DOI 10.34185/1562-9945-5-148-2023-04
УДК 519.2:004.9

Т.О. Гулий, В.Є. Білозьоров

ВИКОРИСТАННЯ МЕТОДУ НЕЛІНІЙНОГО РЕКУРЕНТНОГО АНАЛІЗУ ДО ПОШУКУ DDOS АНОМАЛІЙ ЧАСОВИХ РЯДІВ МЕРЕЖЕВОГО ТРАФІКУ

Анотація. У статті розглядається питання використання методу нелінійного рекурентного аналізу до проблеми пошуку аномалій у мережевому трафіку, що надана у вигляді даних часових рядів знімку трафіку за певний період часу. Описано методика визначення якісної характеристики для вхідних даних та її використання для побудови відповідної рекурентної діаграми (recurrence plots, RP). В якості яких було запропоновано узяти статистку кількості отриманих байтів у секунжу. Для отримання чисельних значень показників RP пропонується використовувати середовище Матлаб та розроблений для цього пакет crptools. Наведені обраховані показники RP дозволили здійснити типізацію отриманих даних та визначити тип якій отримав назву «DDOS RP», що надає змогу виділити деякі види атак типу DoS/DDoS.
Ключові слова: рекурентний аналіз, мережевий трафік, часовий ряд, рекурентна діаграма, параметр затримки, розмірність простору вкладення, RQA аналіз, середовище Матлаб.

Вступ і мета. Виявлення аномалій — це проблема пошуку шкідливих шаблонів у даних, які відрізняються від нормальної поведінки. У комп'ютерних мережах така ненормальна поведінка зазвичай виникає через такі фактори, як відмова в обслуговуванні (DoS), несправність мережевих пристроїв, неправильна конфігурація маршрутизатора, перевантаження трафіку, збій компонентів і вторгнення в мережу.

Відхилення мережевого трафіку від нормальної поведінки відбувається через: зміни шляху та петель маршрутизації, різкі зміни потоку трафіку, варіації затримок трафіку. Ці властивості визначаються та класифікуються шляхом спостереження за потоком трафіку, затримкою пакетів, загальною кількістю надісланих пакетів, співвідношенням байтів, надісланих у кожному напрямку, тривалістю з'єднання, середнім розміром пакетів і часом між надходженнями. Поточне виявлення аномалії спостерігає за статистикою транспортного рівня, і результати порівнюються з пороговим значенням. Проте спостерігається, що характеристики мережевого трафіку, характеристики розподілу ймовірностей

IP-пакетів, щільність трафіку, статистичні характеристики трафіку динамічно змінюються у часовій області.

Нещодавно було запропоновано повторний нелінійний рекурентний аналіз (RQA) для спостереження та вивчення цих нестаціонарних властивостей [1]. З часом, з'явились інші роботи, які виконували аналіз на нестандартизованих даних, та надавали різну до знаходження нових підходів до пошуку аномалії [2,3]. Деякі роботи використовують у якості вхідних данних згорткової нейронної мережі [4]. Проблема досі є досить складною і актуальною.

Мета данної роботи – запропонувати підхід реалізації пошуку аномалій для атаки сервісу типу DDoS/DoS даних мережевого трафіку за допомогою чисельних показників нелінійного рекурентного аналізу.

Огляд проблеми та її аналіз. Основна проблема аналізу часового ряду полягає в тому, що для побудови рекурентної діаграми (RP) необхідно чіткого визначення прихованих параметрів m і τ. m -це розмірність простору вкладення часового ряду, для якого виконується аналіз, а τ - параметр його затримки. До визначення даних параметрів деякі дослідники [2-4] поставилися суто формально, що є помилковим з точки зору синергетики, а тим більше нелінійної динаміки.

У дослідженні [6] представили хороше опитування щодо виявлення аномалій. Опитування щодо методів виявлення аномалій у сферах машинного навчання та статистики можна знайти в [7]. Для виявлення аномалій мережі доступно декілька інструментів. Деякі з цих інструментів використовують чітко визначені правила для аналізу моделей трафіку. Кілька інших методів включають експоненціальне згладжування та прогноз Холта Вінтерса, адаптивну порогову кумулятивну суму, оцінку максимальної ентропії.

Багато з цих детекторів працюють на основі статистики транспортного рівня, включаючи співвідношення байтів, надісланих у кожному напрямку, середній розмір і середній час між надходженнями пакетів. Далі результати порівнюються з пороговими значеннями. Ці значення не залежать від використання мережі та кількості користувачів. Мережеву аномалію можна виявити за допомогою інших моделей, включаючи байєсівську, опорну векторну машину, нейронні мережі тощо. У літературі для вирішення проблеми класифікації мережевого трафіку застосовано кілька методів аналізу даних. [8] запропонував виявлення аномалій мережевого трафіку на основі байтів пакетів. У даній роботі ми запропонуванували інші види візуалізації отриманої ними інформації: рекуре-

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

нтні діаграми. Тому має сенс спрямувати зусилля на більш детальне дослідження параметрів m і τ, а також чисельних показників RP.

