються такі DNS-повідомлення: кожне перше захоплене DNS-повідомлення щодо певного доменного імені в межах TTL-періоду DNS; кожне повторне DNS-повідомлення, отримане хостом в межах TTL-періоду, якщо джерело повідомлення не є локальним DNS-сервером і TTL-період, зазначений в цьому повідомленні, відрізняється від залишку TTL-періоду, в межах якого це повідомлення було отримано.

На етапі моніторингу на основі ознак, вилучених з вхідних DNS-повідомлень щодо певного доменного імені, будуються вектори ознак. Вони мають такий же вигляд, як і шаблон (2).

Виявлення аномалій в DNS-трафіку. На рівні мережного аналізу виявлення бот-мереж здійснюється із залученням апарату штучних імунних систем (ШІС) [29], які оперують з поняттям «свого-чужого» і дозволять виявляти аномальний DNS-трафік, який свідчить про присутність бот-мереж в корпоративній мережі.

Для здійснення розмежування «свій-чужий» в роботі використано концепцію позитивного відбору [30]. Виявлення аномалій включає наступні етапи: побудова шаблонів аномального DNS-трафіку, генерація детекторів, могніторинг і розпізнавання аномалій.

Шаблони аномального DNS-трафіку відповідають поведінці бот-мереж та засновані на знаннях стосовно ознак DNS-трафіка бот-мереж. Значення кожної ознаки в шаблоні були нормалізовані відповідно до розкиду даних. Нормовані дані були розділені на інтервали $d = (\max - \min) / (2^m - 2)$, де m визначає кількість двійкових чисел, використовуваних для кодування. Кожен шаблон був закодований в двійковій формі відповідно до інтервалу n, до якого належить ознака $n \in [1; 2^m - 2)$. Якщо значення ознаки знаходиться поза межами інтервалу [min; max], то вона була закодована як m 0 і 1 відповідно.

Використовуючи алгоритм позитивного відбору, була згенерована множина детекторів D, які мають високу спорідненість із зловмисними шаблонами.

Метою стадії розпізнавання є ідентифікація високої спорідненості детектора з вхідними даними. Це вказує на виявлення аномалій, а отже, наявність бот-мереж в мережі.

Для того, щоб визначити значення спорідненості, використано правило співпадіння r-суміжних бітів. Згідно з цим правилом, два рядки довжини l

співпадають, якщо вони збігаються принаймні в r суміжних бітових позиціях [31]. Співпадіння рядка з детектором розглядається як виявлення аномальної поведінки в DNS-трафіку.

ПРИМІТКА. Вектори ознак нормовані і кодуються таким же чином, як шаблони.

З метою досягнення високої ймовірності виявлення аномалій для створеної множини шаблонів було досліджено оптимальні значення числа детекторів N_R і параметра r. У табл. 2 і на рис. 1 показано, що значно вища ймовірність виявлення аномалій P_f може бути досягнута при значенні r = 32, але в цьому випадку необхідна генерація великої кількості детекторів. Таким чином, використання апарату ШІС дозволяє з високою ефективністю виявляти аномалії в DNS-трафіку.

Таблиця 2 – **Ймовірність виявлення аномалій з різними значеннями параметрів для алгоритма позитивного відбору**

m	**1**	**N_s * 1000**	**N_R * 100**	**Параметр, *r***	**Ймовірність виявлення аномалій, P_f,%**
2	64	10	10	8	0,88
2	64			16	14,15
2	64			32	33,51
2	64		50	8	0,98
2	64			16	29,87
2	64			32	85,77
2	64		100	8	1,02
2	64			16	54,25
2	64			32	99,3
2	64	100	100	8	2,01
2	64			16	15,3
2	64			32	25,6
2	64		500	8	2,55
2	64			16	23,36
2	64			32	72,88
2	64		1000	8	2,9
2	64			16	75,36
2	64			32	99,8

