

УДК 621.324

В.М. Тупкало, Г.А. Кучук, Я.Ю. Стасєва

ЧАСТОТНО-ЧАСОВІ ДОСЛІДЖЕННЯ ТРАФІКА ТЕЛЕКОМУНІКАЦІЙНИХ МЕРЕЖ СИСТЕМ ШВИДКОГО РЕАГУВАННЯ

Проведені частотно-часові дослідження трафіка телекомунікаційних мереж підтримки систем швидкого реагування, які підтвердили його масштабно-інваріантні властивості у надзвичайних ситуаціях.

Ключові слова: трафік, телекомунікаційна мережа, масштабно-інваріантні властивості, система швидкого реагування, надзвичайні ситуації.

Постановка завдання та аналіз літератури

Для прогнозування трафіка телекомунікаційних мереж необхідне проведення дослідження експериментальних даних, що характеризують відповідні процеси, які мають складну структуру, особливо у системах швидкого реагування. Це затрудняє розуміння механізмів міжмережевої взаємодії та обмежує застосування класичних статистичних і динамічних моделей для визначення найбільш істотних особливостей розглядаємих мережевих процесів. При цьому можна відокремити відносно повільні процеси встановлення віртуального з'єднання за допомогою протоколів маршрутизації та обміну параметрами з'єднання між джерелом і приймачем повідомлення [1] від швидких процесів, що у надзвичайних режимах призводять до варіацій затримок при проходженні пакетів через проміжні вузли віртуального з'єднання [2]. Управління інтенсивністю послідовності пакетів у обох випадках здійснюється з використанням механізму зворотного зв'язку на основі обробки інформації про успішну передачу раніше посланих пакетів [3], часові параметри процесів залежать від випадкових затримок, що виникають між передачею пакетів і прийомом підтверджень про їх успішну доставку [4], пов'язаних із проходженням пакетів через буфери пристроїв маршрутизації, а також в результаті зміни маршрутів передачі самих пакетів [5]. Тому статистичний аналіз частотно-часових характеристик є важливою складовою дослідження мережевих процесів, які виникають у телекомунікаційних мережах підтримки систем швидкого реагування.

Для проведення аналізу необхідно провести оцінку характеру зміни моментних функцій на часових інтервалах змінної довжини, що призводить до необхідності агрегування відліків окремих вимірювань з наступним обчисленням моментних функцій різних порядків. При цьому використання вейвлет-перетворення дозволяє визначити властивості мере-

жевих процесів на основі аналізу характеру зміни коефіцієнтів розкладання спостережуваного сигналу в ряди за ортогональним вейвлет-базисом [6]. За допомогою таких базисів можна ефективно виконати процедуру, аналогічну інтерполяції сигналу в часовій області за його уявленням за допомогою перетворення Фур'є. Кожна з функцій базису характеризує як певну частотну складову сигналу, так і локалізацію цієї складової в просторі (часі), де цей сигнал розглядається. Таким чином, на відміну від традиційного перетворення Фур'є забезпечується двовимірне розгортка досліджуваного одновимірного сигналу, а та час розглядаються як дві незалежні змінні. Такий підхід дає можливість аналізувати властивості сигналу разом в часовому і частотному просторах, що дозволяє одержати локальну високо-частотну і глобальну низько-частотну інформацію про властивості досліджуваного сигналу. В цьому випадку вейвлет-перетворення виявляє внутрішню структуру істотно неоднорідних потоків даних і дозволяє визначити масштабно-інваріантні властивості трафіку.

Мета статті – перевірка гіпотези про присутність властивості самоподібності для реалізацій трафіку телекомунікаційних мереж підтримки систем швидкого реагування у надзвичайних ситуаціях у вигляді відліків значень RTT (Round Trip Time) і проведення аналізу характеру нерегулярних затримок, які не зникають навіть при усереднюванні на досить тривалих інтервалах часу.

