

МОДЕЛЮВАННЯ АГРЕГОВАНОГО ТРАФІКУ ПЕРЕДАЧІ ДАНИХ НА ОСНОВІ МОДЕЛІ ON / OFF

Анотація. Розглянуто підхід до розробки моделі агрегованого трафіку, який породжено декількома джерелами з послідовністю ON/OFF інтервалів, що чергуються, та тривалість яких підлягає Парето розподілу. Аналіз статистичних характеристик й оцінка показника Херста процесів, які синтезовані, підтвердили гіпотезу про самоподібну природу трафіку.

Ключові слова: агрегування, ON/OFF модель, розподілення Парето, трафік, параметр Херста.

Постановка проблеми

Постійно прискорюваний розвиток нових інфокомунікаційних технологій у різних сферах діяльності: виникнення нових протоколів передачі даних, розширення апаратних можливостей мережевого встаткування, зростання числа підключених абонентів і сумарний обсяг трафіку. В той же час, такий інтенсивний розвиток спричиняє ряд проблем. Одна з них полягає в тім, що розвиток мережевого обладнання й транспортних протоколів повинен опиратися на адекватні математичні моделі параметрів трафіку й інструменти моделювання мережевих процесів [1, 2]. Характер трафіку визначається певними факторами - від поведження користувачів або прикладного програмного забезпечення, до протоколів передачі й використовуваного обладнання. Вочевидь, що макропараметри мережевого трафіку (на відносно великих часових інтервалах) визначаються людиною. Однак характер трафіку на інтервалах порядку мікросекунд визначаються здебільшого транспортними протоколами, мережевим обладнанням і серверним програмним забезпеченням. З цього приводу, дослідження основних характеристик агрегованого мережевого трафіку, таких як параметр Херста, завантаження мережі на фоні інтенсивного мережевого трафіку, є актуальним завданням [3, 4].

Аналіз останніх досліджень

У роботах [5-7] показано, що процес агрегування трафіку множини окремих джерел в об'єднаній мережі приводить до стрибкоподібних змін інтенсивності трафіку і його можна розглядати як фрактальний процес, статистичні характеристики якого мають властивості масштабної інваріантності.

Формулювання цілей статті (постановка завдання)

Мета роботи полягає в розробці імітаційної моделі трафіку, що утворений на комутаційних вузлах мережі при агрегуванні окремих інформаційних потоків, які спрямовуються по одному маршруту та аналізу характеристик побудованої моделі.

Основна частина

Самоподібний мережевий трафік може бути породжений мультиплексуванням декількох джерел з послідовністю ON/OFF інтервалів, що чергуються, та тривалість яких підлягає Парето розподілу. У мережі з комутацією пакетів ON періоди відповідають переданій серії пакетів (наприклад, IEEE стандарт 802.3), OFF інтервали - періоди з відсутністю передачі пакетів.

Середнє значення розподілу Парето дорівнює:

$$E(x) = \frac{\alpha b}{\alpha - 1} \quad (1)$$

Формула для генерації розподілу Парето

$$X_{PARETO} = \frac{b}{U^{1/\alpha}} \quad (2)$$

де U — рівномірно розподілені значення в діапазоні $(0, 1]$,

α - параметр форми (індекс «хвіста»),

b - мінімальне значення x .

При генеруванні модельного трафіку, бажано апіорно мати навантаження мережі й кожного джерела. З очевидь, що результуюче навантаження L у такому випадку просто дорівнює сумі навантажень L_i , що породжені від кожного окремого джерела i . Дано N джерел,

$$L = \sum_{i=1}^N L_i \quad (3)$$

Таким чином, важливо мати можливість одержання гарної оцінки навантаження, що створюється одним джерелом. Навантаження породжене одним джерелом представляє відношення середньої трива-

лості передачі пакета до суми середньої тривалості передачі пакета й середнього значення тривалості розриву передачі, або це відношення середнього розміру тривалості ON періоду до суми середніх значень тривалості ON й OFF періодів.