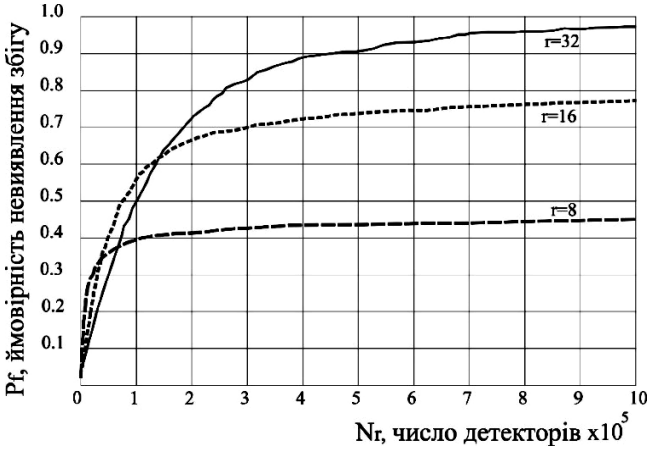

Рис. 1. Залежність ймовірності виявлення аномалій P_F від величини r і числа детекторів N_r

Схема функціонування інформаційної технології в узагальненому вишляді представлена на рис. 2.

3. Експерименти

Для визначення ефективності запропонованої технології виявлення бот-мереж на основі поведінки бот-мереж в корпоративній мережі були проведені серії експериментів. Процедура випробувань була такою ж, як представлено в [16]. В якості бази для дослідження була розгорнута кампусна мережа Хмельницького національного університету.

Було розроблено шкідливе програмне забезпечення, яке емулювало поведінку ботів в DNS-трафіку. Створене програмне забезпечення мало властивості ботів бот-мереж з централізованою архітектурою. Поведінка ботів, які емулювалися, відтворювала чотири типи сценаріїв, які застосовували технології ухилення. Кожен експеримент використовував один сценарій.

Кожен бот в процесі функціонування виконав 600 DNS-запитів. Згенероване програмне забезпечення мало можливість виконувати різні види сценаріїв функціонування бот-мереж відповідно до їх етапів життєвого циклу. Кожен експеримент включав різні способи інфікування (завантаження з вебсайту, завантаження з поштової скриньки, копіювання з флеш-диска). Для початкового з'єднання з C&C-серверами використовувались різні системні порти. Крім того, кожен бот виконував різні шкідливі дії, та всі боти підтримували функцію обслуговування. Через 24 години функціонування боти видалили самі себе. Маючи велику кількість способів відтворення шкідливої активності, були проведені 36 експериментів (9 експериментів для 4 типів технологій ухиленння).

ПРИМІТКА. Всі експерименти були проведені в межах університетської мережі і не могли завдати ніякої шкоди хостам і стабільній роботі мережі.

Також були зареєстровані і використовувались кілька «підроблених» доменних імен, які імітували діяльність C&C-серверів бот-мереж [16]. Крім того, було розроблено програмне забезпечення, яке імітувало активність користувачів і виконувало легітимні DNS-запити.

Головною метою експериментів було з'ясування переваг і недоліків розробленої технології. Для

дослідження ефективності запропонованої інформаційної технології було проведено чотири типи експериментів. Ми були зацікавлені в результатах, отриманих попереднім BotGRABBER [16]; шляхом застосування методу виявлення бот-мереж на основі аналізу поведінки ботів на хості (аналіз на хостовому рівні), описаного в розділі 4.1; шляхом застосування методу на основі DNS для виявлення бот-мереж з використанням штучної імунної системи (аналіз на мережному рівні), описаного в розділі 4.2; новим BotGRABBER, який поєднує аналіз як на хостовому, так і на мережному рівні.

