

## СТАТИСТИЧНИЙ АНАЛІЗ РЕАЛІЗАЦІЇ ТРАФІКУ НА САМОПОДІБНІСТЬ

*В статті розглянуто аналіз і дослідження структури трафіку, отриманого в різних IP-мережах, на предмет виявлення характерних особливостей самоподібності, які необхідно враховувати при забезпеченні якості обслуговування в мережах. Існують різні методи виявлення властивостей стохастичної самоподібності, визначення коефіцієнта Херста. В даній статті описаний один із методів дослідження рядів трафіку і тестових рядів з метою підвищення достовірності отриманих результатів адаптивної маршрутизації.*

**Ключові слова:** еквідистантна, агрегування, самоподібність, інтернет-трафік.

I.V.MULYAR, V.V.ZAVADOVSKY

Khmelnytsky national university

## STATISTICAL ANALYSIS OF TRAFFIC ON SELF-SIMILARITY

*In the article the analysis and study of traffic patterns obtained in different IP networks, to identify the characteristics of self-similarity that must be considered while ensuring quality of service in networks. There are different methods to identify the stochastic properties of self-similarity, Hurst coefficient of determination. This paper describes a series of methods for the study of traffic and test series in order to increase reliability of the results of adaptive routing.*

**Keywords:** equidistantly, aggregation, self-similarity, webtraffic.

## Вступ

Численні дослідження трафіку комп'ютерних мереж свідчать про те, що він має властивість масштабною інваріантності - має особливу фрактальну (самоподібну) структуру, яка зберігається в різних масштабах. В процесі передачі виникають великі сплески при відносно низькому середньому рівні трафіку. Дане явище, яке значно впливає на втрати, затримки і джиттер пакетів, не враховується в класичній теорії телетрафіку. Розрахунок параметрів системи розподілу інформації, призначеної для обробки мережевого трафіку, за класичними формулами дає некоректні, невиправдано оптимістичні результати. Алгоритми обробки трафіку, створені для роботи з найпростішим потоком неефективні для фрактальних потоків. Ситуація, що склалася в сучасних комп'ютерних мережах, наявність великої кількості мережевих маршрутів на яких періодично виникають різкі коливання затримки в передачі даних і великі втрати пакетів, поява нових властивостей мережевого трафіку, необхідність забезпечення високої якості обслуговування додатків, робить актуальним статистичний аналіз мережевого трафіку.

Реалізація трафіку є тимчасовий ряд: рівномірні відрізки часу і відповідна кожному відрізку інтенсивність потоку. При знятті реалізації, в залежності від ситуації, можна відразу отримати такий ряд, або ряд моментів часу проходження кадру і розмір кадру в байтах. У другому випадку необхідно привести вихідну реалізацію до еквідистантної (з постійним кроком  $\Delta t$ ) шкалою по осі часу, тобто провести процедуру агрегування. При цьому величина  $\Delta t$  називатиметься кроком агрегування. Вісь часу розбивається на інтервали часу  $\Delta t$ , кожному з яких ставиться відповідно відношення кількості, що пройшла за даний інтервал часу, інформації (в байтах) до тривалості інтервалу (в секундах).

## Метою досліджень є:

Аналіз і дослідження структури трафіку, отриманого в IP-мережі, на предмет виявлення характерних особливостей самоподібності, які необхідно враховувати при забезпеченні якості обслуговування в мережах.

Для дослідження трафіку була обрана реалізація, отримана з маршрутизатора навчального закладу м. Хмельницького (Хмельницький університет економіки і права (ХУЕП)), тобто маршрутизатора, через який проходять об'єднанні потоки пакетів від безлічі абонентів.

Реалізація трафіку є тимчасовий ряд: рівномірні відрізки часу і відповідна кожному відрізку інтенсивність потоку. При знятті реалізації, залежно від ситуації, можна відразу отримати такий ряд, або отримують ряд моментів часу проходження кадру і розмір кадру в байтах. У другому випадку необхідно привести вихідну реалізацію до еквідистантної (з постійним кроком  $\Delta t$ ). Вісь часу розбивається на інтервали часу  $\Delta t$ , кожному з яких ставиться відповідно відношення кількості, що пройшла за даний інтервал часу, інформації (у байтах) до тривалості інтервалу (в секундах).

