УДК 004.056.5
DOI: 10.25140/2411-5363-2020-2(20)-175-185

Алла Гребенник, Олена Трунова, Володимир Казимир, Максим Міщенко

ВИЯВЛЕННЯ ТА ПРОГНОЗУВАННЯ РІВНЯ ЗАГРОЗ ДЛЯ КОРПОРАТИВНОЇ КОМП'ЮТЕРНОЇ МЕРЕЖІ

Актуальність теми дослідження. *У сучасному світі проблеми надійності інформації, що є в електронному вигляді, загострюють питання її захисту. Незважаючи на те, що в корпоративних комп'ютерних мережах, інформація більш ізольована від зовнішніх впливів, нелінійність та складність протікання процесів, а також інші загрози, що несе в собі не тільки зовнішній, а і внутрішній мережевий трафік, дає підстави до посилення контролю та аналізу мережевих потоків.*

Постановка проблеми*. Підвищення надійності функціонування комп'ютерних мереж залежить не тільки від вчасного виявлення загроз в її інформаційних потоках, а й від задіяння упереджувальних захисних заходів, які насамперед повинні спиратись на обґрунтовані прогнози виникнення шкідливих впливів. Обидві проблеми взаємопов'язані, оскільки для виконання прогнозування необхідно мати достатню та актуальну статистичну базу подій, що відбуваються в мережі.*

Аналіз останніх досліджень і публікацій. *Сучасні підходи до виявлення та прогнозування загроз для комп'ютерних мереж докладно проаналізовані в роботі [6]. Попри снування достатньої кількості методів і моделей вивчення та передбачення поведінки трафіку, найбільш застосовуваними при практичній реалізації залишаються моделі часових рядів.*

Виділення недосліджених частин загальної проблеми. *На сьогодні є багато інформаційних систем, метою яких є виявлення та запобігання мережевим атакам та аномаліям трафіку, більшість із них працює в реальному часі та надає інформацію про загрозу або вживає необхідних дій за фактом настання цієї загрози. Проте такі системи переважно побудовані на сигнатурному методі виявлення. Незважаючи на те, що в останні роки найбільш використовуваними в цій сфері є методи, пов'язані з машинним навчанням та інтелектуальним аналізом даних, більшість підходів мають лише теоретичну основу.*

Постановка завдання. *Враховуючи потребу в практичному застосуванні аномальних методів виявлення загроз інформації, було прийнято рішення про програмну реалізацію модулів інформаційної системи, які б виконували комплекс завдань збору, аналізу, моделювання розвитку подій у мережі, та були адаптовані до її типу та потреб.*

Виклад основного матеріалу. *Принципи функціонування систем виявлення загроз для комп'ютерної мережі базуються на розділенні загроз корпоративної комп'ютерної мережі на два основні класи – вторгнення в мережу та аномальна мережева поведінка. Системи, метою яких є виявлення та запобігання вторгненням базуються на використанні сигнатурних методів, а системи, що проводять аналіз аномальної мережевої поведінки, – на статистичному аналізі мережевого трафіку. У роботі обґрунтовано обрано та виконано програмну реалізацію алгоритмів двох адаптивних методів виявлення аномальної поведінки в потоках мережі, що вирішують зазначену проблему для різних, взаємовиключних умов, з використанням теорії Хаосу та EWMA-статистики. Одним із методів прогнозування стану або рівня загроз комп'ютерних мереж є Баєсова мережа, оскільки цей метод досить тісно пов'язаний із підходами, заснованими на графах атаки, це дозволяє не тільки прогнозувати рівень загроз для корпоративної комп'ютерної мережі, але й досліджувати послідовність їх виникнення, адресу джерела та призначення, тип загрози тощо.*

Висновки відповідно до статті. *Інформаційні технології визначення та забезпечення надійного рівня взаємодії суб'єктів комп'ютерних мереж є однією з актуальних проблем сучасного кіберсередовища. Проблема прогнозування загроз для корпоративної комп'ютерної мережі має менше існуючих рішень, ніж проблема виявлення та усунення загроз, але її вирішення дає змогу вживати завчасних дій до усунення та подальшого вивчення аномалій у мережевих потоках.*

Ключові слова: *комп'ютерна мережа; системи захисту від атак; теорія Хаосу; EWMA-статистики; Баєсова мережа.*

Рис.: 12. Бібл.: 11.

Актуальність теми дослідження. У сучасному світі питання кібербезпеки набуло досить великої актуальності, адже питома вага інформації, що є в електронному вигляді, зростає з кожним днем, гостро ставлячи питання про її захист. З трансформацією розподілених інформаційних систем (ІС) у окремі корпоративні комп'ютерні мережі (ККМ), інформація стала більш ізольованою від зовнішніх впливів, проте нелінійність та складність протікання процесів у таких мережах, а також інші загрози, що несе в собі не тільки зовнішній, а і внутрішній мережевий трафік, дає підстави до посилення контролю та аналізу мережевих потоків.

Постановка проблеми. Для вирішення цих питань необхідно чітко усвідомлювати базові принципи функціонування ККМ, розуміти існуючи рішення забезпечення надійності їх інформаційних потоків та удосконалювати алгоритми виявлення та прогнозування рівня загроз інформації ККМ. Здебільшого в кожній корпоративній мережі функціонує система виявлення атак (СВА), завданнями якої є забезпечення безпечності інформаційних потоків мережі. Від вирішення проблеми раннього виявлення та прогнозування аномальності подій у потоках з метою надання можливості вживання завчасних заходів для їх дослідження та усунення залежить якість СВА і надійність мережі загалом.