У роботі використовувалося програмне забезпечення, розроблене і реалізоване у середовищі типу Матлаб (Октавіа) – crptool [9] під ОС Windows. Для дослідження часових рядів було використано набір даних оцінки методології пошуку аномалій CIC-IDS2017 [5].

Опис даних мережевого трафіку. CIC-IDS2017 набір даних було захоплено з 3 червня 2017 року по 7 червня 2017. Впродовж 5 днів у лабораторних умовах було відтворено ряд базових втручань у внутрішню систему зокрема: веб-атаки, проникнення, ботнет і DDoS/DOS.

У експерименті приймали участь 25 комп'ютерів які симулювали стандартну роботу офісну (пересилання документів, поштові операції, пошук інформації у Google, тощо) і два сервери які працювали як стандартні файлообмінники. На сервери середовища було здійснено атаку DoS/DDOS знято статистку у наступних часових проміжках:

1. 05.07.2017 з 9:47 по 10:10 було здійснено атаку DoS slowloris з адреси 205.174.165.73 на сервер 192.168.10.50
2. 05.07.2017 з 10:14 по 10:35 було здійснено атаку DoS Slowhttptest з адреси 205.174.165.73 на сервер 192.168.10.50
3. 05.07.2017 з 10:43 по 11:00 було здійснено атаку DoS Hulk з адреси 205.174.165.73 на сервер 192.168.10.50
4. 05.07.2017 з 11:10 по 11:23 було здійснено атаку DoS GoldenEye з адреси 205.174.165.73 на сервер 192.168.10.50
5. 07.07.2017 з 15:56 по 16:16 було здійснено атаку DDoS LOIT з адресс 205.174.165.69, 205.174.165.70 205.174.165.71 на сервер 192.168.10.50

Окрім даних, які містять аномально активність, було також зібрано 5 проміжків по 10 хвилин з даних зібраних у понеділок 03.07.2017, які містять схожу за динамікою активність, але не містять втручання. Така динаміка може бути викликана на початку робочого дня, під час роботи програм моніторингу активності або масштабної передачі даних.

Інформація була захоплена за допомогою програмного забезпечення Wireshark. За допомогою даної програми також було зібрано статистику по кількості пакетів отриманих сервером з інтервалом у одну секунду.

Проблеми пошуку параметрів. Задача пошуку аномалій у часових рядах полягає у виявленні незвичайних та несподіваних змін у поведінці системи, яка описується часовим рядом. Аномалії можуть вказувати на виникнення

проблем або надзвичайних подій, що потребують уваги та відповідного реагування. У випадку з відхиленнями у даних мережевого трафіку технічний спеціаліст може зупинити атаку шляхом фільтрації запитів з певних адрес, або проводити системні відключення від системи заражених користувачів з подальшим усуненням проблеми.

Одним з підходів до пошуку аномалій є порівняння між очікуваними значеннями та фактичними значеннями часового ряду. Якщо спостерігається відхилення, то це може вказувати на наявність аномалії. Інші підходи включають виявлення змін точок перелому та застосування різноманітних методів машинного навчання, таких як ансамблеві методи, навчання з учителем та навчання без учителя[].

Нажаль, зі збільшенням систем пошук аномалій може займати все більше інформаційного ресурсу. Моделі засновані на принципах математичної статистики в реальних умовах обмежені у часових, обчислювальних і матеріальних ресурсів. Отже необхідні методології з використанням технологій інтелектуального аналізу даних.

Отже, маємо наступний ряд задач: визначити зі знятих реальних даних мережевого трафіку відхилення у числових показниках, які вказуватимуть на наявність втручання та аномальних процесів у системі; проаналізувати дані показники для того, щоб відокремити для пошуку атаки класу DoS/DDoS.

У якості математичного апарату було обрано нелінійний рекурентний аналіз. Сучасні методології поушку аномалій будуються на нейронних диференційних рівняннях (NeuralODE)[]. Застосування RQA аналізу може бути використаним є цікавим і перспективним саме для таких систем. Запропонована наступна методика обробки реальної інформації мережевих потоків по кожному з інцидентів втручання, яка складалася з наступних кроків:

1. Надання графіку статистики кількості отриманих даних сервером за одну секунду ($Rate$) у нормальних та аномальних .

2. Обрахування значень параметра затримки τ.

3. Визначення значення розмірності простору вкладення m (для min та max-норми) з урахуванням вже визначених у пункті 2 значень параметра затримки τ.

4. Побудова RP діаграми для визначених параметрів τ та m. Значення ε обирається не більше 20%.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

5. Розрахунок для визначених значень τ та m чисельних показників RP діаграми - RQA аналіз.

6. Отримати усі значення чисельних показників RP діаграми, визначити якісні характеристики за якими можна виявляти втручання та аномальні процеси типу атак DoS/DDoS.

Визначення чисельних показників RP діаграми. Графік рекурентності - це квадратна матриця, яка є симетричною відносно головної діагоналі та містить час виникнення станів як у стовпцях, так і в рядках. Кожен елемент матриці відповідає певній парі часів та містить 1, якщо стан повторюється, та 0, якщо ні.