$$L_i = \frac{\overline{ON}_i}{\overline{ON}_i + \overline{OFF}_i} \quad (4)$$

Формула (1) дає середнє значення істинного розподілу Парето. Однак комп'ютерна реалізація, що використовує рівняння (2), формує псевдо-Парето розподіл. Одна із проблем пов'язана з тим, що комп'ютери не можуть генерувати довільно велике значення, однак, будь-який істинний розподіл Парето досить великої довжини буде мати значення, які перевищують діапазон, який згенеровано комп'ютером. Таким чином, це і є розподіл усіченого значення.

Позначимо через S невелике ненульове значення, породжене рівномірним випадковим генератором. Тоді, згенеровані Парето-розподілені значення не перевищать q :

$$q = \frac{b}{S^{1/\alpha}} \quad (5)$$

Тоді, середнє значення розподілу Парето може бути розраховане як:

$$E(x) = \int_b^q x f(x) dx = \int_b^q x \frac{ab^\alpha}{x^{\alpha+1}} dx = ab^\alpha \int_b^q \frac{dx}{x^\alpha} = ab^\alpha \frac{ab}{\alpha-1} \left| \frac{q}{b} = \frac{ab}{\alpha-1} \left[1 - \left(\frac{b}{q} \right)^{\alpha-1} \right] \right. \quad (6)$$

Підставляючи (5) в (6), одержимо

$$E(x) = \frac{ab}{\alpha-1} \left[1 - S^{\frac{\alpha-1}{\alpha}} \right] \quad (7)$$

Рівняння (7) дає середнє значення усіченого значення розподілу Парето. Тепер, якщо є дані навантаження L_i і розмір k пакетів для даного джерела, можемо знайти мінімальне значення періоду OFF.

Для початку, знайдемо середнє значення OFF періоду i . З рівняння (1) одержуємо

$$\overline{OFF}_i = \overline{ON}_i \left(\frac{1 - L_i}{L_i} \right) \quad (8)$$

Позначимо M_{OFF} й M_{ON} мінімуми OFF й ON періодів відповідно. Вище зазначалося, що мінімальний розмір пакета передачі дорівнює одному пакету, тобто, $M_{ON}=1$. Тоді,

$$\frac{M_{OFF}\alpha_{OFF}}{\alpha_{OFF}^{-1}} \left[1 - S \frac{\alpha_{OFF}^{-1}}{\alpha_{OFF}} \right] = k \frac{M_{ON}\alpha_{ON}}{\alpha_{ON}^{-1}} \left[1 - S \frac{\alpha_{ON}^{-1}}{\alpha_{ON}} \right] \left(\frac{1-L_i}{L_i} \right) \quad (9)$$

де α_{ON} — параметр форми для ON періодів, і α_{OFF} — параметр форми для OFF періодів.

Позначаючи $T_{ON} = \frac{\alpha_{ON}^{-1}}{\alpha_{ON}}$ та $T_{OFF} = \frac{\alpha_{OFF}^{-1}}{\alpha_{OFF}}$, одержимо

$$M_{OFF} = k \frac{T_{OFF}}{T_{ON}} \times \frac{1-S^{T_{ON}}}{1-S^{T_{OFF}}} \times \left(\frac{1}{L_i} - 1 \right). \quad (10)$$

Таким чином, з огляду на значення k , L_i , α_{ON} , і α_{OFF} , формула (10) дає значення для M_{OFF} , що приведе до навантаження лінії, наближеного до L_i .

Проте, якщо генерувати трафік, використовуючи наведені вище формули, середні значення для періодів включення й вимикання в згенерованій серії будуть відрізнятися від необхідних. Проблема, полягає в тому, що комп'ютери генерують Парето-розподілені значення за формулою (2). У той час як розподіл Парето припускає безперервний простір значень, генеруються дискретні значення з певною рівномірністю. Парето-розподілення досягається за рахунок більш високої щільності значень у нижній частині шкали. При цьому, якщо будувати функцію розподілу шляхом агрегування значень протягом деякого заданого вікна, то незалежно від того, наскільки великий розмір вікна, в «хвості» відстань між двома сусідніми точками може перевищувати розмір вікна. Це означає, що деякі вікна будуть містити нульові вибірки, навіть якщо кількість значень прагне до нескінченності. Це вносить погрішність у середній розмір періодів включення й вимикання.