© Гребенник А. Г., Трунова О. В., Казимир В. В., Міщенко М. В., 2020

Аналіз останніх досліджень і публікацій. Сучасні підходи до виявлення та прогнозування загроз для комп'ютерних мереж докладно проаналізовані в роботі [6]. Незважаючи на існування достатньої кількості методів і моделей для реалізації поставленого завдання, найбільш застосовуваними, на сьогодні, при практичній реалізації залишаються моделі часових рядів. Методи моделювання та аналізу безпеки комп'ютерних мереж детально висвітлені в роботах [2; 3]. І незважаючи на те, що останніми роками найбільш використовуваними в цій сфері є методи, пов'язані з машинним навчанням та інтелектуальним аналізом даних, більшість підходів мають лише теоретичну основу.

Виділення недосліджених частин загальної проблеми. На сьогодні багато ІС, метою яких є виявлення та запобігання мережевим атакам та аномаліям трафіку, більшість із них працює в реальному часі та надає інформацію про загрозу або вживає необхідних дій за фактом настання цієї загрози. Проте такі системи переважно побудовані на сигнатурному методі виявлення, що налаштований на вже відомі типи порушень, і не враховують особливостей ККМ.

Із розвитком та поширенням методів штучного інтелекту активно почали розвиватися системи, які побудовані на адаптивних методах виявлення та прогнозування загроз ПЗ. Зокрема, у 2014 році Агентством передових оборонних дослідницьких проєктів США, було ініційовано створення класу ІС, побудованих на основі штучного інтелекту, призначених для знаходження, перевірки та виявлення кіберзагроз [9]. Цей клас систем отримав назву Cyber Reasoning Systems (англ. – системи кіберрозсудження) та активно розвивається у сфері кіберзмагань, зокрема Cyber Grand Challenge. Дані системи, працюючи в реальному часі, орієнтовані на висунення гіпотези про наявні загрози для досліджуваного ПЗ, перевірки цієї гіпотези та її підтвердження або відхилення. Проте нині системи такого класу є досить енергомісткими та не розраховані на виявлення та прогнозування рівня загроз для корпоративної комп'ютерної мережі [1].

Вирішення проблеми раннього виявлення та прогнозування загроз для ККМ надало б можливість спеціалісту з кібербезпеки завчасно вживати заходів до їх дослідження та усунення. Також залежно від обраних методів виявлення та прогнозування це допомогло б з певною точністю завчасно ідентифікувати наміри атакувальника, зокрема послідовність його дій та джерела можливих загроз.

Враховуючи все вищезазначене, було прийнято рішення про створення інформаційної системи, яка б здійснювала виявлення та прогнозування рівня загроз для корпоративної комп'ютерної мережі з використанням алгоритмів, що адаптуються до типу та потреб ККМ.

Постановка завдання. Враховуючи потребу в практичному застосуванні аномальних методів виявлення загроз інформації, було прийнято рішення про програмну реалізацію модулів інформаційної системи, які б виконували комплекс завдань збору, аналізу, моделювання розвитку подій у мережі, та були адаптовані до її типу та потреб.

Виклад основного матеріалу. Як загрозу для ККМ можна визначити будь-яку подію, яка призводить до аномальної поведінки або використання вразливостей ККМ, що, у свою чергу, тягне за собою переривання, втручання або знищення будь-якої цінної інформації, послуги або предмета, що належить корпорації-власниці комп'ютерної мережі (КМ).

Таким чином, принципи функціонування систем виявлення загроз для КМ (рис. 1) базуються на розділенні загроз для ККМ на два основні класи – вторгнення в мережу та аномальна мережева поведінка. Системи, метою яких є виявлення та запобігання вторгненням базуються на використанні сигнатурних методів, а системи, що проводять аналіз аномальної мережевої поведінки (Anomaly-Based Intrusion Detection and Prevention Systems) – на статистичному аналізі мережевого трафіку [2].

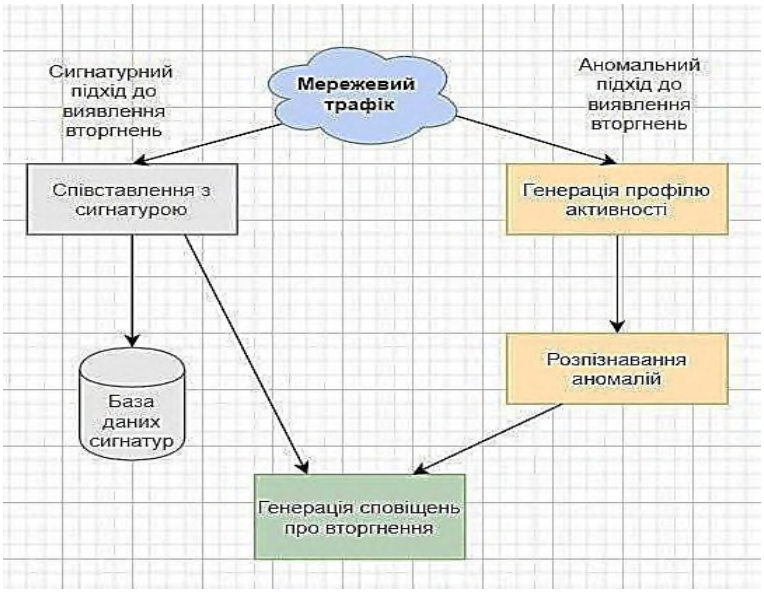

Рис. 1. Схематичний поділ принципів роботи систем виявлення загроз

Обидва підходи виявлення загроз мають свої переваги та недоліки.

До переваг сигнатурного методу виявлення можна віднести високу швидкість обробки та низький рівень хибних спрацювань для загроз, сигнатури яких присутні та оновлені. Це дозволяє цьому методу швидко та точно ідентифікувати підозрілі події. Основним недоліком сигнатурного методу виявлення загроз є неможливість виявити загрози нульового дня, тобто ті, для яких сигнатури відсутні або застарілі [4].

З іншого боку, перевагою методу аномалій є можливість виявлення ще невідомих атак та відсутність необхідності оновлення бази даних (БД) новими сигнатурами та правилами атак. Цей підхід полягає не в чіткому виявленні атак, а у визначенні підозрілої активності, відмінної від нормальної. Але це призводить до хибних спрацьовувань виявлення аномальної поведінки, яке не є атакою, або до пропуску атак, які не збігаються з визначенням аномальної поведінки.