Умови експерименту

Значення RTT було набуто за допомогою стандартної програми ping, яка з періодичністю раз на секунду відправляла контрольний пакет за вказаною IP-адресою і обчислювала значення RTT. Видалені вузли мережі системи швидкого реагування вибиралися за наступними ознаками: вузол повинен бути достатньо видаленим, щоб помітніше виявлялися флуктуації трафіку в мережі; канали зв'язку до вузла

повинні бути достатньо завантажені, але не переважані, оскільки інакше втрачається адекватність одержуваної картини про процеси в віртуальних з'єднаннях досліджуємої телекомунікаційної мережі; маршрут з'єднання до вузла не повинен змінюватися протягом всього експерименту; втрати тестових пакетів повинні бути мінімальні, оскільки при пропуску пакету з'являється пропуск у ряді значень відліків RTT, що небажано і при великому відсотку втрачених пакетів може викликати значні зміни в спостережуваній картині. На основі цих критеріїв для тестування були вибрані вузли, розташовані в різних регіонах України. Для отримання адекватних результатів тривалість реалізації експерименту було близько 2,5 годин, що дозволило одержати від 6000 до 8000 відліків значень RTT. Така тривалість потрібна для того, щоб перевірити присутність асимптотичних ефектів, пов'язаних з проявом протяжних статистичних залежностей. Треба зазначити, що прямий і зворотний маршрути проходження тестового пакету в територіально розподіленій мережі можуть значно розрізнятися, і навіть якщо вони співпадають, то часи проходження пакету до віддаленого вузла і назад можуть бути різного порядку. Для перевірки цієї гіпотези був проведений експеримент між двома маршрутизаторами по величезно завантаженій ділянці цієї ж мережі. Часи отримання і прибуття назад тестового пакету знімалися за допомогою програми tcpdump на обох маршрутизаторах (час на маршрутизаторах було синхронізовано з погрешністю до 50 мс). Аналіз спектральних характеристик одержаних наборів даних після перетворення Фур'є наведений на рис. 1. З цих даних виразно видна перевага низькочастотних гармонік над високочастотними. Такі результати свідчать на користь гіпотези про властивість протяжної залежності досліджуваного трафіку, але ще не є повністю підтверджуючими її.

Для точнішого підтвердження характеру статистичних зв'язків в розглянутих реалізаціях на рис. 2 надані їх автокореляційні функції. З одержаних графіків можна відзначити, що значні кореляції спостерігаються в діапазонах від 150 до 400 с, а самі функції носять яскраво виражений степеневий характер. Такий вигляд автокореляційних функцій узгоджується з характером зміни спектральної щільності.

Для оцінки чисельних характеристик можна використовувати залежності $\log \{ \text{var}[T(m)] \}$ від $\log(m)$, де для кожного $m = 1, 2, \dots$, що розділяють початковий часовий ряд $T = \{T_i; i > 1\}$ на блоки, будується агрегований процес $T(m) = \{T(m)(k)\}$. Якщо процес не має властивості масштабної інваріантності, то при $m \rightarrow \infty$ виконується асимптотична рівність $\text{Var}[T^{(m)}] \sim m^{-1}$, тобто спостерігається лінійна залежність, а для процесів, що мають фрактальні влас-

твості, виконується асимптотичне співвідношення $\text{Var}[T^{(m)}] \sim m^{-\beta}$ при $0 < \beta < 1$, де показник β пов'язаний з параметром Херста H як $H = 1 - \beta / 2$.

Точніші оцінки можна одержати, використовуючи відліки часу прямого і зворотного проходження тестових пакетів. Таке обчислення значення RTT-пакетів вимагає для кожної посылки чотирьох часових відліків.

