

УДК 004.7(045)

Альомар Мхамад Ібрагім Ахмад

Аспірант кафедри інформаційних технологій

Київський національний університет будівництва і архітектури, Київ

ПРОГНОЗУВАННЯ САМОПОДІБНОГО ТРАФІКА У ПАКЕТНИХ МЕРЕЖАХ

***Анотація.** Надано аналіз характеристик самоподібного (фрактального) пакетного трафіку. Запропоновано метод прогнозування самоподібного трафіку, що може бути використаний для моделювання потоків даних у пакетних мережах. Розглянуто алгоритм прогнозування пульсацій у потоці пакетів, що просуваються через порти комутаторів пакетних мереж. Розроблений алгоритм базується на процедурі Бокса-Дженкінса. Результати прогнозу, що отримуються за цим алгоритмом, базуються на великій кількості даних, значення котрих тим більше впливають на результат прогнозу, чим ближче ці дані до моменту прогнозування. Такі умови є характерними під час динамічного керування пропускнуою здатністю пакетних комутаторів.*

***Ключові слова:** комп'ютерні мережі; вузлове обладнання; адаптивне управління; пакетний трафік; тренд трафіку*

***Аннотация.** Представлен анализ характеристик самоподобного (фрактального) пакетного трафика. Предложен метод прогнозирования самоподобного трафика, который может быть использован для моделирования потоков данных в пакетных сетях. Рассмотрен алгоритм прогнозирования пульсаций в потоке пакетов, которые продвигаются через порты коммутаторов пакетных сетей. Разработанный алгоритм базируется на процедуре Бокса-Дженкинса. Результаты прогноза, которые получают согласно этому алгоритму, базируются на большом количестве данных, значения которых тем больше влияют на результат прогноза, чем ближе эти данные к моменту прогнозирования. Такие условия являются характерными во время динамического управления пропускной способностью пакетных коммутаторов.*

***Ключевые слова:** компьютерные сети; узловое оборудование; адаптивное управление; пакетный трафик, тренд трафика*

***Abstract.** The analysis of descriptions fractal package traffic is given. It is shown that fractal processes have a hyperbolically going out coefficient of correlation. At averaging of fractal process dispersion of it selective middle diminishes slower than size that reverses to the size of selection. At that time as for traditional stationary casual processes dispersion diminishes inversely proportional to the sample size. It is shown also, that a row that forms the successive values of cross-correlation function of fractal process goes away. It is proposed the method of fractal traffic prediction that can be used for the design of streams given in package networks. The algorithm of prediction of pulsations is considered in the stream of packages that move up through ports of switchboards of package networks. The worked out algorithm is based on procedure of Box-Djenkins. The results of prognosis, that turn out after this algorithm, are based on plenty of data value of that the anymore influence on the result of prognosis, than nearer these data to the moment of prognostication. Such terms are characteristic during the dynamic management of package switchboards a carrying capacity.*

***Keywords:** Networks; node equipment; adaptive control; package traffic; trend of traffic*

Постановка проблеми

У багатьох задачах інженерії трафіка бажано мати можливість його прогнозування. Проте внаслідок непередбачуваних пульсацій пакетного

трафіка здійснювати його прогнозування виявляється нетривіальним завданням. Результати експериментальних досліджень вказують на можливість моделювання пакетного трафіка у вигляді фрактального (самоподібного) процесу.

Мета статті

Мета статті – на основі аналізу фрактальних властивостей пакетного трафіку запропонувати метод його прогнозування. Розробити відповідний алгоритм прогнозування пульсацій у потоці пакетів, що просуваються через порти комутаторів пакетних мереж. Розроблений алгоритм має базуватися на процедурі Бокса-Дженкінса.

Аналіз фрактальних властивостей трафіка в пакетних мережах

Процесам, які моделюють потоки пакетів у сучасних мережах передачі даних, притаманний ряд властивостей, що значно відрізняє їх від потоків, розглянутих у класичній теорії телетрафіка. При дослідженні потоків пакетів звичайно мають справу з асимптотично самоподібними процесами, вичерпний аналіз котрих наведено у [1]. У рамках цієї ж роботи під час дослідження потоків будемо користуватися комплексом взаємозалежних понять: самоподоба, масштабування, довгострокова залежність та розподіл з вагомим хвостом (РВХ). Ці поняття характеризують властивості, якими самоподібні процеси відрізняються від класичних випадкових процесів, наприклад пуасонівського.

Неформально самоподібний (або фрактальний) процес визначається як випадковий процес, статистичні характеристики якого проявляють властивості масштабування. Графічне відображення будь-якого самоподібного процесу істотно не змінюється при розгляді у різних масштабах за шкалою часу. Зокрема, на відміну від випадкових процесів, у яких відсутні фрактальні властивості, не відбувається швидкого «згладжування» самоподібного процесу при його усередненні за шкалою часу. Це означає, що такий процес при усередненні зберігає схильність до сплесків.

Нехай $\{X_k; k = 0;1;2;...\}$ – півнескінченний відрізок стаціонарного випадкового процесу X дискретного аргументу (часу) $t=1, \dots, k$, що моделює потік пакетів у каналі транспортування даних певної пакетної мережі. З огляду на стаціонарність такого процесу і припущення про існування й скінченність його двох перших моментів, уведемо такі позначення:

$m = E[X_t]$ – середнє значення (або математичне сподівання) процесу;

$\sigma^2 = E[X_t - m]^2$ – дисперсія процесу;

$R(k) = E[(X_{t+k} - m)(X_t - m)]$ – кореляційна

функція процесу;

$r(k) = R(k) / R(0) = R(k) / \sigma^2$ – коефіцієнт кореляції процесу.