Отже, для підвищення надійності та захищеності потоків мережі необхідно в системах захисту поєднувати обидва підходи, що й було вирішено зробити у створюваній у цій роботі інформаційній системі.

Архітектурною особливістю розроблюваної системи (рис. 2) є те, що ІС встановлюється окремо від досліджуваної мережі та серверу БД, що забезпечує автономність кожної частини системи. Це пов’язано з тим, що виявлення атак або аномальної поведінки вимагає постійної реєстрації всіх подій контрольованої мережі, завантажує обчислювальні потужності, вимагає великих обсягів дискового простору для зберігання зібраних даних та, як наслідок, знижує швидкодію роботи системи.

Як компонент для збору статистичної інформації про кількість та тип вторгнень у розроблюваній ІС використана мережева система виявлення вторгнень із відкритим кодом Snort. За рахунок цього вдасться підсилити компоненти виявлення аномалій мережевого трафіку та розширити базу знань про наявні загрози для корпоративної мережі.

Для обміну даними між мережею та сервером БД встановлено агент Telegraf, який поставляється у вигляді пакетного ПЗ для ОС FreeBSD, на основі якої працює мережевий екран pfSense.

Програмний мережевий екран PfSense, побудований на базі ядра FreeBSD, встановлюється на вузол, що фізично знаходиться одразу перед роутером. Основною перевагою цього рішення є наявність відкритого API, розробленого для роботи з PfSense – FauxAPI [5], що забезпечує доступ до даних трафіку, які проходить через мережевий екран.

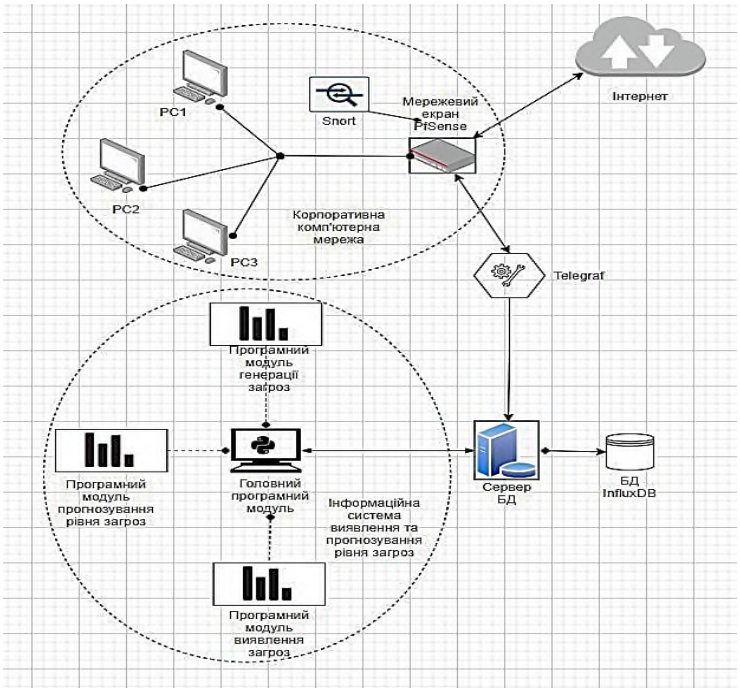

Рис. 2. Загальна архітектура системи прогнозування та виявлення загроз ККМ

Для формування бази знань, про виконані атаки на КМ, використовується IDS Snort, що встановлена як додатковий пакет для мережевого екрана pfSense, вона генерує оповіщення про здійснені атаки у форматі IDMEF (Intrusion Detection Message Exchange Format), записує їх до окремого лог-файла, дані з якого періодично записуються до БД за допомогою агента Telegraf.

У якості БД використано базу даних часових рядів InfluxDB, яка призначена для зберігання великих обсягів однотипних даних [8]. InfluxDB буде встановлена на окремий сервер, що знаходиться за межами ККМ, таким чином забезпечивши ізольованість даних та функціонування незалежно від внутрішніх збоїв мережі.

Сама ІС прогнозування та виявлення рівня загроз буде включати в себе модуль комунікації з InfluxDB для отримання даних трафіку з мережевого екрана ККМ, окремі модулі аналізу трафіку з метою виявлення аномалій трафіку та модуль передбачення рівня загроз, виведення отриманих статистичних даних і прогнозів. Для тестування роботи та збільшення обсягу даних аналізу використовується модуль для генерації аномального трафіку, розроблений із застосовуванням бібліотеки Scapy. Програмна частина ІС написана мовою програмування Python, ізольованість від досліджуваної мережі, дозволяє системі функціонувати незалежно від внутрішніх збоїв мережі.

Основним джерелом інформації про загрози для розроблюваної системи є мережевий трафік, який надходить до ККМ з глобальної мережі Інтернет. Таким чином, у якості вхідних даних для побудови прогнозу рівня загроз для корпоративної мережі, будемо використовувати набір кількісних та якісних параметрів трафіку, що проходить через мережу за одиницю часу.

Для виявлення аномалій мережевого трафіку в певний момент часу, необхідно сформувати часові ряди параметрів трафіку. Значення цих параметрів мають бути кількісною мірою трафіку, що проходить за одиницю часу, та надавати інформацію як про зовнішню взаємодію з мережею, так і відповіді мережі на здійснені взаємодії. Також обрані параметри для забезпечення можливості аналізу тенденцій та аномалій мережевого трафіку мають бути неперервними величинами.

Експертним шляхом були відібрані найважливіші, на наш погляд, параметри мережевого трафіку, що задовольняють заданим вище умовам. Серед цих параметрів: назва інтерфейсу, кількість байтів, що надійшло на вхід/вихід інтерфейсу за одиницю часу; кількість помилок на вході/виході та загальна.

