УДК 004.83

С.О.Сушко, В.П.Мойсєєнко

ДОСЛІДЖЕННЯ ЕНТРОПІЇ МЕРЕЖЕВОГО ТРАФІКУ МЕТОДОМ ЧАСОВОГО ВІКНА

Розглядається зв'язок між ентропією мережевого трафіка і рівнем мережевої активності , що викликає відмови в обслуговуванні. Визначення ентропії на основі методу рухомого вікна дають змогу проводити вимірювання у режимі, наближеному до режиму реального часу, и своєчасно сигналізувати про порушення нормальної роботи мережі, зокрема про DDoS-атаки.

Рассматривается связь между энтропией сетевого трафика и уровнем сетевой активности, вызывающей отказ в обслуживании. Определение энтропии на основе метода движущегося окна позволяет проводить измерения в режиме, близком к режиму реального времени и своевременно сигнализировать о нарушении нормальной работы сети, в частности о DDoS-атаках.

Determining the correspondence between the entropy of network traffic and the level of network activity, such as denial of service, are considered. Definition of entropy based on the method of moving the window allows measurements in near to real time and in a timely manner to signal properly, the network, in particular the DDoS-attacks.

Останнім часом спостерігається зростання кількості розподілених атак на глобальні комп’ютерні мережі. Значна їх частина спрямована на порушення доступності або «розподілену відмову в обслуговуванні» (Distributed Denial of Service, DDoS) і може мати за ціль будь-який ресурс, пов'язаний с мережею Інтернет, включаючи сервери DNS, сервіси електронної пошти, он-лайнові проекти або інтерфейси маршрутизації. Атака зводиться до перевантаження хосту або мережевого ресурсу шляхом наповнення системи великою кількістю мережевих пакетів. Реалізація атак здійснюється програмними агентами, розміщеними на хостах, які зловмисник скомпрометував раніше. Об’єм трафіку розподіленої атаки може досягати 10 Гбіт/с, а її наслідком стає не тільки вихід із ладу окремих хостів або служб, а й зупинка роботи кореневих DNS-серверів і часткове або повне припинення роботи сервісів. На жаль, розподілені атаки, спрямовані на відмову в обслуговуванні, складно попередити, особливо, якщо система передбачає загальнодоступні вхідні з’єднання. Ускладнює боротьбу й те, що атака не потребує високої кваліфікації нападника і може здійснюватися з значної кількості адрес, що часто розташовані у різних сегментах мережі і належать різним операторам зв'язку.

Щодо раннього виявлення DDOS-атак, то у літературі склалося чітке переконання, що ефективним індикатором аномального поводження мережі є інформаційний аналіз мережевого трафіку завдяки його інформативності та потенціальної можливості реагування у реальному часі. Для виявлення мережевих аномалій застосовуються відомі математичні методи обробки, аналізу та моделювання сигналів: класичний статистичний, факторний і кластерний аналіз, спектральний аналіз, вейвлет-аналіз, цифрову обробку сигналів, моделювання часових рядів і інтелектуальний аналіз даних. Практичні

розробки в області виявлення порушень мережевих аномалій ведуть як університетські наукові центри (Ohio University, Columbia University, МГУ та ін.), так і великі комерційні компанії (Cisco, CA, ISS, Symantec).

У даній роботі виявлення шкідливої мережевої активності базується на обчисленні ентропії мережевого трафіку. Ще у 1994 р. В. Леланд виявив, що в умовах DDOS-атак агрегований мережевий трафік стає самоподібним [1]. У 2003 р. Файнштейн і Шнахенберг [2] показали, що як індикатор атак можна розглядати такі статистичні характеристики мережевого потоку, як вибіркове середнє, вибіркову дисперсію та критерій згоди Пірсона χ^2 або інформаційно-теоретичну міру – ентропію трафіку. Останнє обумовлене тим, що у разі нормальної роботи комп'ютерної мережі спектр IP-атрибутів мережевих пакетів достатньо широкий, а потік є хаотичним и невпорядкованим. Під час же розгортання DDOS-атак «запас безладу» у мережі поступово втрачається, бо її «наводнює» значна кількість схожих пакетів. Кількісно такі процеси традиційно характеризують за допомогою ентропії розподілу ймовірностей К.Шеннона. Отже, DDOS-атаки неминуче мають призводити до зменшення ентропії мережевого трафіку [3-5].