Щоб виправити цю погрішність, розрахункові значення ON й OFF періодів необхідно помножити на ваговий коефіцієнт C . Таким чином, формула (10) приймає вид

$$M_{OFF} = k \frac{C_{ON}}{C_{OFF}} \times \frac{T_{OFF}}{T_{ONN}} \times \frac{1 - S^{T_{ON}}}{1 - S^{T_{OFF}}} \times \left(\frac{1}{L_i} - 1 \right). \quad (11)$$

Якщо ми вибираємо α_{ON} й α_{OFF} тими ж самими, то рівняння (11) спрощується:

$$M_{OFF} = k \left(\frac{1}{L_i} - 1 \right). \quad (12)$$

Це, однак, може обмежити застосування генератора трафіку в реальних умовах. Імовірність наявності дуже великого OFF періоду вище, ніж імовірність наявності дуже великого ON періоду. Це означає, що параметр форми α_{ON} повинен бути більше, ніж α_{OFF} .

На основі моделі, що була запропонована розроблена схема імітаційного моделювання в системі OPNET і проведений статистичний аналіз отриманих процесів. Трафік надсилався від кількох вузлів одному одержувачеві. Змінювались значення α для розподілення Парето та співвідношення довжин періодів ON та OFF. Результати аналізу однієї з реалізацій представлені на рисунках нижче.

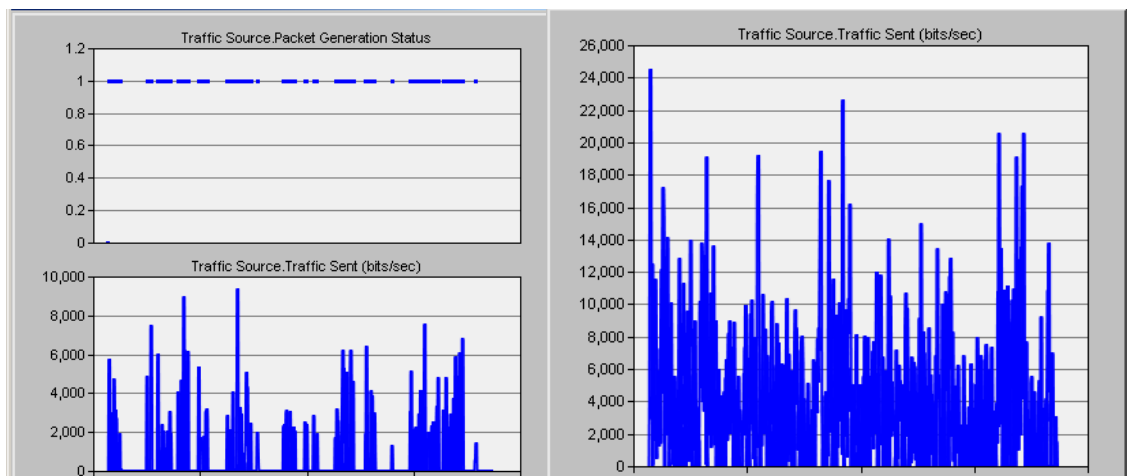


Рисунок 1 - Вихідний трафік для одного з вузлів мережі та агрегований для всієї системи

На рисунку 2 представлені розрахунки показників Херста різними методами.