Отже, розроблювана ІС у ролі вхідних даних приймає часові ряди показників мережевого трафіку (кількість вхідних/вихідних пакетів, байт) та часові ряди сповіщень про загрози у форматі IDEMF. На виході система надає інформацію про ймовірні загрози у графічному вигляді та у вигляді сповіщень, і зберігає її в бази даних.

У дослідженні були використані два адаптивні методи виявлення кіберзагроз, що вирішують зазначену проблему для різних, взаємовиключних умов, з використанням:

– теорії Хаосу;

– EWMA-статистики.

Розглядаючи комп'ютерну мережу як складну динамічну систему, процеси в якій протікають нелінійно, можемо зробити припущення про можливість застосування Теорії Хаосу для раннього виявлення нормального (хаотичного) та підозрілого (нехаотичного) перебігу процесів у ККМ.

Для виявлення хаотичності скористаємося підходом з обчисленням експоненти Ляпунова λ, що для потоку динамічної системи $F^t(x_0) = x_t$, визначається таким чином:

$$\lambda(x) = \lim_{t\to\infty} \frac{1}{t} \cdot \ln\left\| d_x F^t \right\|. \quad (1)$$

Загалом додатне значення експоненти вказує на хаотичну поведінку потоку, нульове – на незмінність поведінки, а від'ємне – на наявність нехаотичної поведінки.

У якості потоку даних, як зазначалося вище, системою використовується значення вхідного трафіку, а саме кількість отриманих пакетів за одиницю часу. Увесь трафік має бути розділений на дві вибірки: старий та новий трафік. Після розподілу виконується прогноз нових значень, що базується на вибірці старого трафіку. Обчислюючи помилки Δx_k для кожного з передбачень за формулою (1), будемо визначати поведінку їх зміни за допомогою експоненти Ляпунова $\Delta x_k = x_k - x_k^n$, де x_k^n – передбачене значення x_k.

Якщо значення експоненти буде додатним, тобто зміна помилки відбувається хаотично – трафік є нестабільним, що є властиво для ККМ, процеси в якій протікають нелінійно. При нульовому значенні експоненти, зміна помилки передбачення відсутня, а отже, і зміна трафіку відсутня. Якщо експонента набуває від'ємного значення, зміна помилки передбачення не є хаотичною. Отже, трафік стабілізувався і можна зробити висновок, що така зміна може бути викликана DDoS-подібною атакою.

Для тестування результатів роботи алгоритму було оброблено два типи трафіку: нормальний трафік та спричинений SYN і UDP flood атаками, та отримані відповідні їм експоненти Ляпунова (рис. 3-6).

Отже, цей метод дозволяє виявити атаки на ранньому етапі, що впливають на рівень мережевого трафіку, викликаючи різкий хаотичне зростання його показників. До таких атак можна віднести DDoS атаки, flooding та brute force атаки. Недоліком цього методу є неможливість його застосування для дальніх горизонтів прогнозування, оскільки відповідно до Теорії Хаосу, динаміка системи значною мірою залежить від початкових умов, що робить довгострокове прогнозування неможливим [11].

Метод з використанням EWMA-статистики полягає в застосуванні зваженого ковзкого середнього до часових рядів мережевого трафіку та визначення аномальних значень за формулою:

$$X_t \geq \beta\mu_{t-1}, \quad (2)$$

де X_t – значення параметру мережі (вхідного трафіку) в момент часу t;

β – коефіцієнт перевищення;

μ_{t-1} – середнє значення параметра за методикою EWMA для масиву даних до моменту часу t.

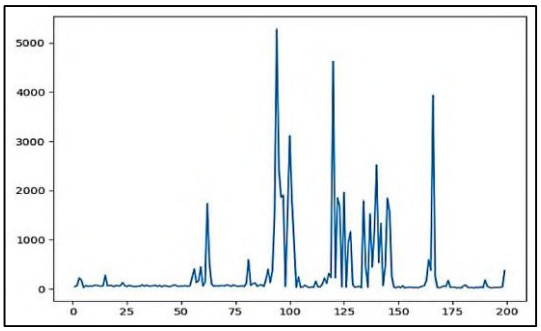

Рис. 3. Нормальний трафік

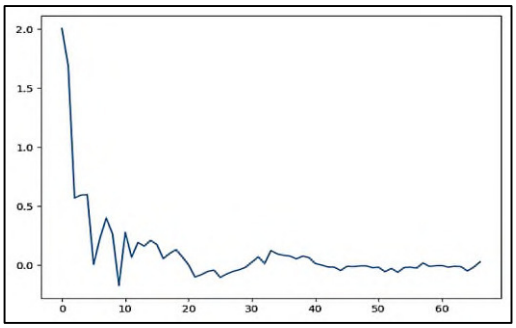

Рис. 4. Значення експоненти Ляпунова для нормального трафіку

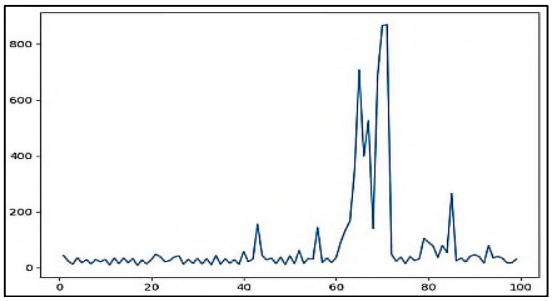

Рис. 5. Трафік з SYN та UDP flood

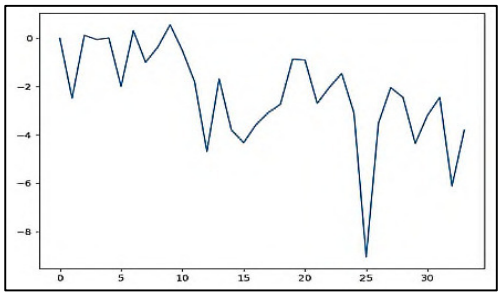

Рис. 6. Значення експоненти Ляпунова для SYN та UDP flood

На рис. 7-8 можна спостерігати графічне зображення застосування зазначеного методу виявлення аномалій до часових рядів зміни значень параметрів мережевого трафіку. У точках, де значення трафіку перевищує розрахункове значення ковзкої середньої, є підстави стверджувати про аномалії трафіку.