На сьогодні існує два основних транспортних протоколи передачі даних у IP мережі: TCP і UDP. Загальні поля пакетів, генерованих обома протоколами, – IP-адреса відправника пакету, IP-адреса його отримувача, вихідний порт, вхідний порт призначення. При цьому кожен протокол характеризується своїм набором керуючих прапорців, що відповідають за встановлення з'єднання, синхронізацію, початок передачі даних і т.д. Оскільки під час DDoS-атаки пакети надсилаються з багатьох джерел в основному на один вхідний порт, то у подальшому для нас важливими будуть два параметри, а саме вихідна IP-адреса джерела пакета і вхідний порт призначення (ці характеристики вже й раніше зарекомендували себе як добрі кандидати для виявлення хробаків, сканування та ін. [3]). Принцип такої структуризації мережевих пакетів графічно подано на рис. 1.

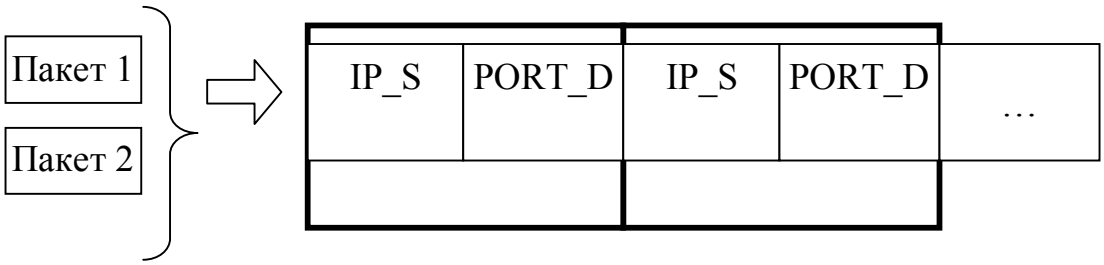

Рис. 1. Структуризація мережевих пакетів

Для досліджень атак типу «відмова в обслуговуванні» (DDOS) була створена розподілена мережа, в якій проводились навантажувальні тести з використанням програм netmap, nping та WPE Pro. Локальна мережа працювала у звичайному режимі та зазнавала впливу DDOS-атаками, різного рівня інтенсивності, максимально наближеними до реальних DDOS вторгнень.

Статистична інформація збиралася за допомогою спеціально розробленої програми-аналізатора на основі відкритих бібліотек WinPcap та jpcap. За допомогою таймера високої роздільної здатності аналізатора здійснювалося перехоплення трафіка, що надходив на мережевий інтерфейс сервера, де було розміщено сервіс, що атакується. Базова інформація, зібрана за період проведення дослідження, – частота пакетів кожного типу, виміряна як кількість пакетів за секунду.

Параметри моделювання представлені у табл. 1.

Табл. 1

Параметри моделювання

Номер послідовності	**Розмір послідовності (пакетів)**	**Час, за який було зібрано пакети (хв.)**	
		Звичайний режим роботи мережі	**Моделювання атаки**
1	**50 000**	**15**	**2**
2	**100 000**	**30**	**3**
3	**150 000**	**60**	**5**
4	**200 000**	**75**	**6**
5	**250 000**	**84**	**9**

Тестовий трафік складався з суміші реального мережевого трафіку та спеціально змодельованого, за параметрами наближеного до трафіку при атаках типу DDOS. Таким чином вдалося змоделювати ситуацію на сервері при атаці, що є близькою до реальної.

За Шенноном ентропія трафіка залежіть від ймовірностей p_i появи пакетів при їх передачі:

$$H(x) = -\sum_{i=1}^{n} p_i \cdot \log_2 p_i ,$$

де в якості ймовірності появи p_i пакету i-го типу може виступати його частота $f_i = \frac{n_i}{N}$, n_i – кількість пакетів i-го типу типу, N – загальна кількість пакетів трафіку.