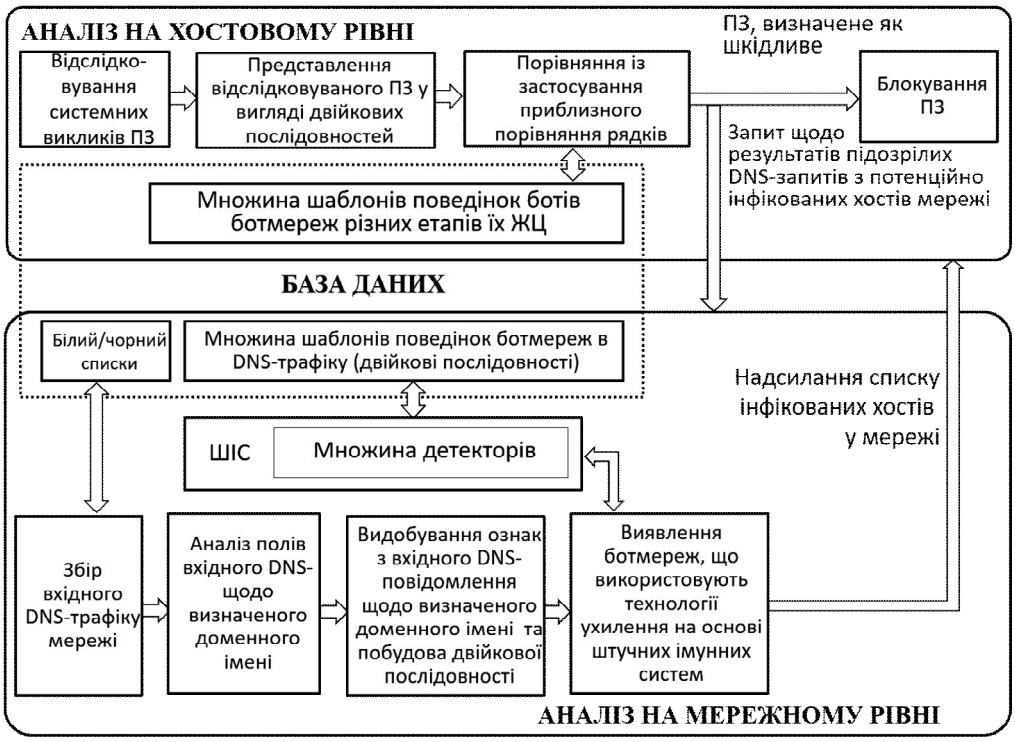

Рис. 2. Схема функціонування технології виявлення бот-мереж

Таблиця 3 демонструє результати експериментів, згаданих вище. Як видно з результатів, попередня система BotGRABBER в загальному показала хороші і навіть кращі результати, в порівнянні з методом виявлення бот-мереж на основі аналізу на хостовому рівні.

Проте, попередній BotGRABBER продемонстрував найгірші результати для виявлення технології ухилення бот-мереж «потік доменів». Якщо проаналізувати результати виявлення бот-мереж на хостовому рівні, використовувані окремо, то можна побачити, що вони є досить низькими. І, нарешті, використання системи BotGRABBER, яка включає в себе аналіз як на хостовому рівні, так і на мережному, продемонстрував більш значні результати виявлення, в порівнянні з попереднім.

Таблиця 3 – **Експериментальні результати виявлення бот-мереж: виявлення (TP) і хибні спрацювання (FP)**

Назва технології ухилення	**Виявлення за допомогою BotGRABBER:**							
	попереднього [16]		**нового, тільки аналіз на рівні:**					
			хостовому		**мережному**		**хостовому та мережному**	
	TP,%	**FP,%**	**TP,%**	**FP,%**	**TP,%**	**FP,%**	**TP,%**	**FP,%**
періодична зміна IP-відображення	99	2	91	3	99	1	99	1
«потік доменів»	89	6	77	6	90	6	96	6
«швидкозмінні» мережі	97	4	84	5	95	3	99	4
DNS-тунелювання	94	4	80	5	90	4	96	4
Всього	**94,75**	**4**	**83**	**4,75**	**93,5**	**3**	**97,5**	**3,75**

Таким чином, результати технології виявлення бот-мереж на основі поведінки бот-мереж в корпоративній мережі (реалізовано у вигляді системи BotGRABBER) показали високу ефективність (до 97%). Варто зауважити, що як попередня, так і нова система BotGRABBER показала аналогічні значення хибних спрацювань на рівні близько 3-5%.

Висновки

У роботі представлено нову технологію виявлення бот-мереж на основі аналізу поведінки бот-

мереж в корпоративній мережі. Виявлення бот-мереж здійснюється шляхом комбінації двох розроблених методів: за допомогою аналізу на мережному рівні та аналізу на хостовому рівні.

Перший метод дозволяє аналізувати поведінку програмного забезпечення на хості, що може вказувати на можливу присутність бота безпосередньо на хості і надає можливість здійснювати виявлення шкідливого програмного забезпечення.

Другий метод включає в себе моніторинг і аналіз DNS-трафіку, що дозволяє зробити висновок про інфікування мережних хостів ботами бот-мережі.