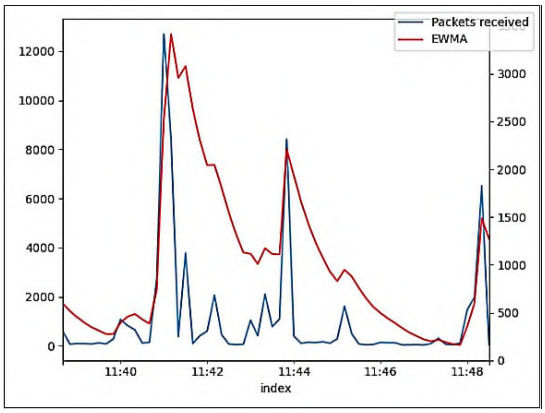

Рис. 7. «Кількість отриманих пакетів» вихідна та розрахована з EWMA

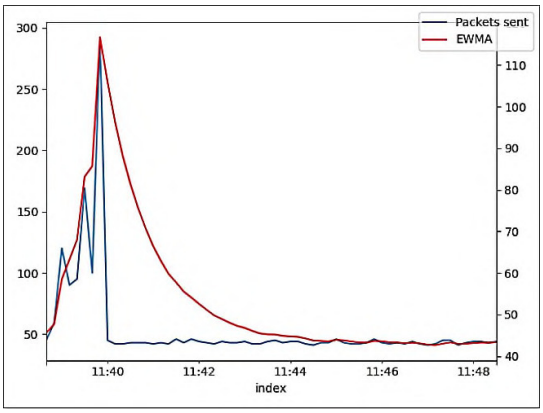

Рис. 8. «Кількість відправлених пакетів» вихідна та розрахована з EWMA

Відповідно до двох наведених вище адаптивних методів виявлення аномалій в інформаційній системі було розроблено програмні модулі ChaosDetector та EWMADetector. Для порівняння їх роботи було проведено тестування виявлення атак для часових проміжків 10, 20, 30, 40, 50, 60 хвилин спостереження значень вхідного трафіку. Як параметр було обрано кількість вхідних пакетів за одиницю часу. Для кожного з алгоритмів порівнювався час їх виконання та кількість виявлених аномалій. Результати порівняння зображені на рис. 9.

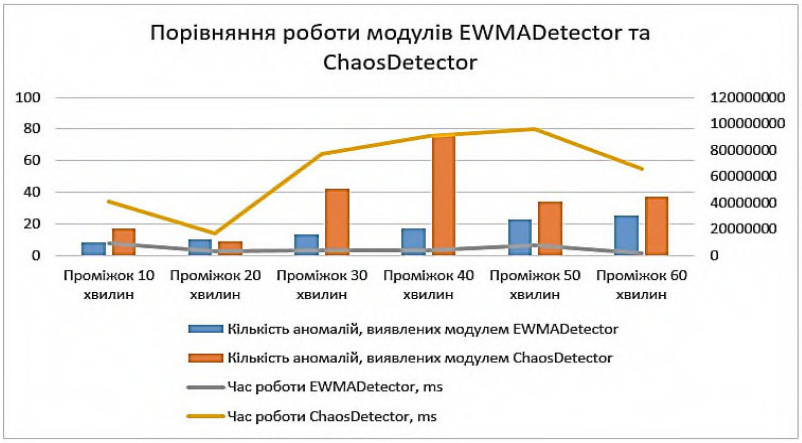

Рис. 9. Порівняння роботи модулів EWMADetector та ChaosDetector

Як бачимо з графіка, різниця в кількості аномалій, виявлених детекторами, зростає зі збільшенням часового проміжку аналізу даних. Це пояснюється тим, що виявлення аномалій на основі Теорії Хаосу не призначене для використання з далекими горизонтами прогнозування й може давати хибні результати зі збільшенням горизонту.

Разом з тим час роботи детектора, побудованого на основі використання EWMA статистики, значно менший, ніж час детектора, побудованого на використанні Теорії Хаосу. Це пояснюється складністю обчислень модулю ChaosDetector, а саме необхідністю для кожної ітерації робити прогнози наступного значення та вираховувати значення експоненти Ляпунова для отриманих помилок.

Можна зробити висновок, що оптимальний проміжок часу для виявлення аномалій мережевого трафіку становить 20 хвилин, адже для нього кількість аномалій, виявлених обома детекторами, найбільш близька за значеннями, а їх час роботи один із найменших серед усіх інтервалів.

Крім моніторингу трафіку, до завдань мережевого адміністрування треба віднести завдання прогнозування стану або рівня загроз комп'ютерних мереж.

Одним із найбільш гнучких та точних методів прогнозування є прогнозування рівня загроз за допомогою ймовірнісних мереж Байєса. Цей метод досить тісно пов'язаний із підходами, заснованими на графах атаки [8]: Баєсова мережа переважно побудована на основі графа атак. Відмінна риса Баєсових мереж – це умовні змінні та ймовірності, які відображені в моделі.

Баєсова мережа – це імовірнісна графова модель, яка складається зі змінних та зв'язків між ними. Ця мережа являє собою спрямований ациклічний граф з вузлами, представленими дискретними або неперервними змінними, та ребрами, що відображають зв'язки між ними. Вузли утримують стани випадкових величин та форму умовної ймовірності [7].

Для набору випадкових змінних $X = \{x_1, \dots, x_n\}$ у мережі Баєса, функція спільної щільності ймовірності визначається за формулою 3, де $P_a(x_i)$ представляє відповідне значення ймовірності змінних у батьківських вузлах мережі, а $P(x_i/Pa(x_i))$ – умовна ймовірність у дочірніх вузлах.

$$P(x_1, \dots, x_n) = \prod_{i=1}^{n} P\left(\frac{x_i}{Pa(x_i)}\right). \quad (3)$$

У Баєсових мережах, ймовірності зв'язків оновлюються у міру надходження нової інформації за допомогою теореми Баєса. Таким чином, з появою нових загроз для мережі, побудована мережа Баєса оновлюється та надає актуальні прогнози про рівні загроз.