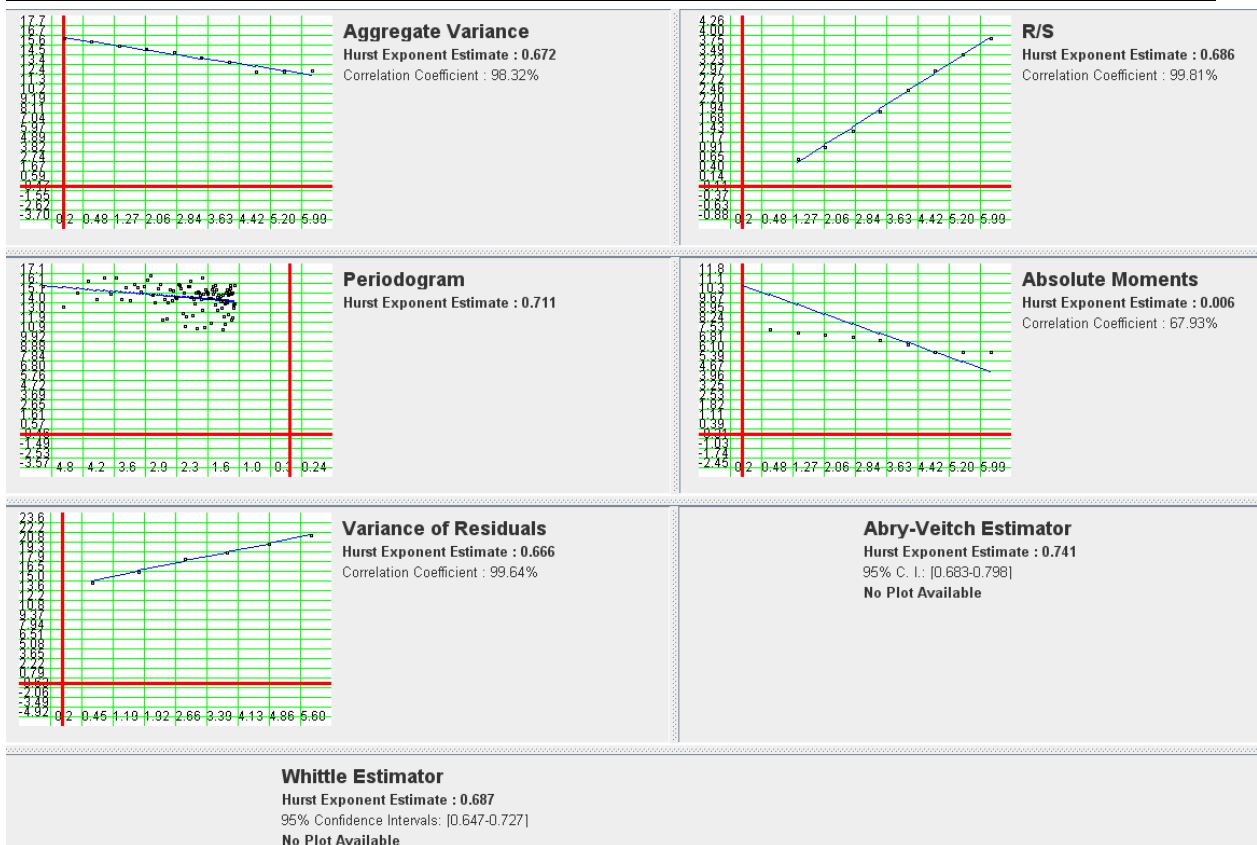


Рисунок 2 - Показники Херста для одного з джерел трафіку

Отриманий агрегований трафік від 50 джерел представлено на рисунку 3.