На основі розробленої технології було покращено ефективний інструмент виявлення бот-мереж BotGRABBER. Він здатний виявляти боти, які використовують такі технології ухилення від виявлення як періодична зміна IP-відображення, «потік доменів», «швидкозмінні» мережі та DNS-тунелювання.

Використання розробленої системи дозволяє виявляти інфіковані ботами бот-мереж хости і локалізувати шкідливі програми з високою ефективністю – до 96%, а також демонструє низькі помилкові спрацьовування – на рівні 3-5%. Особливість запропонованого підходу полягає в тому, що виявлення бот-мереж є «невидимим» для власників бот-мереж.

СПИСОК ЛІТЕРАТУРИ (REFERENCE)

1. Komar, M., Kochan, V., Sachenko, A. and Ababii, V. (2016), “Improving of the security of intrusion detection system”, *2016 International Conference on Development and Application Systems (DAS)*, pp. 315–319.
2. Harsha, T., Asha, S. and Soniya, B. (2016), “Feature selection for effective botnet detection based on periodicity of traffic”, *Information Systems Security*: 12th International Conference, ICISS 2016, Jaipur, India, December 16-20, 2016, Proceedings, pp. 471–478, DOI: https://doi.org/10.1007/978-3-319-49806-5_26.
3. Zuzcak, M. and Sochor, T. (2017), “Behavioral analysis of bot activity in infected systems using honeypots”, *Communications in Computer and Information Science*, Springer, Cham, vol. 718, pp. 118-133.
4. Sochor, T. and Zuzcak, M. (2015), “Attractiveness Study of Honeypots and Honeynets in Internet Threat Detection”, *22nd Int. Conf. Computer Networks: Communications in Computer and Information Science*, Springer International, Cham, 2015, pp. 69-81.
5. Wang, H., Jia, Q., Fleck, D., Powell, W., Li, F. and Stavrou, A. (2014), “A moving target DDoS defense mechanism”, *Computer Communications*, vol. 46, pp. 10-21.
6. Javadianasl, Y., Manaf, A. A. and Zamani, M. (2017), “A Practical Procedure for Collecting More Volatile Information in Live Investigation of Botnet Attack”, *Multimedia Forensics and Security*, Springer, pp. 381-414.
7. Khattak, S., Ramay, N. R., Khan, K. R., Syed, A. A. and Khayam, S. A. (2014), “A taxonomy of botnet behavior, detection, and defense”, *IEEE communications surveys & tutorials*, vol. 16, no. 2, pp. 898-924.
8. Wang, P., Wu, L., Aslam, B. and Zou, C. C. (2015), “Analysis of Peer-to-Peer botnet attacks and defenses”, *Propagation phenomena in real world networks*, Springer International Publishing, pp. 183-214.
9. Bhuyan, M. H., Bhattacharyya, D. K. and Kalita, J. K. (2015), “An empirical evaluation of information metrics for low-rate and high-rate DDoS attack detection”, *Pattern Recognition Letters*, vol. 51, pp. 1-7.
10. Hoque, N., Bhuyan, M. H., Baishya, R. C., Bhattacharyya, D. K. and Kalita, J. K. (2014), “Network attacks: Taxonomy, tools and systems”, *Journal of Network and Computer Applications*, vol. 40, pp. 307-324.
11. Wang, B., Zheng, Y., Lou, W. and Hou, Y. T. (2015), “DDoS attack protection in the era of cloud computing and software-defined networking”, *Computer Networks*, vol. 81, pp. 308-319.
12. Pathan, A. S. K. (2016), *Security of self-organizing networks*, MANET, WSN, WMN, VANET, CRC press, 638 p.
13. Branitskiy, A. and Kotenko, I. (2015), “Network Attack Detection Based on Combination of Neural, Immune and Neuro-Fuzzy Classifiers”, *IEEE 18th International Conference on Computational Science and Engineering (CSE)*, pp. 152-159.
14. Komar, M., Sachenko, A., Bezobrazov, S. and Golovko, V. (2017), “Intelligent Cyber Defense System Using Artificial Neural Network and Immune System Techniques”, Ginige A. et al. (eds), *Information and Communication Technologies in Education, Research, and Industrial Applications. ICTERI 2016*, pp. 36-55.
15. Bezobrazov, S., Sachenko, A., Komar, M. and Rubanau, V. (2016), “The methods of artificial intelligence for malicious applications detection in Android OS”, *International Journal of Computing*, vol. 15, no. 3, pp. 184-190.
16. Pomorova, O., Savenko, O., Lysenko, S., Kryshchuk, A. and Bobrovnikova, K. (2016), “Antievasion Technique for the Botnets Detection Based on the Passive DNS Monitoring and Active DNS Probing”, *International Conference on Computer Networks: Springer International Publishing*, pp. 83-95.
17. Schiller, C., R. Binkley and J. Botnets (2017), The Killer Web Application, *Syngress Publishing*, 464 p.
18. Yadav, S. and Reddy, A.L.N. (2011), “Winning with DNS failures: Strategies for faster botnet detection”, *Proc. of the 7th International ICST Conference on Security and Privacy in Communication Networks*, pp. 446-459.
19. Salusky, W. and Danford, R. (2007), *Know your enemy: Fast-flux service networks. The Honeynet Project*, available at: http://www.honeynet.org/book/export/html/130.
20. Nazario, J. and Holz, T. (2008), “As the Net Churns: Fast-Flux Botnet Observations”, *Conference on Malicious and Unwanted Software (Malware08)*, pp. 24-31.
21. *DAMBALLA. Botnet Communication Topologies. Understanding the intricacies of botnet command-and-control* (2019), available at: https://www.damballa.com/downloads/r_pubs/ WP_Botnet_Communications_Primer.pdf.
22. Bilge, L., Kirda, E., Kruegel, C. and Balduzzi, M. (2011), “EXPOSURE: Finding Malicious Domains Using Passive DNS Analysis”, *NDSS*, pp. 1-17.
23. Farnham, G. and Atlasis, A. (2013), *Detecting DNS Tunneling. SANS Institute InfoSec Reading Room*, pp. 1-32.
24. Dietrich, C.J., Rossow, C., Freiling, F. C., Bos, H., van Steen, M. and Pohlmann, N. (2011), “On Botnets that use DNS for Command and Control”, *Proceedings of European Conference on Computer Network Defense*, pp. 9-16.
25. Guy, J. (2009), *A study of DNS*, available at: http://armatum.com/blog/2009/ a-study-of-dns/.
26. Jorma, Tarhio and Esko, Ukkonen. (1993), “Approximate BoyerMoore String Matching”, *SIAM Journal on Computing*, vol. 22, no. 2, pp. 243-260.