Іншим визначенням графіку повторення є $N \times N$ матриця, що складається лише з чорних та білих точок з позначенням, що чорна точка відображає повторення, разом з двома осями часу. Математично, графік повторності можна виразити наступним рівнянням.:

$$R_{i,j} = \theta\left(\varepsilon_i - \left\|\vec{x_i} - \vec{x}_j\right\|\right), \quad \vec{x}_j \in \mathbb{R}^m \qquad i,j = 1, \ldots, N$$

де розглядаються стани $\mathbf{x}_i$ мають кількість N; ε_i - порогове значення відстані (околиця); $\|.\|$ - норма; θ $(\cdot)$ – функція Хевісайду. Коли відстань між двома станами, тобто $\mathbf{x}_i$ та $\mathbf{x}_j$, менша за порогове значення ε , визначається повторення.

Для показників, які використовуємо, представимо їх визначення, що надано у рекурентному аналізі. Для обчислення характеристик, які використовуються у RP, значення порога ε є фіксованими.

Міра рекурентності (recurrence rate, RR)

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j}^{m,\varepsilon}$$

показує щільність рекурентних точок, просто підраховуючи їх. N - кількість точок на траєкторії у фазовому просторі. Дана міра показує можливість знаходження рекурентної точки в RP (ймовірність повторення стану).

Наступна міра розглядає діагональні лінії. Частотний розподіл довжин l діагональних ліній в RP $P^{\varepsilon}(l) = \{l_i; i = 1, \ldots, N_l\}$, де N_l - абсолютна кількість діагональних ліній (кожна лінія надається тільки один раз). Процеси зі стохастичною поведінкою можуть породжувати дуже короткі діагоналі чи взагалі не породжувати їх, тоді як детерміністські процеси дають довгі діагоналі і малу кількість окремих рекурентних точок.

Отже, відношення рекурентних точок

$$DET = \frac{\sum_{l=l_{min}}^{N} lP^{\varepsilon}(l)}{\sum_{i,j=1}^{N} R_{i,j}^{m,\varepsilon}},$$

називається *мірою детермінізму (determinism, DET)* або передбачуваності системи. Слід зазначити, що цей захід не має значення реального детермінізму процесу. Порогове значення $l_{\min}$ виключає діагональні лінії, утворені тангенціальним рухом траєкторії у фазовому просторі. Якщо $l_{min} = 1$, *DET=RR*. Більше значення детермінізму вказує більш діагональну лінію в RP і, отже, більш сильну передбачуваність системи.

Діагональні структури показують час, протягом якого ділянка траєкторії підходить досить близько до іншої ділянки траєкторії. Таким чином, ці лінії дозволяють судити про розбіжність елементів траєкторій.

Середня довжина діагональних ліній

$$L = \frac{\sum_{l=l_{min}}^{N} lP^{\varepsilon}(l)}{\sum_{l=l_{min}}^{N} P^{\varepsilon}(l)},$$

це середній час, протягом якого дві ділянки траєкторії проходять близько одна до одної, і може розглядатися як середній час передбачуваності. Чим більше значення L, тим менша випадковість, тобто легше визначити поведінку ознаки системи.

Міра ентропії (entropy, ENTR) співвідноситься з ентропією Шеннона (Shannon) частотного розподілу довжин діагональних ліній

$$ENTR = -\sum_{l=l_{min}}^{N} p(l)\ln p(l), \qquad p(l) = \frac{P^{\varepsilon}(l)}{\sum_{l=l_{min}}^{N} P^{\varepsilon}(l)},$$

та відображає складність детерміністської складової у системі. Велике значення ентропії має на увазі періодичність системи, а низьке - хаотичність. Інакше кажучи, велика ентропія слідує за складнішою системою.

Міра завмирання (laminarity, LAM)

$$LAM = \frac{\sum_{v=v_{min}}^{N} vP^{\varepsilon}(v)}{\sum_{i,j=1}^{N} R_{i,j}^{m,\varepsilon}},$$

визначається відношенням кількості рекурентних точок, що утворюють вертикальні лінії, до загальної кількості рекурентних точок. LAM характеризує наявність станів завмирання системи (тобто коли рух системи фазової траєкторії зупиняється або просувається дуже повільно). Ламінарність обчислює ймовірність того, що стан залишиться для наступного тимчасового кроку.

Середня довжина вертикальних структур

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

$$TT = \frac{\Sigma_{v=v_{min}}^{N} vP^{\varepsilon}(v)}{\Sigma_{i,j=1}^{N} P^{\varepsilon}(v)},$$

називається *мірою часу зупинки (trapping time, TT)* і характеризує середній час, який система може провести у певному стані або тривалість часу, протягом якого кожен стан перебуває у пастці.

Проведемо пошаговий розрахунок для нормального та аномального часового ряду. Як було зазначено в методології розрахунки велися для міри отриманої кількості даних у байтах у 1 секунду. Назвемо данні зібранні з даних процессів $Rate1$ та $Rate2$.