Програмну реалізацію прогнозування рівня загроз з використанням мереж Баєса необхідно розділити на дві частини:

1. «Офлайн режим» – побудова мережі Баєса на основі сповіщень, отриманих від встановлених IDS систем та передбачення загалом рівня загроз для системи.

2. «Онлайн-режим» – передбачення ймовірності майбутніх загроз, шляхом подачі на вхід побудованої мережі Баєса інформації про нові загрози.

На вхід цього алгоритму подаються сповіщення системи IDS у форматі IDMEF, що включають у себе як мінімум такі поля:

- Time – час створення сповіщення;
- AlertType – тип сповіщення;
- Source – адреса джерела атаки;
- Destination – адреса призначення атаки.

Сповіщення агрегуються за їхнім типом для послідовних часових значень їх створення (формула (4)):

$$A_{t1}[AlertType] = A_{t2}[AlertType] = \cdots = A_{tn}[AlertType], \quad (4)$$

де $t_1, \dots, t_n$ – послідовні часові значення створення сповіщень.

Для агрегованого сповіщення A_0 типу $AlertType_A$, зберігаються значення адреси джерела атаки та адреси призначення атаки кожного зі сповіщень A_t.

Після агрегації сповіщень за типом, виконується обчислення кореляції між двома послідовними агрегованими сповіщеннями A_0 типу $AlertType_A$ та B_0 типу $AlertType_B$. Сповіщення про загрози A_t і B_t вважаються корельованими, якщо виконується одна з двох умов:

1)$\{A_t[Source] = B_t[Source], A_t[Destination] = B_t[Destination]\}$;

2)$\{A_t[Destination] = B_t[Source]\}$.

Для кожної пари корельованих сповіщень обчислюються ймовірності їх виникнення за формулою:

$$P(B/A) = \frac{Corr(A,B)}{Corr\ (A,*)}, \quad (5)$$

де $Corr(A,B)$ – кількість зв'язків між сповіщеннями A та B,

$Corr\ (A,*)$ – загальна кількість зв'язків, спричинена сповіщенням A (може бути виражена у кількості *Source* адрес для сповіщення A).

У результуючій Баєсовій мережі вузлами виступають агреговані сповіщення, а ребрами між ними – обчисленні ймовірності їх виникнення.

Створені за вказаним алгоритмом мережі зображені на рис. 10, 11.

Рис. 10. Баєсова мережа сповіщень про загрози для періоду 3 дні

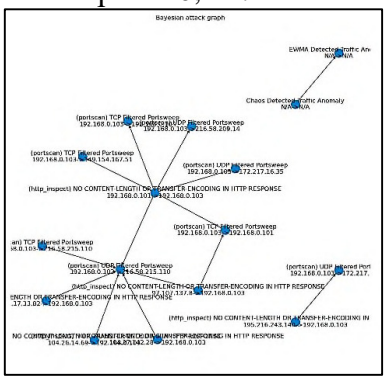

Рис. 11. Баєсова мережа сповіщень про загрози для періоду 7 днів

Відзначимо, що кількість агрегованих сповіщень про загрози зростає зі збільшенням періоду агрегації. Баєсова мережа добудовується, як тільки надходить інформація про нову загрозу.

Отримана в офлайн режимі Баєсова мережа використовується в онлайн-режимі для прогнозування рівня загроз за обраний проміжок часу t. Для цього, крім періоду, для якого створена мережа, передається назва останнього сповіщення про загрозу з БД.

У результаті обчислень умовних ймовірностей змінних, алгоритм повертає ймовірності появи усіх можливих загроз, що виступають змінними в побудованій мережі Баєса. Приклад виводу реалізованого програмного модуля наведено на рис. 12.

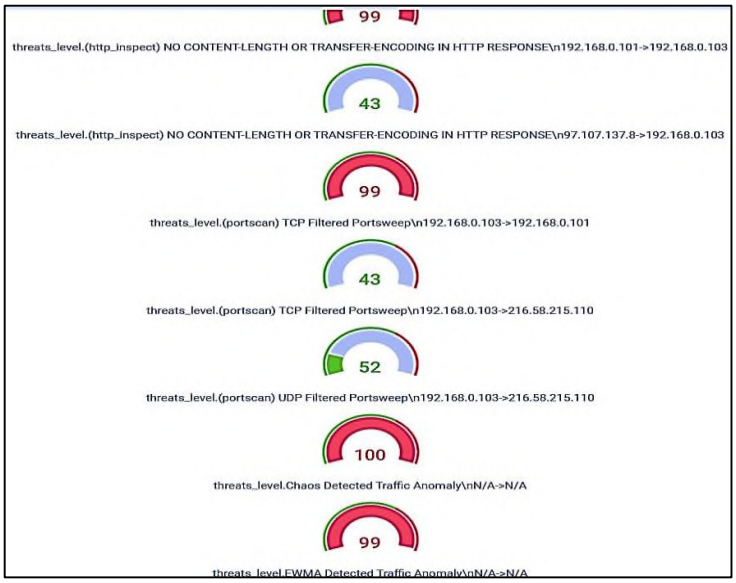

Рис. 12. Результат прогнозування рівня загроз на основі Баєсової мережі, побудованої за період 3 дні

Останньою загрозою в наведеному прикладі стала аномалія, виявлена за допомогою теорії Хаосу. Оскільки вона зафіксована та записана до БД модулем виявлення загроз, ймовірність її появи дорівнює 100 %. Відповідно, вона була передана на вхід алгоритму для прогнозування рівня загроз, а алгоритм, у свою чергу, вивів ймовірності появи всіх інших загроз, наявних у побудованій Баєсовій мережі, що обчислюються за формулою умовних ймовірностей.