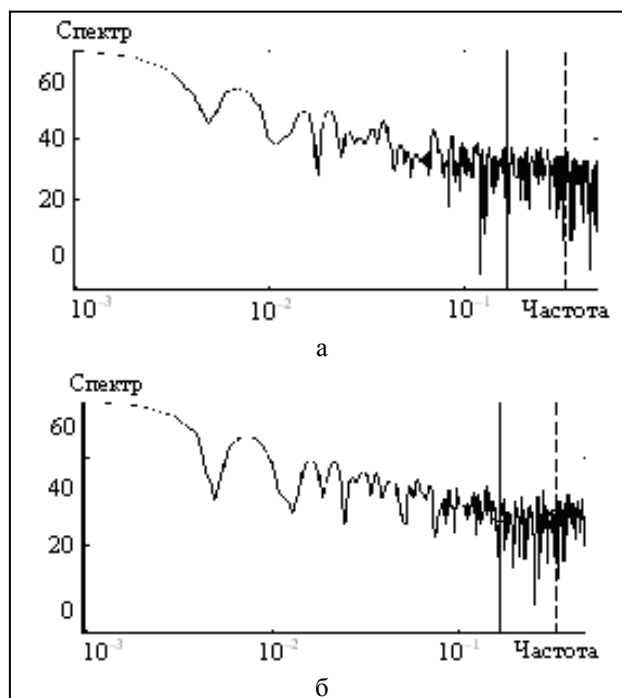


Рис. 1. Спектр реалізації відліків RTT для першого (а) і другого (б) маршрутів

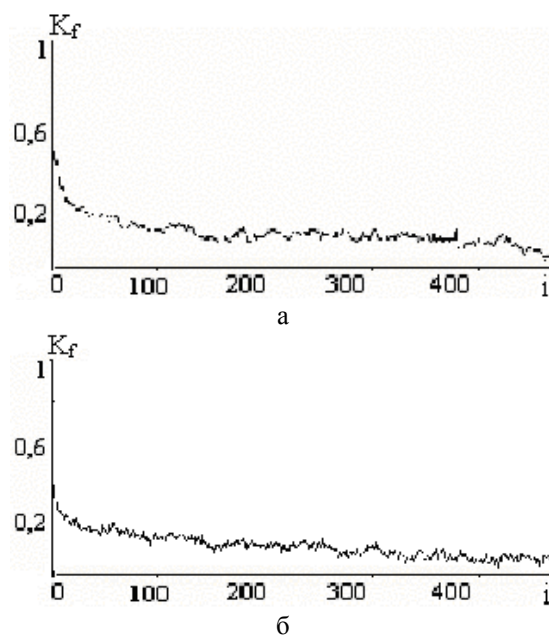


Рис. 2. Автокореляційна функція реалізації відліків RTT для першого (а) і другого (б) маршрутів

Коли посилається тестовий пакет, то записується час його відправки T_1 , отримання приймачем пакету відповідає відліку T_2 , час відправки зворотного пакету – це відлік T_3 , отримання пакету джерелом визначається моментом часу T_4 .

В результаті можна відзначити:

величина затримки дорівнює $T_{RTT} = (T_4 - T_1)$ (рис. 3,а);

значення $(T_2 - T_1)$ складає час проходження пакету в прямому напрямку (рис. 3,б);

значення $(T_4 - T_3)$ – у зворотному (рис. 3,в).

Значення T_1, T_4 легко одержати з файлу трасування, який фіксується на вузлі-джерелі, а значення T_2, T_3 – з файлу трасування, який фіксується на вузлі-приймачі.

Несиметричний характер одержаних залежностей, який бачимо на рис. 3, є очевидним. Затримка при зворотному проходженні тестового пакету мінімальна через те, що канал в прямому напрямку сильно завантажений, а у зворотному напрямку йдуть тільки пакети підтвердження TCP-протоколу. Такий характер TCP-з'єднання треба враховувати при різних дослідженнях.

Для оцінки впливу величини дисперсії для різних масштабів спостереження можна побудувати залежність чинника Фано для масштабно-інваріантного процесу. В цьому випадку в подвійному логарифмічному масштабі графік є прямою, за нахилом якої можна судити про ступінь фрактальності процесу. Так, для звичайного процесу, наприклад пуассонівського, виходить пряма, паралельна осі абсцис.