Дана реалізація Інтернет-трафіку отримана з маршрутизатора Хмельницького університету економіки і права (маршрутизатор між внутрішньою мережею ХУЕП та мережею Інтернет). Дані при вхідному трафіку записувалися 14.09.2014 протягом доби, з інтервалом в 5 хвилин, фіксувалися інтервали часу, кількість Ethernet кадрів і кількість інформації в байтах за ці інтервали часу (рисунок 1, рисунок 2). Всього 288 відліків. (Середня інтенсивність трафіку 4200,52 кбайт / с; максимальна інтенсивність 7620,79 кбайт / с; мінімальна інтенсивність 30,69 кбайт / с).

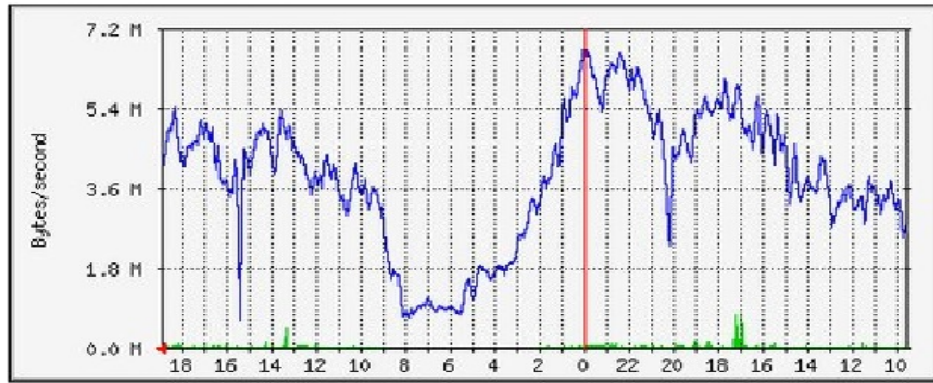


Рис. 1. Інтенсивність трафіку ХУЕП

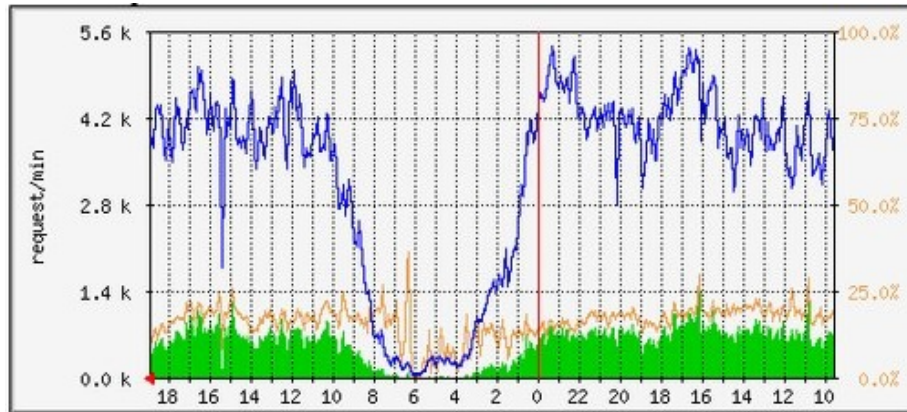


Рис. 2. Кількість пакетів що проходять

Для аналізу трафіку на самоподібність використовуємо ряд, що володіє властивостями самоподібності і випадкову послідовність. У першому випадку можна сформувати послідовність ФБД (фрактальний броунівський рух) із заданим коефіцієнтом Херста (рисунок 1.3) в пакеті Matlab. У якості випадкового процесу виберемо реалізацію класичного білого гауссівського шуму і дисперсією (рисунок 1.4). Отримані послідовності агрегуються з кроком 3 (рисунок 1.5)