Таким чином, розглянутий метод дозволяє не тільки прогнозувати рівень загроз для ККМ, але й досліджувати послідовність їх виникнення, адресу джерела та призначення, тип загрози тощо.

Для моніторингу станів трафіку та аномалій застосовується створена статистична модель, а для прогнозування рівня загроз – мережа Баєса. Алгоритм є адаптивним, оскільки його робота не залежить від конфігурації ККМ.

Висновки відповідно до статті. Інформаційні технології визначення та забезпечення надійного рівня взаємодії суб'єктів комп'ютерних мереж є однією з актуальних проблем сучасного кіберсередовища. Проблема прогнозування загроз для ККМ має менше існуючих рішень, ніж проблема виявлення та усунення загроз.

Прогнозування рівня загроз для ККМ дає змогу вживати завчасних дій до їх усунення та подальшого вивчення.

Реалізований алгоритм прогнозування рівня загроз за допомогою мереж Баєса надає змогу спеціалістам з кібербезпеки досліджувати їх джерела та послідовність виникнення.

Майбутні удосконалення розробленої інформаційної системи вбачаються в розширенні модулів виявлення та прогнозування атак новими алгоритмами, застосування хмарної інфраструктури для розгортання та тестування модулів системи, тестування роботи системи на високонавантажених комп'ютерних мережах.

Подяка. *Робота проведена та фінансована в межах проєкту HATO CyRADARS (Cyber Rapid Analysis for Defense Awareness of Real-time Situation – CyRADARS) – grant agreement number: G5286.*

Список використаних джерел

1. Cyber Reasoning Systems: Automating Cyber Warfare. Medium. 2016. URL: https://medium.com/@joey_rideout/cyber-reasoning-systems-automating-cyber-warfare-3329f339edeb.

2. Моделювання та аналіз безпеки розподілених систем : навч. посіб. [для студ. спец. 121 «Інженерія програмного забезпечення»] / В. В. Литвинов та ін. Чернігів : Чернігів. нац. технол. ун-т, 2016. 254 с.

3. Методи аналізу та моделювання безпеки розподілених інформаційних систем : монографія / В. В. Литвинов та ін. ; за заг. ред. проф. С. М. Шкарлета. Чернігів : Чернігів. нац. технол. ун-т. 2017. 206 с.

4. Cybersecurity Spotlight – Signature-Based vs Anomaly-Based Detection. URL: https://www.cisecurity.org/spotlight/cybersecurity-spotlight-signature-based-vs-anomaly-based-detection.

5. FauxAPI – v1.3. URL:https://github.com/ndejong/pfsense_fauxapi.

6. Husák M., Komárková J., Bou-Harb E., Čeleda P.Survey of Attack Projection, Prediction, and Forecasting in Cyber 5. Security. *IEEE Communications Surveys Tutorials*. September 2018. Vol. 21, No. 1. P. 640-660, URL: https://is.muni.cz/repo/1434138/2019-COMST-survey-of-attack-projection-prediction-forecasting.pdf.

7. Husák M. Predictions of Network Attacks in Collaborative Environment (PhD Dissertation). Brno, 2019. 144 c.

8. Sheyner O., Haines J., Jha S. Lippmann R., J. M. Wing J.M. Automated generation and analysis of attack graphs. In: *Proceedings of the 2002 IEEE Symposium on* Security *and* Privacy *(S&P'02)*. 2002. P. 273-284.

9. Avgerinos T., Brumley D., Davis J. and etc. The Mayhem Cyber Reasoning System. *Security&Privacy*. 2018. P. 52-60. URL: http://users.umiacs.umd.edu/~tdumitra/courses/ENEE657/Fall19/papers/Avgerinos18.pdf.

10. What infrastructure and application monitoring can solve for you: URL: https://www.influxdata.com/customers/infrastructure-and-application-monitoring.

11. Мартінзон О. С., Грабар О. І. Теорія хаосу: URL: https://conf.ztu.edu.ua/wp-content/uploads/2017/06/139-2.pdf.

References

1. Cyber Reasoning Systems: Automating Cyber Warfare. Medium. (2016). Retrieved from https://medium.com/@joey_rideout/cyber-reasoning-systems-automating-cyber-warfare-3329f339edeb.

2. Lytvynov, V. V., Kazymyr, V. V., Stetsenko, I. V., Trunova, O. V., Skiter, I. S. … Nekhai, V. V. (2016). *Modeliuvannia ta analiz bezpeky rozpodilenykh system [Modeling and analysis of safety-related systems]*. Chernihiv: ChNTU [in Ukrainian].

3. Shkarlet, S. M. (Ed.). (2017). *Metody analizu ta modeliuvannia bezpeky rozpodilenykh informatsiinykh system [Methods of analysis and modeling of safety-related information systems]*. Chernihiv: ChNTU [in Ukrainian].

4. Cybersecurity Spotlight – Signature-Based vs Anomaly-Based Detection. Retrieved from https://www.cisecurity.org/spotlight/cybersecurity-spotlight-signature-based-vs-anomaly-based-detection.

5. FauxAPI – v1.3: Retrieved from https://github.com/ndejong/pfsense_fauxapi.

6. Husák, M., Komárková, J., Bou-Harb, E., Čeleda, P. (September, 2018). Survey of Attack Projection, Prediction, and Forecasting in Cyber 5. Security. *IEEE Communications Surveys Tutorials,* 21, 1. Retrieved from https://is.muni.cz/repo/1434138/2019-COMST-survey-of-attack-projection-prediction-forecasting.pdf.

7. Husák, M. (2019). *Predictions of Network Attacks in Collaborative Environment* (PhD dissertation). Brno.

8. Sheyner, O., Haines, J., Jha, S. Lippmann, R., J. M. Wing, J. M. (2002). Automated generation and analysis of attack graphs. In: *Proceedings of the 2002 IEEE Symposium on* Security *and* Privacy (S&P'02) (pp. 273-284).