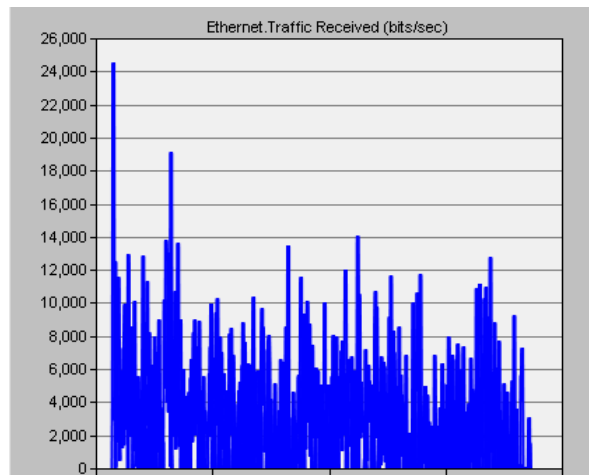


Рисунок 3 - Отриманий агрегований трафік

Параметр Херста може бути мірою оцінки довгострокової залежності часового ряду. Оцінка параметра не тільки може допомогти зробити висновок про самоподібність процесу, але й дозволить надалі

застосувати до нього ряд математичних методів по прогнозуванню фрактальності процесів .

Як й очікувалося, значення параметра вказало на існування довгострокової залежності й само подібності (рисунок 2 та 4).