Розраховуючи ентропію за таким класичним алгоритмом, потрібно перераховувати частоти усіх пакетів у разі надходження нового пакета. При великій кількості пакетів це суттєво знижує швидкість обчислень. Зрозуміло, що для своєчасного блокування DDOS-атаки при підозрі на виникнення нештатних ситуацій, ентропія має розраховуватися не просто швидко, а надзвичайно швидко. Отже, обчислення ентропії розумно прискорити за

допомогою, наприклад, методу рухомого вікна, коли трафік сканується рухомим вікном, накладеним на певну кількість пакетів (рис. 2). Розрахунки ентропії здійснюються за наступним алгоритмом. Вибирається вікно розміром W пакетів, обчислюються і зберігаються частоти f_i кожного типу пакетів та визначається базове значення ентропії H_0 для перших W пакетів трафіку. Далі вікно зсувається на величину DW вправо вздовж послідовності пакетів трафіку, фіксуються частоти $f_{before,input}$ і $f_{now,input}$ пакетів, що зайшли у вікно, і частоти $f_{before,output}$ і $f_{now,output}$ пакетів, що залишились поза межами вікна після його зсуву. За цих умов поточне значення ентропії дорівнює

$$H_i = H_{i-1} + \Delta H ,$$

де приріст ΔH відображає зміну ентропії при зсуві вікна.

$$\Delta H = -\sum_{i=1}^{n} p_i \cdot \log_2 p_i + \sum_{j=1}^{l} p_j \cdot \log_2 p_j ,$$

що згідно з наведеним алгоритмом дає наступну розрахункову формулу

$$\Delta H = -f_{before,input} \cdot \log f_{before,input} - f_{before,output} \cdot \log f_{before,output}$$
$$+ f_{now,input} \cdot \log f_{now,input} + f_{now,output} \cdot \log f_{now,output} .$$

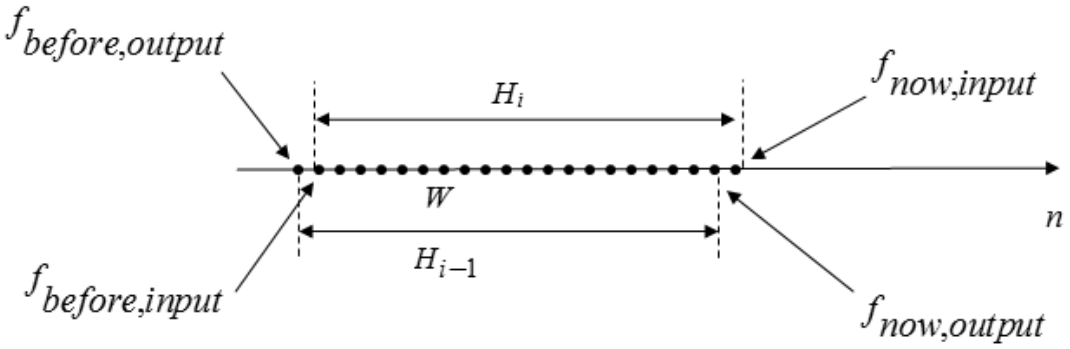

Рис. 2 Схема методу рухомого вікна

Такий підхід дозволяє визначати ентропію без повного перерахунку частот усіх пакетів з надходженням нового пакету і має прискорити швидкодію алгоритму. Для порівняння швидкості роботи алгоритмів вимірювався час обчислення ентропії на вищезазначених послідовностях (табл.2). Тестування проводилось на процесорі Intel Core i5. Аналіз отриманих даних свідчить про значно більшу швидкість обчислень за допомогою методу рухомого вікна. Це дозволяє віддати перевагу на користь методу рухомого вікна. У той же час недолік методу порівняно з класичними розрахунками – збільшення потреб в оперативній пам’яті на 20-25%. Об’єм виділеної оперативної пам’яті – у середньому 65МБ.