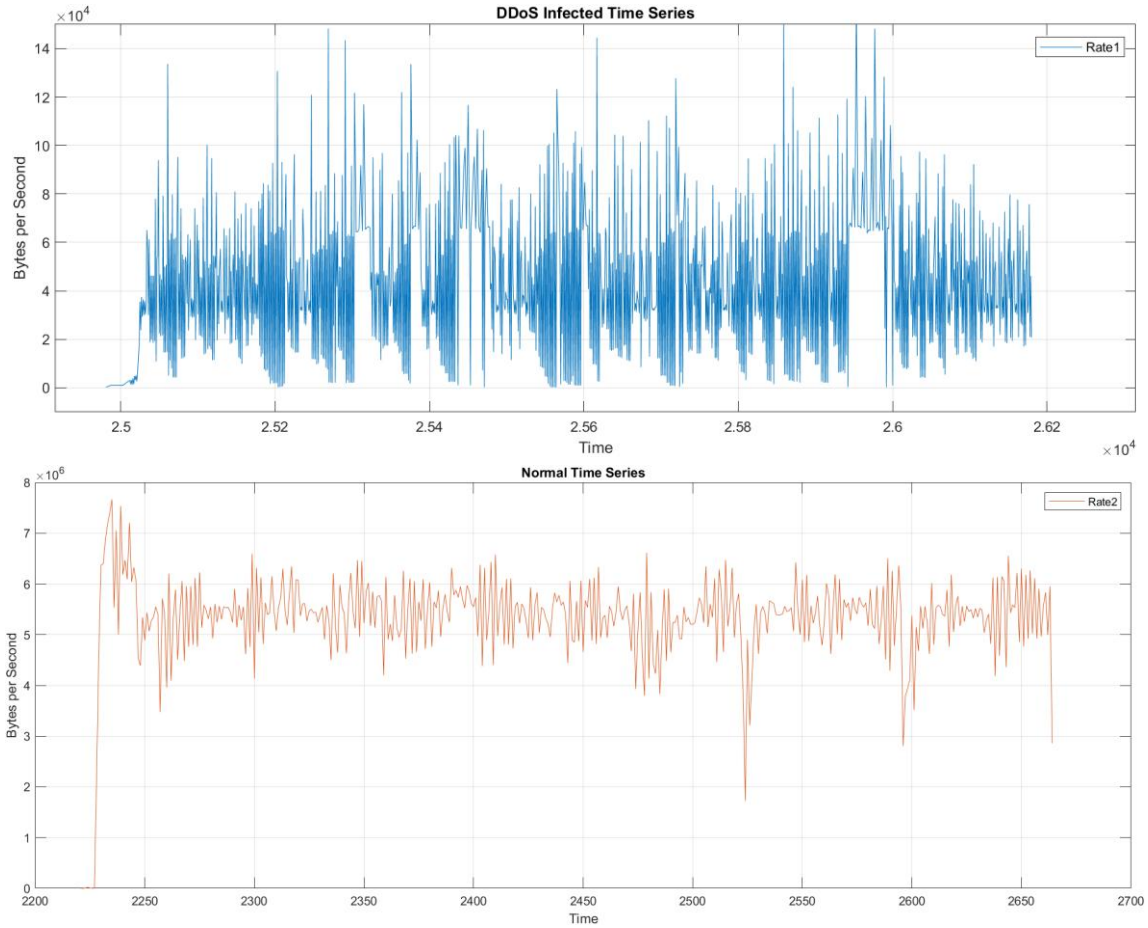

Рисунок 1 - Траєкторія поведінки сигналів у т. $Rate1$(синя), $Rate2$(червона)

Розрахунки для точки $Rate1$.

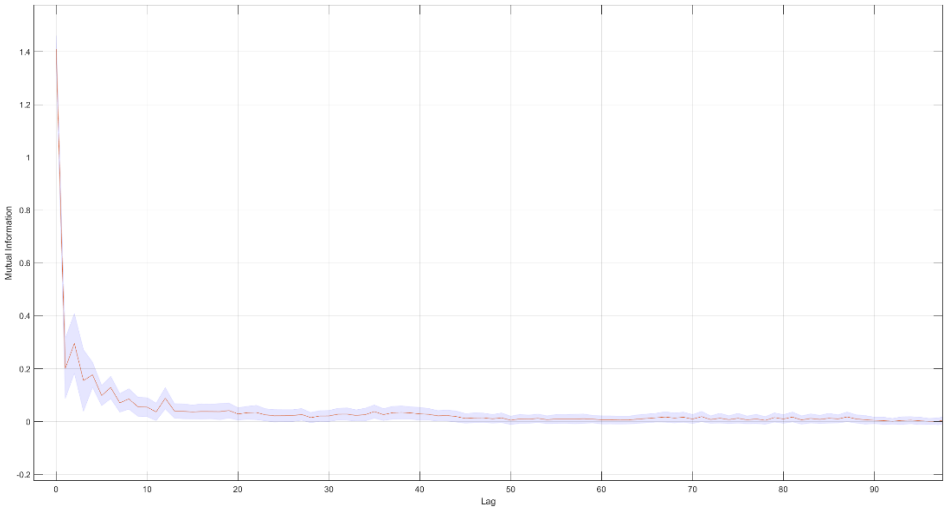

Рисунок 2 - Визначення параметра затримки τ для точки $Rate1$ (значення 28, 50, 78)