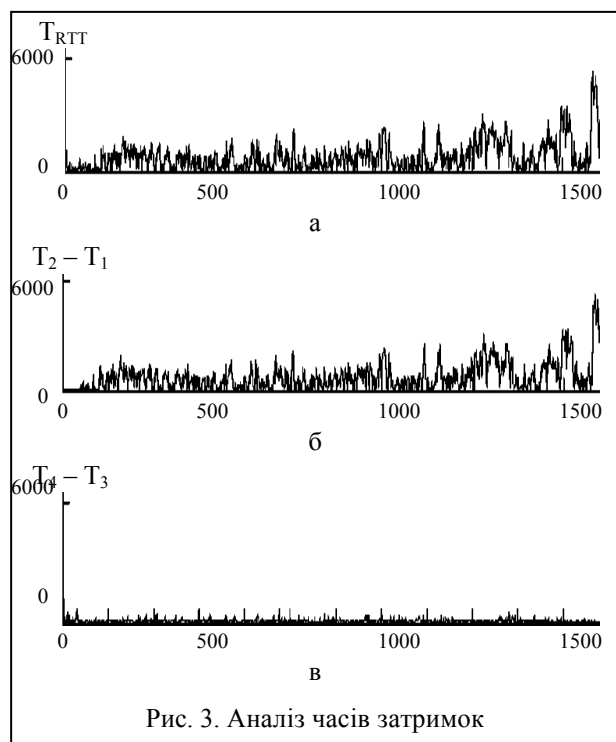


Рис. 3. Аналіз часів затримок

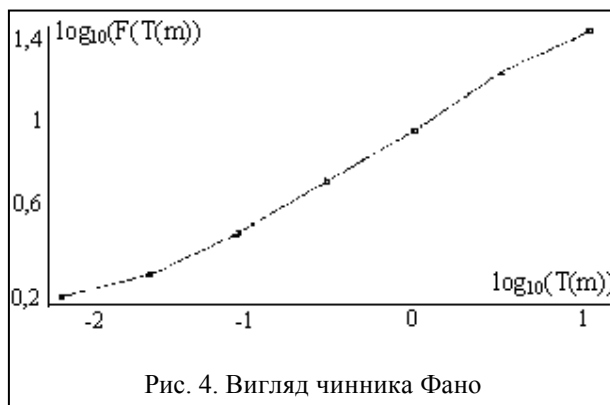


Рис. 4. Вигляд чинника Фано

Побудуємо чинник Фано для реалізації знятого трафіку (рис. 4). З одержаної залежності видно фрактальний характер досліджуваного процесу, а за її нахилом можна визначити значення параметра Херста H .

Як вже згадувалося вище, вейвлет-перетворення дозволяє виявити внутрішню структуру істотно неоднорідних потоків даних і визначити масштабно інваріантні властивості трафіку [6].

Рівень заповнення буферних пристроїв є важливим чинником, що впливає на затримки при передачі пакетного трафіку. Розглянемо підхід до дослідження характеру зміни черг в буфері маршрутизатора з використанням вейвлет-перетворення. Нехай вхідний трафік в буфері представляється значеннями випадкових відліків інтенсивності мережевого трафіку $X(t)$ на інтервалах заданої тривалості L_i , які характеризують приріст завантаження буфера в i -й інтервал спостереження. Поточному інтервалу надамо індекс $i = 0$. Тоді завантаження в каналі за k інтервалів буде дорівнювати

$$F_k = \sum_{i=-k+1}^0 L_i. \quad (1)$$

Якщо швидкість обслуговування пакетів в буфері є постійною величиною, рівною C , то черга у буфері визначатиметься як

$$Q_0 = \max \{Q_{-k} + F_k - Ck, F_{k-1} - (k-1)C, \dots, F_0\}, \quad (2)$$

де Q_0 і Q_{-k} – черги в буфері при розгляданні поточного та k -го часового інтервалу відповідно.

Проаналізуємо характер зміни значення Q_0 . Якщо розглядати довільний інтервал спостереження, тобто $k \in Z_+$, і в якийсь момент $Q_{-k} \geq 0$, то

$$Q_0 \geq \sup_{k \in Z_+} (F_k - kC).$$

Позначимо через t момент часу, найближчий зліва (тобто той, що передує) поточному інтервалу спостереження, коли черга в буфері була порожня (якщо $Q_0 = 0$, то $t = 0$).

Тоді можна записати нерівність, вірну для всіх моментів спостереження:

$$Q_0 = F_t - tC = \sup_{k \in Z_+} (F_k - kC).$$

Якщо до моменту k буфер був пустим, тобто не було черги, то

$$Q_0 = \sup_{k \in Z_+} (F_k - kC). \quad (3)$$

Визначимо інтервал спостереження, у якому з найбільшою ймовірністю відбувається перевищення деякого порогу заповнення буфера:

$$k^* = \arg \sup_{k \in Z_+} P(F_k - kC > b), \quad (4)$$

де b – поріг заповнення буфера.

Величину k^* назвемо критичним інтервалом і за її допомогою визначимо ймовірність перевищення значення порогу

$$q(b) = P(F_{k^*} - k^*C > b). \quad (5)$$

Для визначення оцінки $q(b)$ розглянемо ті моменти часу, яким відповідають координати для вейвлет-перетворення при використанні базисних функцій Хаара.

Для цих моментів можна записати вираз:

$$\bar{Q}_0 = \sup_{m \in \{0, \dots, n\}} (F_{2^m} - C2^m), \quad (6)$$

який апроксимує (3).