Табл. 2

Час обчислення ентропії

Номер послідовності	Час обчислення за «класичним» методом (мс)	Час обчислення за методом рухомого вікна (мс)
1	**41**	**39**
2	**138**	**60**
3	**348**	**101**
4	**670**	**120**
5	**1070**	**150**

Експерименти виявляють суттєву залежність ентропії від розміру вікна W (рис. 3). Цей параметр алгоритму індивідуальний для кожної мережі і має визначатися експериментально. Тому первинною задачею стає встановлення оптимального значення розміру вікна W, за якого значення ентропії, обчислені різними алгоритмами, будуть збігатися. У нашій роботі ця проблема проаналізована за допомогою побудови графіків залежності ентропії досліджуваного трафіку (рис.4) від розмірів вікна W для п'яти різних послідовностей пакетів. Зокрема, ентропія трафіку найбільш близька до значень, отриманих за класичним алгоритмом (табл. 3), при розмірі вікна $W = 6$ або $W = 7$ пакетів. Отже, у подальшому вважатимемо, що розмір вікна $W = 6$ пакетів є оптимальним.

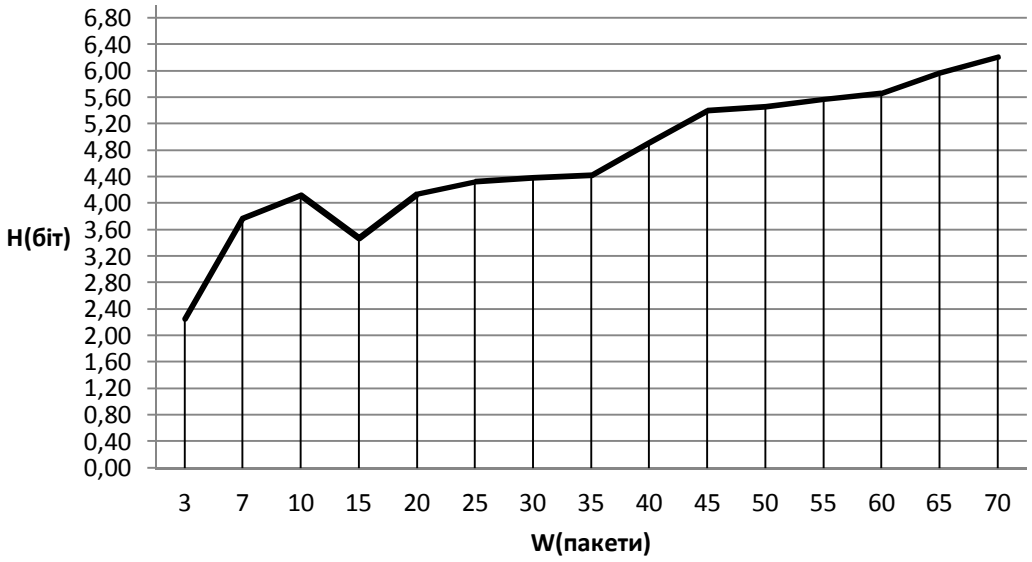

Рис. 3 Залежність ентропії H трафіка від розміру вікна

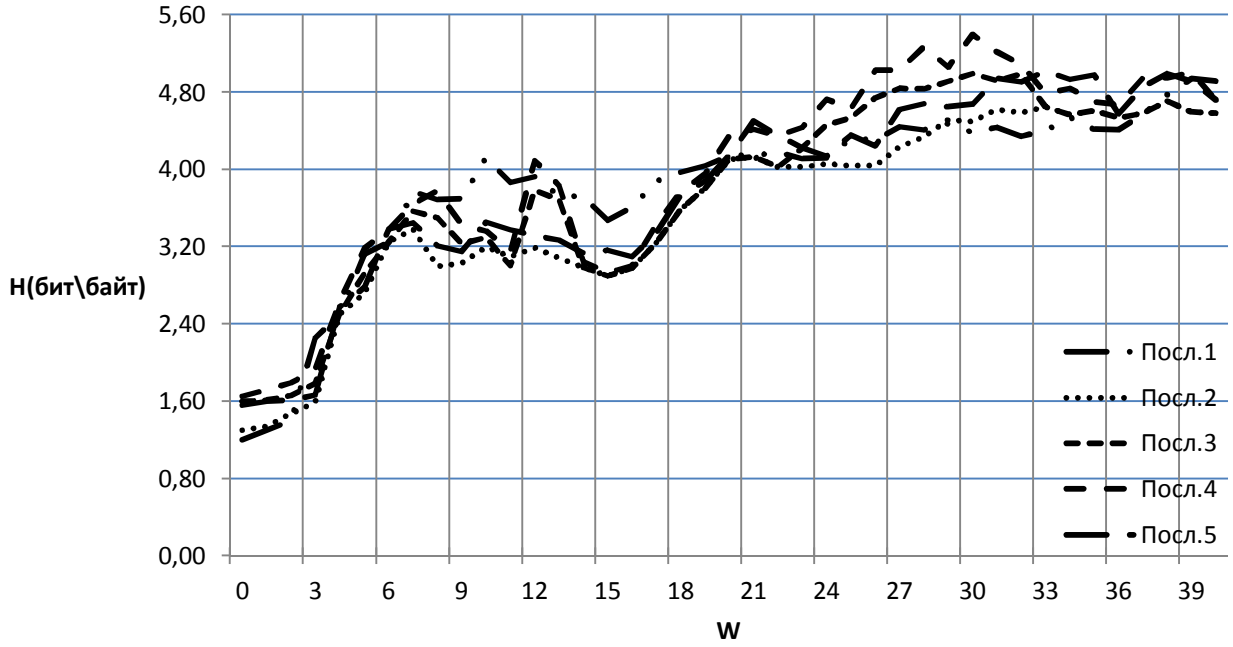

Рис. 4. Залежність ентропії H трафіка від розміру W вікна для різних послідовностей мережевих пакетів.

Для ілюстрації можливості використання ентропії трафіку як детектора DDOS-атак ентропія обчислювалась для мережевого трафіка без атак і при емуляції атаки типу DDOS . Для цього було задіяно десять машин, одна з яких використовувалась як сервер, що обробляв запити клієнтів. З решти дев'яти машин на сервер відправлялися пакети доти, поки не досягалась необхідна кількість пакетів зі максимально можливою швидкістю для мережі Fast Ethernet. Обчислені значення ентропії трафіку без атак та при DDOS-атаці можна порівняти у табл.3.

Табл.3

Ентропія *H* мережевого трафіку без атаки і при DDoS-атаці

Послідовність	**Ентропія (у бітах), обчислена за класичним алгоритмом**	**Ентропія (у бітах) без атаки**	**Ентропія (у бітах) при DDoS-атаці**