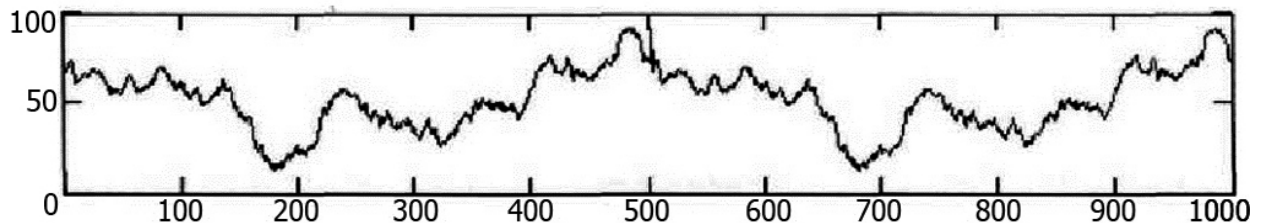


Рис. 3. Фрактальний броунівський рух  $H=0.7$



Рис. 4. Білий гауссівський шум

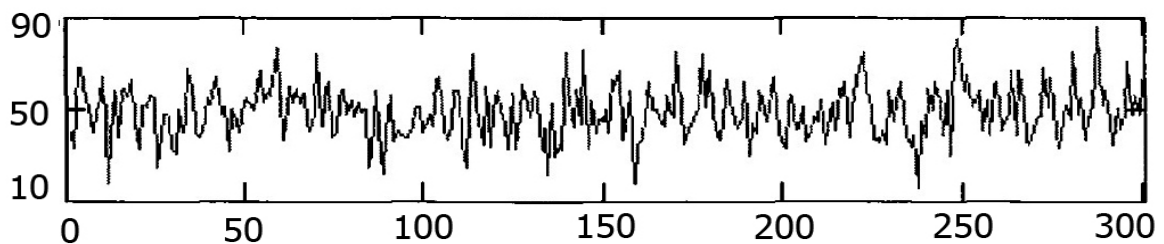


Рис. 5. Білий гауссівський шум, агрегований по 3-ом значенням.

**Розрахунок показника Херста досліджуваних реалізацій**

Параметр Херста є індикатором ступеня самоподібності, а також свідчить про наявність у процесу таких властивостей, як персистентність, антиперсистентність і тривала пам'ять. У літературі описані різні методи визначення цього параметра, причому в різних джерелах одні й ті ж методи описуються по-різному, з різними обмеженнями до досліджуваних рядів. Одним із популярних методів для розрахунку показника Херста є графік  $R/S$  – статистики.

**Графік  $R/S$  – статистики**

Грунтуючись на дослідженні різних явищ, Херст розробив нормовану безрозмірну міру, здатну описати змінність. Цю міру він назвав нормованим розмахом ( $R/S$ ). Для заданого набору спостережень  $X_j (j = 1, \dots, n)$  з вибірковою середнім:

$$\bar{X} = \frac{1}{n \sum_{j=1}^n X_j} \quad (1)$$

і вибірковою дисперсією:

$$S(n) = \frac{1}{n \sum_{j=1}^n (X_j - \bar{X})^2} \quad (2)$$

вводиться поняття розмаху:

$$R(n) = \max D_j - \min D_j \quad (3)$$

де

$$D_j = \sum_{i=1}^j X_i - j\bar{X} \quad (4)$$

Тобто різниця між мінімальним і максимальним відхиленням.

Мінливість випадкового процесу на інтервалі  $N$  визначається нормованою безрозмірною характеристикою, яку Херст назвав нормованим розмахом і показав, що для багатьох природних процесів справедливе емпіричне відношення:

$$\left( \frac{R(n)}{S(n)} \right) \approx cn^H \quad (5)$$

де  $c$  – позитивна константа, не залежна від  $N$   
Прологарифмуючи обидві частини, отримаємо:

$$\ln \left( \frac{R(n)}{S(n)} \right) \approx H \ln(n) + \ln(c) \quad (6)$$

при  $n \rightarrow \infty$

Таким чином, параметр Херста -  $H$  можна оцінити за графіком залежності (рисунок 1.6) і, використовуючи отримані точки, підібрати пряму лінію з нахилом  $H$ .

$R/S$  – метод не надто точний, оскільки дає оцінку тільки рівня самоподібності в тимчасовому ряді. Тому даний метод може використовуватися для перевірки, чи є часовий ряд самоподібним, і якщо є, отримує позитивну оцінку. У випадку

$$\left( \frac{R(n)}{S(n)} \right) \approx cn^{0,5} \quad (7)$$

може розглядатися як характеристика процесу, який не володіє властивістю самоподібності та є короткочасно залежним. Для розрахунку і побудови графіків був використаний пакет Mathcad і Excel, для дуже довгих послідовностей використовувалася програма Fractan 4. На графіку наведено залежності і відповідні ліній тренда, з рівняння лінії тренду визначається коефіцієнт Херста (коефіцієнт нахилу) і приводиться похибка  $R^2$  (коефіцієнт детермінації).