Для значення \bar{Q}_0 справедливі нерівності:

$$\bar{Q}_0 \leq Q_0;$$

$$P(Q_0 > b) \geq P(\bar{Q}_0 > b).$$

Аналогічно виразам (4) і (5) введемо наступні характеристики черги в буфері, які розглядаються в дискретні моменти часу:

$$k_d^* = \arg \sup_{k \in \{2^m : m \in \{0, \dots, n\}\}} P(F_k - kC > b);$$

$$q_d(b) = P(F_{k_d^*} - Ck_d^* > b).$$

Через зроблені допущення про властивості \bar{Q}_0 маємо:

$$q_d(b) \leq P(\bar{Q}_0 > b) \leq P(Q_0 > b).$$

Якщо на досліджуваній вибірці, що складається з $T = 2n$ відліків трафіку $X(t)$, значення L_i в (1) відповідають моментам часу $T - 1 + i$, тобто

$$L_i = X(T - 1 + i), \quad I = -1, \dots, -T + 1, \quad (7)$$

то для оцінки величини завантаження буфера F_k можна використовувати властивості вейвлет-перетворення сплесків. Позначаючи вершину бінарного дерева вейвлет-перетворення для $X(t)$ як $U_{0,0}$ (індекс

масштабу без порушення загальності вибраний позитивним), можна записати такий вираз для значень завантаження

$$F_{2^{n-i}} = 2^{-i/2} U_{i,2^i-1}, \quad i = 0, \dots, n. \quad (8)$$

Визначимо величину

$$S_i = F_{2^{n-i}} - C2^{n-i}.$$

Зіставимо умові виконання нерівності $S_i > b_i$ подію, яку позначимо як e_i . Припустимо, що завантаження буфера відбувається в результаті дії незалежних джерел, тому величину S_i можна представити у вигляді суми незалежних випадкових змінних

$$S_i = x_0 + \dots + x_i, \quad i \in [1, n].$$

Звідки для будь-яких значень i із заданого вище інтервалу вірна нерівність, сформульована для умовної ймовірності події e_i

$$P(e_i | e_{i-1}, \dots, e_0) \geq P(e_i), \quad (9)$$

яке можна записати у вигляді

$$P(S_i < b_i | S_l < b_l, l = 0, \dots, i-1) \geq P(S_i < b_i).$$

Введемо дві нові змінні:

$$Y_i = S_i | e_{i-1}, \dots, e_0, \quad i \geq 1;$$

$$Z_i = S_i | e_i, \dots, e_0, \quad i \geq 1.$$

Між цими змінними існують такі залежності:

$$Y_0 = Z_0 = S_0 = x_0;$$

$$\dot{Y}_{i+1} = Z_i + x_{i+1},$$

де x_{i+1} – статистично незалежна від S_j ($j < i$) змінна.

Позначимо як $F(x)$ функцію розподілу випадкової змінної z , тобто

$$F(x) = P(z < x).$$

Тоді для доказу нерівності (9) досить показати, що для відповідних функцій розподілу змінних Y_i та S_i для будь-яких x вірна нерівність

$$F_{Y_i}(x) \geq F_{S_i}(x), \quad i = 0, 1, 2, \dots \quad (10)$$

Для доказу (10) використовуємо метод математичної індукції. Виконання нерівності

$$F_{Y_0}(x) \geq F_{S_0}(x)$$

є очевидним.

Нехай нерівність (10) є вірною для довільного індексу i . Покажемо, що це призведе до виконання нерівності для індексу $i + 1$.

Відповідно до формули Байеса вірогідність $P(e_i | A)$ після того, як мала місце подія A , задовольняє співвідношенню

$$P(e|A) \cdot P(A) = P(e) \cdot P(A|e).$$

Враховуючи, що

$$F_{Z_i}(x) = P(S_i < x | e_{i-1}, \dots, e_0) = P(S_{i-1} + x_i < x | e_{i-1}, \dots, e_0),$$

запишемо

$$F_{Z_i}(x) = \begin{cases} \frac{F_{Y_i}(x)}{F_{Y_i}(b_i)}, & \text{якщо } x < b_i \\ 1, & \text{якщо } x > b_i \end{cases} \geq F_{Y_i}(x). \quad (11)$$

Відзначимо, що $X_i + 1$ статистично незалежна від Z_i випадкова змінна. Тому, враховуючи (10) і (11) і використовуючи для позначення функцій щільності розподілу змінних Z_i і X_i відповідно функції $f_{Z_i}(y)$ і $f_{X_i}(y)$, можна записати:

$$\begin{aligned} F_{Y_{i+1}}(x) &= P(Y_{i+1} < x) = P(Z_i + x_{i+1} < x) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{x-x_{i+1}} f_{Z_i}(Z_i) f_{x_{i+1}}(X_{i+1}) dZ_i dX_{i+1} = \\ &= \int_{-\infty}^{\infty} F_{Z_i}(x - x_{i+1}) f_{x_{i+1}}(X_{i+1}) dX_{i+1} \geq \\ &\geq \int_{-\infty}^{\infty} F_{Y_i}(x - x_{i+1}) f_{x_{i+1}}(X_{i+1}) dX_{i+1} \geq \\ &\geq \int_{-\infty}^{\infty} F_{S_i}(x - x_{i+1}) f_{x_{i+1}}(X_{i+1}) dX_{i+1} = \\ &= P(S + x_{i+1} < x) = F_{S_{i+1}}(x). \end{aligned}$$