1	**3,19**	**3,37**	**1,23**
2	**3,23**	**3,28**	**1,5**
3	**3,27**	**3,23**	**1,92**
4	**3,25**	**3,29**	**1,7**
5	**3,34**	**3,37**	**1,85**

Аналіз отриманих даних підтверджує чутливість ентропії трафіку до DDoS-атак: по перше, з розвитком атаки її значення зменшуються майже на

68%, а по-друге, падіння ентропії спостерігається і при аномаліях, що не проявляють себе як значні відхилення у загальному обсязі трафіку.

Слід відзначити, що отримані результати не впливає використання буферу мережевої плати та розмір пакету. У запропонованій постановці задачі обчислювальна ентропія змінюється при надходженні пакету з визначеними параметрами. Фактор часу не враховувався.

Таким чином, обчислення ентропії трафіку доцільно застосовувати у системах виявлення вторгнень як базовий показник мережевої активності. Безпосередньо обчисленням передують декілька підготовчих етапів: вибір параметрів мережі, визначення розміру вікна для початку роботи алгоритму. Визначення ентропії на основі методу рухомого вікна дають змогу проводити вимірювання у режимі, наближеному до режиму реального часу, и своєчасно сигналізувати про порушення нормальної роботи мережі, зокрема DDoS-атаки.

Список літератури

1. Leland W.E., Taqqu M.S., Willenger W., Wilson D.V. On the Self-Similar Nature of Ethernet Traffic. IEEE Transactions on Networking, 1994, v. 2, Feb. p.1 – 15.
2. Feinstein L., Schnackenberg D. Statistical Approaches to DDoS Attack Detection and Response. Proceedings of the DARPA Information Survivability Conference and Expostion (DISCEX'03), April 2003.
3. Seong Soo Kim. Real-time Analysis of Aggregate Network Traffic for Anomaly Detection. http://cesg.tamu.edu/wp-content/uploads/2012/02/TAMU-ECE-2005-02.pdf
4. Lee W., Xiang D. Information-Theoretic Measures for Anomaly Detection. Conference: IEEE Symposium on Security and Privacy. – 2001, 130 – 143 p.