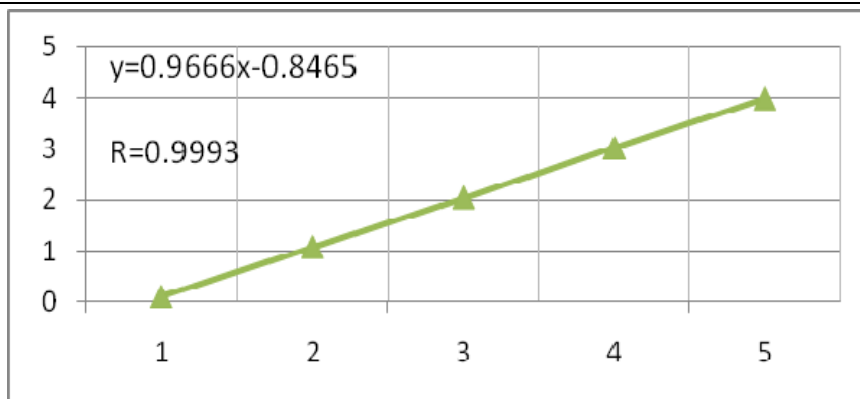


Рис. 6. Оцінка параметрів Херста – H

### Висновки

У результаті аналізу статистичних властивостей трафіку показано, що досліджувані ряди: інтенсивність трафіку і міжкадрового інтервалу Хмельницького університету економіки і права підкоряються розподілу Гаусса (можна пояснити особливостями технології 802.11). Ряди з великим кроком агрегування підкоряються розподілу Гаусса.

При аналізі розподілу тривалості кадрів виявлено, що в мережі найчастіше зустрічаються кадри певного розміру: або з максимальним або з мінімальним розмірами. Можна з певною часткою ймовірності визначити, що тривалість обслуговування кадру постійна. Результати проведеного дослідження дають можливість розробити аналізатор інтенсивності трафіку, який, на відміну від авторегресійних фільтрів, при обчисленні використовує характеристики і властивості самоподібності передбаченого ряду. Це дозволяє компенсувати затримки часу, що виникають при передачі службових пакетів, що переносять інформацію про стан мережі (з відносною похибкою 0,05 -0,2 для трафіку, що володіє властивостями стохастичної самоподібності).

### Література

1. Постановка задачи для построения квазистатических адаптивных алгоритмов маршрутизации / А. М. Комиссаров // Проблемы техники и технологии телекоммуникаций: Сб. докл. VIII Междунар. научн.-техн. конф. - Уфа, УГАТУ, 2007. С. 116-119.
2. Коммутаторы локальных сетей D-Link: учебное пособие. 4-е изд. - М., 2006. - 162с
3. Комиссаров А.М. «Адаптивная маршрутизация в сетях передачи данных с учетом самоподобия трафика», 2011.
4. Платунова С.М. Исследование структуры сети на имитационных моделях // Сб. тезисов докладов VII конференции молодых ученых. Вып. 1. Труды молодых ученых. – 2010. – С. 90.
5. Платунова С.М. Учебно-исследовательская подсистема автоматизированного анализа характеристик сети компьютерных классов // Сб. трудов I Всероссийского конгресса молодых ученых. – 2012. – С. 51–54 Problems of technology and telecommunications tehnologii: Sat.
6. Березка М.П. "Маршрутизация сообщений в сети передачи данных АСУ "Экспресс-2". Журнал «Автоматика, связь, информатика», №10, 2000г с.40-7. ISSN 0005-2329

### References

1. Statement of the problem to construct a quasi-static adaptive routing algorithm / AM Commissioners // Dokl. VIII Intern. nauchn.-tehn. Conf. - Ufa, USATU, 2007, pp 116-119.
2. LAN switches D-Link: a tutorial. 4th ed. - M., 2006. - 162c
3. Commissioners AM "Adaptive routing in data networks based on self-similar traffic," in 2011.
4. Platonova SM Research on the structure of the network simulation models // Proc. Abstracts VII conference of young scientists. Vol. 1. Proceedings of the young scientists. - 2010. - P. 90.
5. Platonova SM Teaching and research subsystem is the automated analysis of network features computer classes // Proc. of the I All-Russian Congress of Young Scientists. - 2012. - P. 51-54
6. Birch MP "Routing messages in a data network ACS" Express-2. "Journal" Automation, Communication, Computer Science », №10, 2000 s.40-7. ISSN 0005-2329

Рецензія/Peer review : 8.12.2014 р.

Надрукована/Printed : 1.1.2015 р.