27. Guy, J. (2009), *Dns part ii: visualization*, available at: http://armatum.com/ blog/2009/dns-part-ii/.
28. Pomorova, O., Savenko, O., Lysenko, S., Kryshchuk, A. and Bobrovnikova, K. (2015), "A technique for the botnet detection based on DNS-traffic analysis", *International Conference on Computer Networks*, Springer Int. Publishing, pp. 127-138.
29. Dipankar, D. (2013), "Artificial immune systems", *Encyclopedia of Sciences and Religions*, pp. 136–139 .
30. Zhang, F. and Qi, D. (2012), "A positive selection algorithm for classification", *J. Comput. Inf. Syst*, pp. 207–215.
31. Goswami, M. and Bhattacharjee, A. (2014), "Detector generation algorithm for self-nonself detection in artificial immune system", *International Conference for Technology on Convergence of Technology (I2CT)*, pp. 1–6.

Received (Надійшла) 11.11.2019
Accepted for publication (Прийнята до друку) 27.11.2019

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Лисенко Сергій Миколайович – кандидат технічних наук, доцент кафедри комп'ютерної інженерії та системного програмування, Хмельницький національний університет, Хмельницький, Україна;
Sergii Lysenko – PhD, Associate Professor of Computer Engineering & System Programming Department, Khmelnytskyi National University, Khmelnytskyi, Ukraine;
e-mail: sirogyk@ukr.net; ORCID ID: http://orcid.org/0000-0001-7243-8747

Бобровнікова Кіра Юліївна – кандидат технічних наук, доцент кафедри комп'ютерної інженерії та системного програмування, Хмельницький національний університет, Хмельницький, Україна;
Kira Bobrovnikova – PhD, Associate Professor of Computer Engineering & System Programming Department, Khmelnytskyi National University, Khmelnytskyi, Ukraine;
e-mail: bobrovnikova.kira@gmail.com; ORCID ID: https://orcid.org/0000-0002-1046-893X