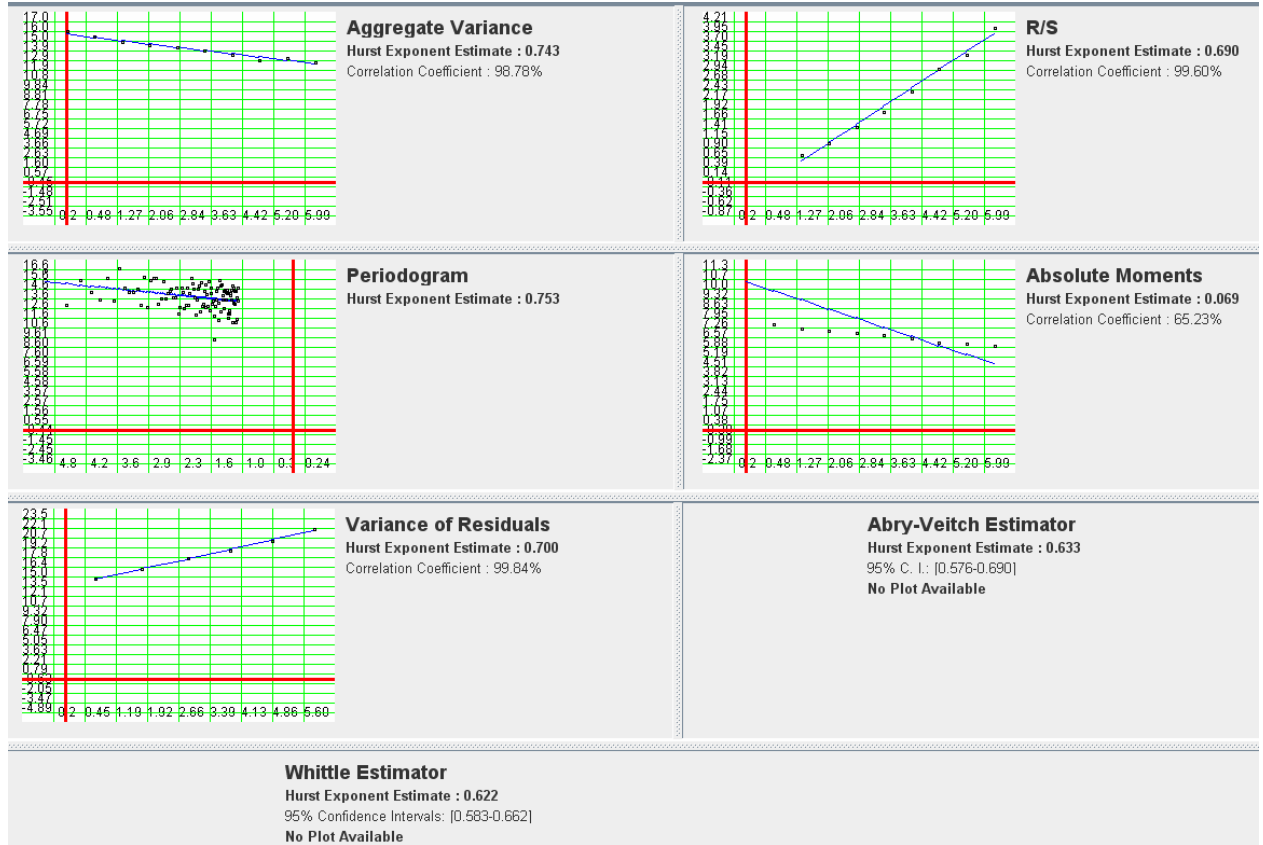


Рисунок 4 - Показники Херста для агрегованого трафіку

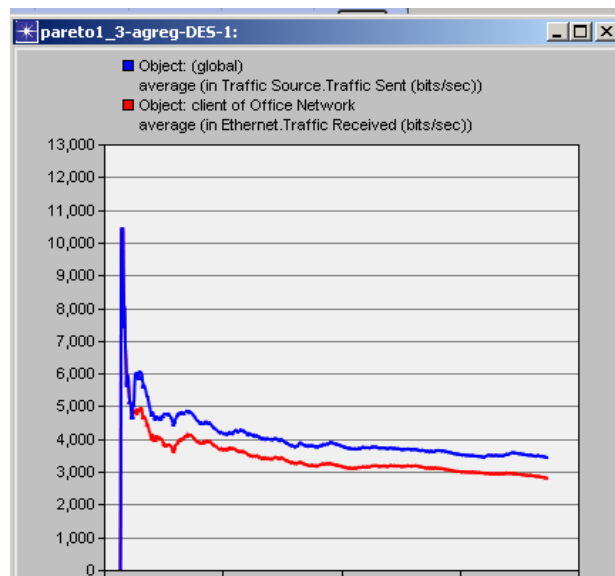


Рисунок 5 - Отриманий та відправлений трафік при значенні $\alpha_{ON} = 6$ та $\alpha_{OFF}=0.8$

Запропоновано метод генерації трафіку із заданими ступенем самоподібності й довгострокової залежності на основі ON-OFF моделі й комплексна модель оцінки характеристик агрегованого трафіку, що використовує оцінку середньої інтенсивності трафіку, а також ступеня його самоподібності й довгострокової залежності.

Висновки та перспективи подальших досліджень

Одержала подальший розвиток математична ON/OFF-модель для агрегованого потоку даних. На основі розроблених математичних моделей було проведено імітаційне моделювання у середовищі OPNET. Оцінка показника Херста різними методами та аналіз статистичних характеристик трафіку й синтезованих процесів підтвердили гіпотезу про самоподібну природу розглянутого потоку даних.

Результати моделювання достатньо повно узгоджуються з результатами відомих числових та експериментальних досліджень та дають можливість прогнозування навантаження в телекомунікаційних мережах.

ЛІТЕРАТУРА

1. Столлингс В. Современные компьютерные сети : пер. с англ. - СПб.: Питер, 2003. - 783 с.
2. Таненбаум Э. Компьютерные сети : пер. с англ. - СПб.: Питер, 2003. - 992 с.
3. Иванов И.П. Мониторинг ресурсов узлов корпоративной сети /Иванов И.П., Бойченко М.К. // Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. - № 2. – Москва, 2010. - С. 114-120.
4. Бойченко М.К. Доступность ресурсов транспортных подсистем компьютерных сетей /Бойченко М.К., Иванов И.П., Кондратьев А.Ю.// Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. - № 3. – Москва, 2010. - С. 103-118.
5. Крылов, В.В. Теория телетрафика и ее приложения [Текст] / В.В. Крылов, С.С. Самохвалова. – СПб.: БХВ-Петербург, 2005. – 288 с.
6. Олифер, В.Г. Компьютерные сети. Принципы, технологии, протоколы [Текст] / В.Г. Олифер, Н.А. Олифер. – СПб.: Питер, 2012. – 943 с.
7. Куроуз, Дж. Компьютерные сети. 2-е изд [Текст] / Дж. Куроуз, К. Росс. – СПб.: Питер, 2004. – 765 с.