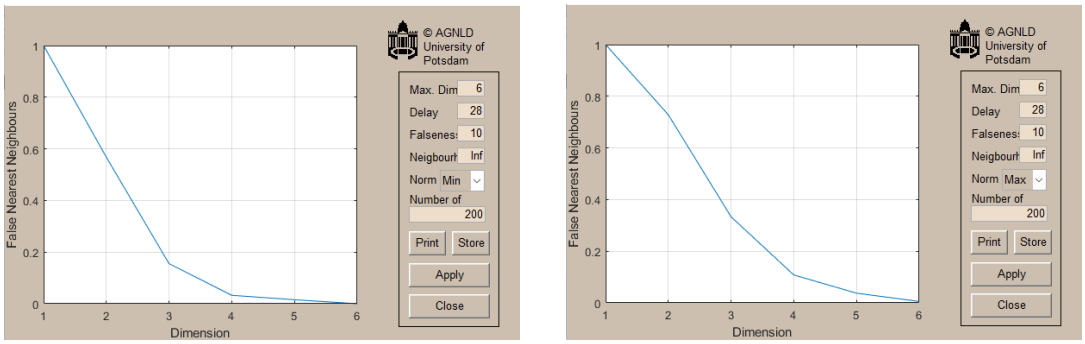

Рисунок 3 - Обрахування m для точки $Rate1$, τ =28(перше значення ціле нижче 0.1) m =4 зліва та m=5 справа (надано для обох норм)

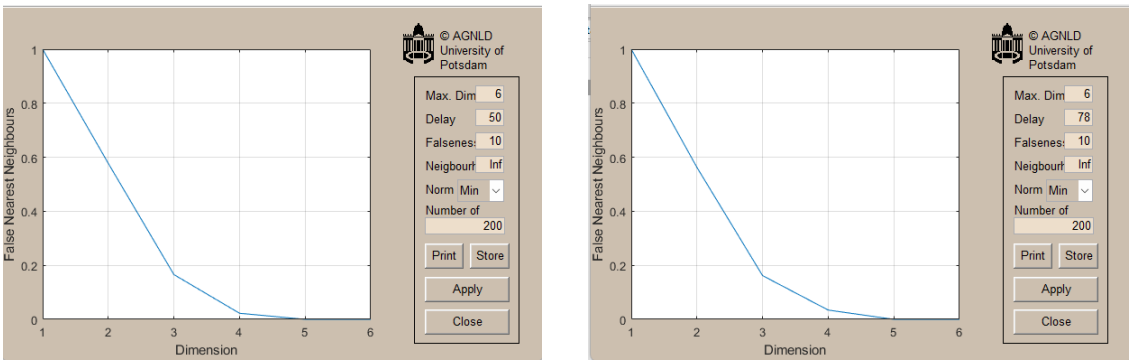

Рисунок 4 - Обрахування m для точки $Rate1$, τ =50 (зліва), 78(справа) m =4 зліва та m=4 справа (надано для норми min)

Результати щодо норми max такі, як на рис.3 для норми min m=5.

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Для розрахунку рекурентної діаграмии було викроистано інший інструмент MatLab:

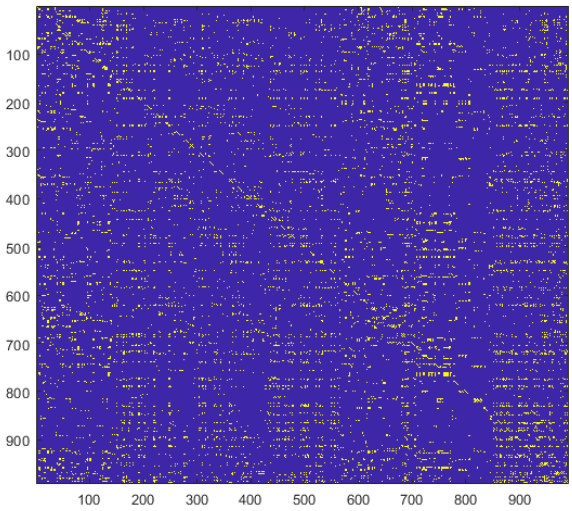

Рисунок 5 - Рекурентна діаграма сигналу O1 (m=5, τ =28, ε=0.2)

Програма видала наступні значення показників RQA:

Таблиця 1

DDoS LoIC	RR	DET	ENTR	L	LAM	TT
$Rate1$	0.046105	0.24954	0.06717	2.2379	0.019056	2.29866

1. Для точки $Rate2$ значення параметра затримки τ аналогічно рис.2. було визначено та мало наступні значення: τ =21, 43, 87.