Харченко В'ячеслав Сергійович – доктор технічних наук, професор, завідувач кафедри комп'ютерних систем, мереж та кібербезпеки, Національний аерокосмічний університет «ХАІ», Харків, Україна;
Vyacheslav Kharchenko – Full Doctor, Full Professor, Head of Departments of Computer Systems, Networks and Cybersecurity, National Aerospace University "Kharkiv Aviation Institute", Kharkiv, Ukraine;
e-mail: v.kharchenko@csn.khai.edu; ORCID ID: https://orcid.org/0000-0001-5352-077X

Методы выявления бот-сетей в компьютерной системе

С. Н. Лысенко, К. Ю. Бобровникова, В. С. Харченко

Аннотация. **Объект**. Процесс обнаружения бот-сетей в корпоративных сетях на основе анализа сетевого трафика и поведения программного забезпечееня компьютерных системах. **Предмет**. Методы обнаружения бот-сетей в компьютерных системах. **Цель**. Повышение достоверности обнаружения бот-сетей путем разработки методов их обнаружения бот-сетей в корпоративных сетях. **Результаты**. Предложен новый подход к выявлению бот-сетей в корпоративных сетях на основе анализа поведения ботов. Обнаружение бот-сетей осуществляется путем применения разработанных двух методов: с помощью анализа на сетевом уровне и на хостовой уровне. Первый метод позволяет анализировать поведение программного обеспечения на хосте, что может указывать на возможное присутствие бота непосредственно на хосте и обнаружения вредоносного программного обеспечения, тогда как второй метод включает в себя мониторинг и анализ DNS-трафика, также позволяет сделать вывод об инфицировании сетевых хостов ботами бот-сети. На основе предложенных методов был разработан эффективный инструмент обнаружения бот-сетей - BotGRABBER. Он способен обнаруживать боты, которые используют такие методы уклонения от обнаружения: периодическая смена IP-отображения (cycling of IP mapping), «поток доменов» (domain flux), "быстро меняющиеся» сети (fast flux) и DNS-туннелирования (DNS tunneling). **Выводы**. Использование разработанной системы позволяет обнаруживать хосты, инфицированные ботами бот-сетей и локализовать вредоносные программы с высокой эффективностью - до 96%, а также демонстрирует низкий уровень ложных срабатываний - на уровне 3-5%. Особенность предлагаемого подхода состоит в том, что обнаружение бот-сетей является «невидимым» для владельцев бот-сетей.

Ключевые слова: бот; бот-сеть; DNS-трафик; поведение ботов бот-сетей; вредоносное программное обеспечение; искусственные иммунные системы; детекторы; положительный алгоритм отбора.

Methods for detecting bot nets in computer systems

S. Lysenko, K. Bobrovnikova, V. Kharchenko

Abstract. **Object**. The process of the botnets detection in the corporate area networks based on network traffic analysis and on the of computer systems software's behavior. **Subject**. Methods for botnets detection in computer systems. **Goal**. Increasing of the botnet detection efficiency by developing new methods for its detection in the corporate networks. **Results**. A new approach for the botnet detection in the corporate area networks based on the analysis of the bots' behavior is proposed. The detection of botnets is accomplished by applying the developed two methods: by means of network-level and host-level analysis. The first method allows you to analyze the behavior of the software on the host, which may indicate the possible presence of the bot directly on the host and the detection of malicious software, while the second method involves monitoring and analysis of DNS traffic, which also allows to make a conclusion about infection of network hosts with botnets. Based on the proposed methods, an effective tool for botnet detection - BotGRABBER - was developed. It is capable of detecting bots that use such evasion methods as IP mapping, fast flux, domain flux, and DNS tunneling. **Conclusions**. The usage of the developed system allows to detect the hosts infected with botnet and localize malware with high efficiency - up to 96%, and also shows low rate of false positives 3-5%. A feature of the proposed approach is that the detection of botnets is "invisible" to botnet owners.

Keywords: bot; botnet; DNS traffic; evasion technique; malware; artificial immune systems; detectors; positive selection algorithm.