Отримана нерівність доводить справедливість (9). Тепер, використовуючи одержані співвідношення, можна записати оцінку для ймовірності перевищення чергою \bar{Q} порогу b . Використовуючи (6), (7) і (9), одержимо нерівність

$$\begin{aligned} q_d(b) &= P(\bar{Q} > b) = 1 - P(\bar{Q} < b) = 1 - P\left(\bigcap_{i=0}^n e_i\right) = \\ &= 1 - P(e_0) \prod_{i=1}^n P(e_i | e_{i-1}, \dots, e_0) \leq 1 - \prod_{i=1}^n P(e_i) = u(b), \end{aligned}$$

яка визначає верхню межу для досліджуваної ймовірності перевищення в тому випадку, якщо події e_i можна вважати незалежними.

Враховуючи нерівність (5) та її наслідок у формі нерівності $\bar{Q} \leq Q$, запишемо, що

$$P(\bar{Q} > b) \leq P(Q > b).$$

Тому величина $u(b)$ є верхньою межею і для вірогідності $P(Q > b)$. В результаті маємо таке співвідношення, в якому як обмеження виступають припущення про допустимість апроксимації (5) і використовуються позначення з (7):

$$P(Q > b) \approx u(b) = 1 - \prod_{i=0}^n P(F_{2^{n-i}} < b + C2^{n-i}).$$

Для набору реалізацій трафіку, прирости якого характеризуються протяжними залежностями з параметром Херста $H = 0,7$, у табл. 1 наведені значення $u(b)$ і оцінки ймовірності перевищення межі b_i з використанням даних програми моделювання ТСП-з'єднання.

Таблиця 1

Оцінка ймовірності перевищення межі b_i

b_i	0	100	200	500	1000
$u(b)$	0,7765	0,1914	0,0491	0,0018	0,0002
P_{b_i}	0,7103	0,1556	0,0322	0,0014	0,0001

Висновки

З наведених результатів виходить, що можливість використання визначеної вище оцінки перевищення верхньої межі підтверджується для вельми великого діапазону зміни величини b_i , що дозволяє рекомендувати її для налагодження параметрів мережевих пристроїв маршрутизації або управління трафіком в магістральних каналах, де найсильніше виявляються протяжні статистичні залежності. Наведені частотно-часові дослідження трафіка телекомунікаційних мереж підтримки систем швидкого реагування підтвердили його масштабно-інваріантні властивості у надзвичайних ситуаціях Метою подальших досліджень у цьому напрямку є розробка моделей розрахунку непараметричних статистик другого порядку процесів, які відбуваються у віртуальних каналах мереж підтримки систем швидкого реагування при надзвичайних ситуаціях.

СПИСОК ЛІТЕРАТУРИ

1. Стеклов В.К., Беркман Л.Н. Телекомунікаційні мережі. – К.: Техніка, 2001. – 392 с.
2. Leland W., Taqqu M., Willinger W. On the self-similar nature of IP-traffic // IEEE/ACM Transactions on Networking. – 1997. – № 3. – P. 423 – 431.
3. Захаров Г.П., Симонов М.В., Яновский Г.Г. Службы и архитектура ШЦСИО. – М.: Эко-Трендз, 1993. – 102 с.
4. Варакин Л.Е. Введение в теорию инфокоммуникаций. Ч.1 // Электросвязь. – 2000. – № 2 (14). – С. 2 – 11.
5. Кучук Г.А., Иванов В.К., Шалапин В.Н. Моделирование эластичного трафика точечного источника мультисервисной сети // Радиотехника. – Х.: ХТУРЭ, 2005. – Вып. 143. – С. 137 – 141.
6. Кучук Г.А. Фрактальный гауссовский шум в трафиковых трассах // Системи обробки інформації: Зб. наук. пр. – Х.: ХВУ, 2004. – Вип. 3. – С. 91 – 99.

Надійшла 3.05.2006

Рецензент: д-р техн. наук професор В.І. Карпенко, Харківський університет Повітряних Сил ім. Івана Кожедуба.