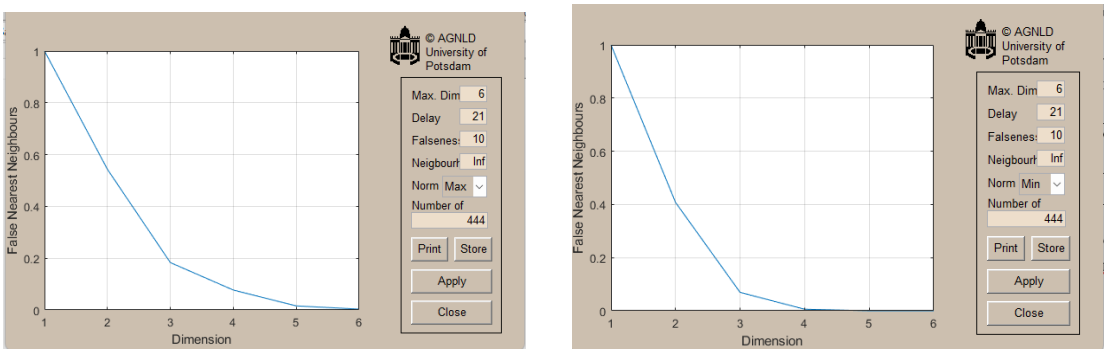

Рисунок 6 - Обрахування m для точки $Rate2$, τ =21(перше значення ціле нижче 0.1) m =5 зліва та m=4 справа (надано для обох норм)

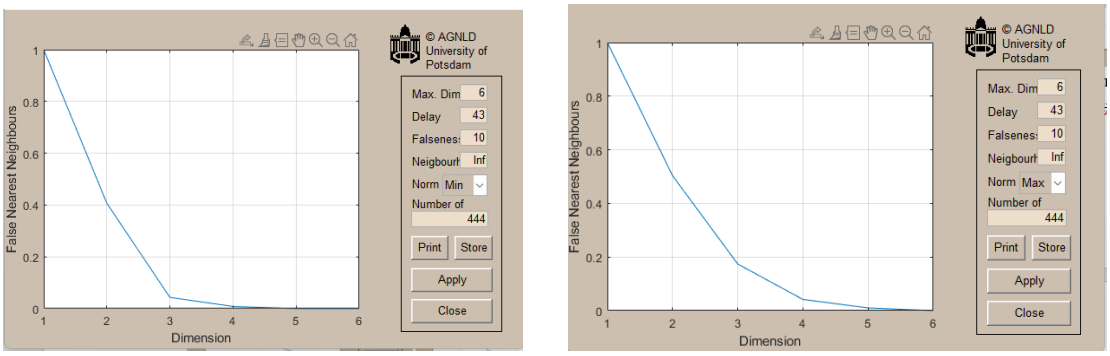

Рисунок 7 - Обрахування m для точки $Rate2$, τ = 43(перше значення ціле нижче 0.1) m =4 зліва та m=5 справа (надано для обох норм)

Результати щодо точки O2 для τ =87, m =5 для max норми та m=4 для min норми. Такім чином вони аналогічні для усіх значень параметру затримки.

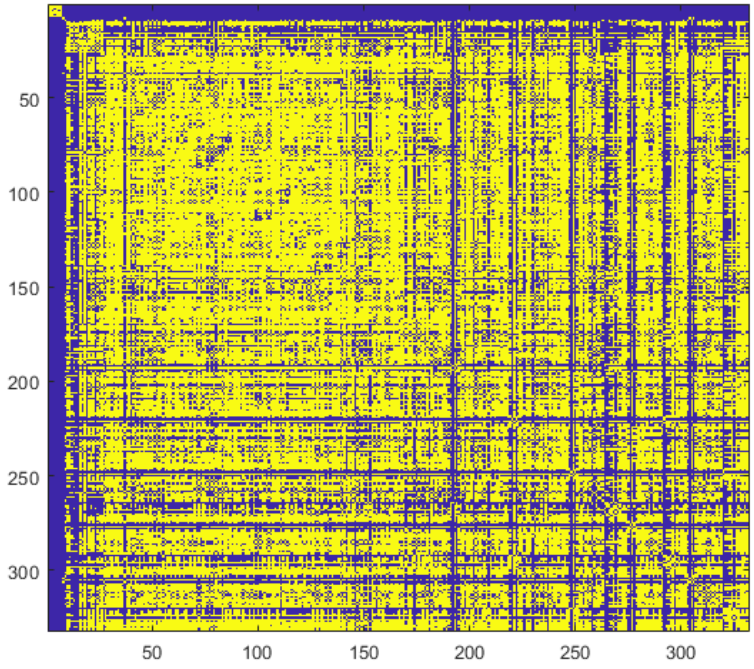

Рисунок 8 - Рекурентна діаграма RP сигналу $Rate2$ (m=5, τ =21, ε=0.2)

Якщо дивитися на рис. 5 та рис. 8, то можна помітити лише деяку незначну структуру схожість. Але зробити якісь висновки досить складно адже процеси відбуваються з різною інтенсивністю.

Таблиця 2

DDoS LoIC	RR	DET	ENTR	L	LAM	TT
$Rate2$	0.62454	0.87469	0.33398	4.52934	0.86912	5.64760

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

Після першого аналізу отриманої інформації можна зробити висновок, що використання RP діаграми для будь-якого візуального аналізу є досить важким завданням. Наявними є також деякі проблеми з визначення значень параметру затримки з представленого графіку (Рис.2).

Використовуючи надану методику, розраховуємо чиселні показники для усіх нормальних та аномальних даних мережевого трафіку. Усі данні будуть зведені у єдину таблицю груповану по маркуванню. Результати обробки мережевих даних, що описані у першому розділі, представлено у підсумковій таблиці 3.