9. Avgerinos, T., Brumley, D., Davis, J. and etc. (2018). The Mayhem Cyber Reasoning System. Security&Privacy. Retrieved from http://users.umiacs.umd.edu/~tdumitra/courses/ENEE657/Fall19/papers/Avgerinos18.pdf.

10. What infrastructure and application monitoring can solve for you. Retrieved from https://www.influxdata.com/customers/infrastructure-and-application-monitoring/

11. Martinzon, O. S., Hrabar, O. I. (2017). T*eoriia khaosu [Chaos theory]*. Retrieved from https://conf.ztu.edu.ua/wp-content/uploads/2017/06/139-2.pdf.

UDC 004.056.5

Alla Grebennyk, Olena Trunova, Volodymyr Kazimir, Maksym Mishchenko

DETECTION AND FORECASTING OF THE THREAT LEVEL FOR A CORPORATE COMPUTER NETWORK

Urgency of the research. *In today's world, the question of the reliability of electronic information exacerbates the issue of its protection. Despite the fact, that in corporate computer networks information is more isolated from external influences, the nonlinearity and complexity of processes in such networks, as well as a number of threats posed not only by external but also internal network traffic, gives grounds to strengthen control and network flow analysis.*

Target setting. *Improving the operational reliability of computer networks depends not only on the timely detection of threats in its information flows, but also on the use of preventive measures, which, above all, should be based on reasonable forecasts of harmful effects. Both problems are interrelated, as forecasting requires a sufficient and up-to-date statistical database of events, which happen in the network.*

Actual scientific researches and issues analysis. *Modern approaches to the detection and prediction of threats to computer networks are analyzed in detail in [6]. Despite the existence of a sufficient number of methods and models for the implementation of the task, time series models are the most used in practical realization.*

Uninvestigated parts of general matters defining*. Currently, there are a number of information systems designed to detect and prevent network attacks and traffic anomalies, most of which work in real time and provide information about the threat or take the necessary action upon the occurrence of this danger. However, such systems are mainly based on the signature detection method. Although the most widely used methods in recent years have been machine learning and data mining, most approaches have only a theoretical basis.*

The research objective. *Taking into account the need for practical application of anomalous methods of detecting information threats, it was decided to create the program implementation of information system modules that perform a set of tasks of collecting, analyzing, modeling network events, and were adapted to its type and needs.*

The statement of basic materials. *The principles of operation of threat detection systems for a computer network are based on the division of dangers to the corporate computer network into two main classes - network intrusion and abnormal network behavior. Systems, which are based on the detecting and preventing intrusions, use of signature methods, and systems that analyze abnormal network behavior - based on statistical analysis of network traffic. Among such methods, the software implementation of algorithms of two adaptive methods for detecting anomalous behavior in network flows, which solve this problem for different mutually exclusive conditions using Chaos theory and EWMA-statistics, is selected and performed. Bayesian network is one of those methods for predicting the state or level of dangers to computer networks, because this method is closely related to approaches based on attack graphs: a distinctive feature of Bayesian networks is the conditional variables and probabilities that are reflected in the model. This allows not only predicting the level of threats to the corporate computer network, but also investigating the sequence of their occurrence, source and destination, type of threat, and so on.*

Conclusions. *Information technologies for determining and ensuring a reliable level of interaction between the subjects of computer networks is one of the urgent problems of the modern cyber environment. The problem of predicting dangers to the corporate computer network has less existing solutions than the problem of detecting and eliminating threats, but its solution allows you to take early action to eliminate and further study the anomalies in network flows.*

Keywords: computer network, attack protection systems, Chaos theory, EWMA statistics, Bayesian network.

Fig.: 12. References: 11.

Гребенник Алла Григорівна – аспірант, Інститут проблем математичних машин і систем НАН України (просп. Глушкова, 42, м. Київ, 03187, Україна).
Hrebennyk Alla – PhD student, the Institute of Mathematical Machines and Systems Problems National Academy of Science of Ukraine (42 Academician Glushkova Av., 03187 Kyiv, Ukraine).
E-mail: grebennik.alla@gmail.com
ORCID: http://orcid.org/0000-0002-7464-1412

Трунова Олена Василівна – кандидат педагогічних наук, доцент, доцент кафедри інформаційних технологій та програмної інженерії, Чернігівський національний технологічний університет (вул. Шевченка, 95, м. Чернігів, 14035, Україна).
Trunova Olena – PhD in Pedagogical Sciences, Assistant Professor, Assistant Professor of Department of Information Technology and Software Engineering, Chernihiv National University of Technology (95 Shevchenka Str., 14035 Chernihiv, Ukraine).
E-mail: e.trunova@gmail.com
ORCID: http://orcid.org/0000-0003-0689-8846

Казимир Володимир Вікторович – доктор технічних наук, професор, професор кафедри інформаційних та комп'ютерних систем, Чернігівський національний технологічний університет (вул. Шевченка, 95, м. Чернігів, 14035, Україна)
Kazymyr Volodymyr – Doctor of Technical Sciences, Professor, Professor of Information and Computer Systems. Chernihiv National University of Technology (95 Shevchenka Str., 14035 Chernihiv, Ukraine).
E-mail: vvkazymyr@gmail.com
ORCID: http://orcid.org/0000-0001-8163-1119
Scopus Author ID: 56644727300

Міщенко Максим Валерійович – студент магістратури, Чернігівський національний технологічний університет (вул. Шевченка, 95, м. Чернігів, 14035, Україна).
Mishenko Maxim – master, Chernihiv National University of Technology (95 Shevchenka Str., 14035 Chernihiv, Ukraine).
E-mail: mak22101996@gmail.com

Гребенник А., Трунова О., Казимир В., Міщенко М. Виявлення та прогнозування рівня загроз для корпоративної комп'ютерної мережі. *Технічні науки та технології.* 2020. № 2 (20). С. 175-185.