Таблиця 3

Показники RQA у аномальних процесах						
	RR	DET	ENTR	L	LAM	TT
DDoS LoIC	0.04611	0.24954	0.06717	2.23791	0.019056	2.29866
DoS slowloris	0.00134	0,38128	0.05415	2.12443	0.991322	2.10536
DoS Slowhttptest	0.00913	0.56233	0.00124	12.1245	0.017162	6.41234
DoS Hulk	0.04694	0.27281	0.04524	2.40560	0.426120	2.75914
DoS GoldenEye	0.04797	0.29281	0.06428	3.81642	0.049356	4.1282
Показники RQA у нормальних процесах						
	RR	DET	ENTR	L	LAM	TT
Normal1	0.624542	0.87469	0.33398	4.52934	0.86912	5.64760
Normal2	0.00412	0.00831	1.08273	85.12433	0.12543	47.12932
Normal3	0.22784	0.51283	1.69012	25.09128	0.71293	10.09120
Normal4	0	0.95190	4.68395	112.10353	0.99153	215.12093
Normal5	0.00084	0.72612	0.91243	2.02089	0.02132	2.33333

Безпосередній порівняльний аналіз значень, який знайшов відображення у таблиці 3, дозволив виявити деякі нечіткі межи щодо ряду обрахованих чисельних показників. Ці показники дозволяють виконати виділити аномальну поведінку під час атаки DDoS у переважній більшості випадків. У подальшому данні числові характеристики можуть використовуватися для уточнення методології пошуку аномалій у якості додаткової фільтрації.

Під час атаки DoS/DDoS данні не є однорідними. Вдалось виділити один стандартний тип, який має суттєву різницю від нормальної поведінки системи за параметрами: міра рекурентності, міра детермінізму та міра ентропії. Ці характеристики є наступними:

1. показники RR мають усі значення біля нуля [0.044-0.048]
2. показники DET, ENTR значення різні, але з одного невеликого діапазону (наприклад, [0.27-0.3] або [0.056-0.071]);

У результаті обчислювального експерименту, який полягав у проведенні процедури – розрахунків чисельних показників рекурентних діаграм (RP)– авторам вдалося виділити один тип – «DDOS-RP», якій відрізняє аномальний процесс від нормального та визначає саме атаки типу DoS/DDoS.

У наступному пропонується розглянути задачу налаштування та навчання нейроної мережі з використанням числових характеристик RQA. З розвитком Neural-ODE починають розвиватись методології додаткового налаштування нейронних мереж заснованих на методах аналізу нелінійних систем. Також запропроновано розглянути інші види атак зокрема атаки типу botnet та різні види втручання.

Висновки. 1. У роботі наведено методику розрахунків чисельних показників рекурентної діаграми (RP) та визначення нейрофізіологічного значення її параметрів — міра рекурентності (RR ймовірність повторення стану), міра детермінізму (DET більше значення детермінізму вказує більш сильну передбачуваність системи), середня довжина діагональних ліній(чим більше значення L, тим менша випадковість, тобто легше визначити поведінку ознаки системи), міра ентропії (ENTR відображає складність детерміністської складової у системі. Велике значення ентропії має на увазі періодичність системи, а низьке - хаотичність. Інакше кажучи, велика ентропія слідує за складнішою системою), міра завмирання (LAM ламінарність обчислює ймовірність того, що стан залишиться для наступного тимчасового кроку), міра часу зупинки (TT характеризує середній час, який система може провести у певному стані або

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

тривалість часу, протягом якого кожен стан перебуває у пастці) для аномальних данних мережевого трафіку і нормальних.

2. Аналіз отриманої інформації дозволив визначити деякі види аномальної активності типу DoS/DDoS, які мають наступні ознаки:

3. показники RR мають усі значення біля нуля [0.044-0.048]

4. показники DET, ENTR значення різні, але з одного невеликого діапазону (наприклад, [0.27-0.3] або [0.056-0.071]);

3. В результаті обчислювального експерименту, вдалося виділити один тип рекурентних діаграм – «DDOS-RP», якій відрізняє аномальний процес від нормального та визначає саме атаки типу DoS/DDoS.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Palmeieri, F. & Fiore, U. Network anomaly detection through non-linear analysis. Computers Security, 2010, 29(7), 737-55.

2. Somenath Mukherjeea, Rajdeep Ray, Rajkumar Samantac, Mofazzal H. Khondekar ,Goutam Sanyal Nonlinearity and chaos in wireless network traffic Chaos, Solitons and Fractals 96 (2017) 23–29

3. N. Jeyanthi, R. Thandeeswaran, J. Vinithra RQA Based Approach to Detect and Prevent DDoS Attacks in VoIP Networks CYBERNETICS AND INFORMATION TECHNOLOGIE Volume 14, No 1 11-24 DOI: 10.2478/cait-2014-0002

4. Yun Chen Shijie Sum, Hui Yang Convolutional Neural Network Analysisof Recurrence Plots for Anomaly Detection International Journal of Bifurcation and Chaos, Vol. 30, No. 1 (2020) 2050002 (13 pages)

5. Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, “Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization”, 4th International Conference on Information Systems Security and Privacy (ICISSP), Portugal, January 2018

6. Chandola, V.; Banerjee, A. & Kumar, V. Anomaly detection: A survey. ACM Computing Surveys, 2009,41(3), 1-58.

7. Hodge, V. & Austin, J. A survey of outlier detection methodologies. Art. Intel. Revi., 2004, 22(2), 85-126

8. Mahoney, M. Network traffic anomaly detection based on packet bytes. In Proceedings of ACM Symposium on Applied Computing, 2003, pp. 346-50.

9. Mekler A.A. Application of the Apparatus for Nonlinear Analysis of Dynamic Systems for EEG Signal Processing // Actual Problems of Modern Mathematics: Scientific Notes. p / ed. prof. Kalashnikova E.V., ed. LGU them. A.S. Pushkin, St. Petersburg, 2004. T. 13 (issue 2), p. 112-140.

REFERENCES

1. Palmeieri, F. & Fiore, U. Network anomaly detection through non-linear analysis. Computers Security, 2010, 29(7), 737-55.

2. Somenath Mukherjeea, Rajdeep Ray, Rajkumar Samantac, Mofazzal H. Khondekar ,Goutam Sanyal Nonlinearity and chaos in wireless network traffic Chaos, Solitons and Fractals 96 (2017) 23–29

3. N. Jeyanthi, R. Thandeeswaran, J. Vinithra RQA Based Approach to Detect and Prevent DDoS Attacks in VoIP Networks CYBERNETICS AND INFORMATION TECHNOLOGIE Volume 14, No 1 11-24 DOI: 10.2478/cait-2014-0002

4. Yun Chen Shijie Sum, Hui Yang Convolutional Neural Network Analysisof Recurrence Plots for Anomaly Detection International Journal of Bifurcation and Chaos, Vol. 30, No. 1 (2020) 2050002 (13 pages)

5. Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization", 4th International Conference on Information Systems Security and Privacy (ICISSP), Portugal, January 2018

6. Chandola, V.; Banerjee, A. & Kumar, V. Anomaly detection: A survey. ACM Computing Surveys, 2009,41(3), 1-58.

7. Hodge, V. & Austin, J. A survey of outlier detection methodologies. Art. Intel. Revi., 2004, 22(2), 85-126

8. Mahoney, M. Network traffic anomaly detection based on packet bytes. In Proceedings of ACM Symposium on Applied Computing, 2003, pp. 346-50.

9. Mekler A.A. Application of the Apparatus for Nonlinear Analysis of Dynamic Systems for EEG Signal Processing // Actual Problems of Modern Mathematics: Scientific Notes. p / ed. prof. Kalashnikova E.V., ed. LGU them. A.S. Pushkin, St. Petersburg, 2004. T. 13 (issue 2), p. 112-140.

Received 17.10.2023.
Accepted 21.10.2023.

Using the method of nonlinear recursive analysis for detecting DDOS anomalies in time series data

This research endeavors to address this gap by determining a qualitative characteristic for server network traffic and use it to construct the corresponding recurrence plot (RP). The goal of this study is to develop and assess a novel technique based on nonlinear recursive analysis to detect Distributed Denial of Service (DDoS) anomalies in network traffic time series data. With the increasing frequency of DDoS attacks on modern digital

ISSN 1562-9945 (Print)
ISSN 2707-7977 (Online)

infrastructures, there is a pressing need for more efficient and accurate detection methods.

There has been some attempts to apply nonlinear analysis to network traffic [2-4], but those studies lack critical steps in determining parameters for embedding space dimension m *and delay time* τ *. More recent studies have explored machine learning and deep learning approaches [7], which offer improved accuracy but can be computationally intensive and require extensive training data. Despite the advancements, there remains a need for a method that is both accurate and efficient, especially in real-time detection scenarios.*

The researchers employed nonlinear recursive analysis by estimating RQA parameters and determining a qualitative characteristic of data points of DDoS attack contained in CIC-IDS2017 dataset. A technique for determining hidden information for this series and its use for constructing the corresponding recurrence diagram (RP) at the points of information retrieval are described. It is shown that the use of RP has significant drawbacks associated with the visualization of information on a computer monitor screen, so another way of research is proposed - the calculation of numerical indicators of RP

The given calculated RP indicators made it possible to typify the received data and determine the type, which was named "DDOS-RP", which makes it possible to distinguish some types of DoS/DDoS type attacks. The study concludes by recommending further exploration of this method in diverse network environments and against more complex DDoS attack patterns.

Key words: recurrent analysis, network traffic, time series, recurrent diagram, delay parameter, dimension of the nesting space, RQA analysis, Matlab environment

Гулий Тарас Олександрович – аспірант, кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара.

Білозьоров Василій Євгенійович – доктор фізико-математичних наук, професор кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара.

Hulyi Taras - postgraduate student, department of computer technologies, Oles Honchar Dnipro National University.

Belozyorov Vasily - Doctor of Physical and Mathematical Sciences, Professor, Department of computer Technologies, Oles Honchar Dnipro National University.