Позначимо через $\{X^{(m)}\}$ усереднений за блоками довжини m процес X , компоненти якого визначаються рівністю

$$X_k^{(m)} = \frac{1}{m} \sum_{i=km-m+1}^{km} X_i \quad (1)$$

При моделюванні мережного трафіка значення X_k інтерпретуються як кількість пакетів (рідше – як сумарний об'єм даних у байтах), що надійшли в канал транспортування пакетів протягом k -го інтервалу часу. Вихідний процес при цьому вже є усередненим. У деяких випадках, коли є необхідність уникнути такого початкового усереднення, розглядається точковий процес або потік подій, тобто послідовність моментів надходження одиничних пакетів у канал.

Випадковий процес $X(t)$ є самоподібним з параметром Херста H , якщо $X(t)$ та моделюючий його процес $a^{-H}Z(at)$ мають ідентичні кінцевомірні розподіли ймовірностей для всіх $a > 0$. Параметр Херста визначає ступінь самоподоби процесу. Чим ближче цей параметр до одиниці, тим більш яскраво проявляються фрактальні властивості. Навпаки, рівність $H=0.5$ свідчить про відсутність самоподоби.

Самоподібні процеси мають гіперболічно згасаючий коефіцієнт кореляції виду

$$r(k) = \frac{1}{2} \left((k+1)^{2H} - 2k^{2H} + (k-1)^{2H} \right), \quad (2)$$

або для асимптотично самоподібних процесів - кореляційну функцію виду

$$R(k) \approx k^{2H-2} L(t) \text{ при } k \rightarrow \infty, \quad (3)$$

де $L(t)$ – повільно мінлива функція на нескінченності для всіх $x > 0$. Отже ряд, що утворений послідовними значеннями кореляційної функції самоподібного процесу, розходиться.

Ця властивість характеризує практично всі самоподібні процеси й відрізняє їх від процесів без довгострокової залежності, у яких кореляційна функція убуває за показовим законом, а ряд, утворений послідовними значеннями цієї кореляційної функції, не розходиться.

Довгострокова залежність є причиною виникнення пульсацій (сплесків) процесу, однак дозволяє говорити про деяку передбачуваність “поведінки” процесу у невеликих межах часу. З точки зору теорії черг важливим наслідком корельованості потоку є неприйнятність оцінок параметрів черги, заснованих на припущенні про однаковий і незалежний розподіл інтервалів у вхідному потоці.

При усередненні самоподібного процесу дисперсія його вибіркового середнього загасає повільніше, ніж величина, що обернена розміру

вибірки, у той час, як для традиційних стаціонарних випадкових процесів має місце таке:

$$\sigma^2(X^{(m)}) = \frac{1}{m} \sigma^2(X), \quad (4)$$

тобто дисперсія зменшується обернено пропорційно об'єму вибірки.

Властивість повільно спадної дисперсії говорить про можливість істотних, не згладжуваних усередненням, «викидів» (або сплесків) у випадковому процесі, і зв'язує поняття самоподоби з таким поняттям, як розподіл з вагомими хвостами. Важливий наслідок властивості повільно загасаючої дисперсії полягає в тому, що у випадку здійснення класичних статистичних тестів (наприклад, обчислення довірчих інтервалів) щодо розподілів з вагомими хвостами не слід користуватися загальноприйнятою мірою середньоквадратичного відхилення σ (оскільки така міра у даному випадку буде некоректною).

З даною властивістю зв'язане й «нехарактерне» поведіння IDC – індексу дисперсії (або індексу розкиду) для відліків процесу, котрий також іноді називають фактором Фано. IDC визначається як відношення дисперсії кількості оброблених пакетів на заданому часовому інтервалі T до математичного сподівання цієї величини:

$$F(T) = \frac{\text{Var}[N(T)]}{E[N(T)]}, \quad (5)$$

де $N(T)$ – кількість пакетів досліджуваного потоку, що були оброблені на інтервалі T .

Для самоподібних процесів логарифм індексу розкиду $F(T)$ лінійно зростає:

$$\ln[F(T) - 1] = (2H - 1) \ln T + y, \quad (6)$$

$$\text{де } y = \ln \left[\frac{2K}{\alpha(1-\alpha)} M_r(\alpha) B^{-\alpha/2} \right],$$

$$M_r(x) = \frac{\Gamma(1+x/2)\Gamma(1-x)}{\Gamma(1-x/2)}. \quad (7)$$

У багатьох випадках фрактальний процес доцільно характеризувати розподілом із, так званим, вагомим хвостом. Випадкова величина Z має розподіл з вагомим хвостом (PBX), якщо ймовірність $P[Z > x] \sim cx^\alpha$ при $x \rightarrow \infty$, тобто хвіст розподілу при зростанні значень x загасає за степеневим законом. Прикладом розподілу з вагомим хвостом є розподіл Парето. Для $0 < \alpha < 2$ величина Z має нескінченну дисперсію, а для $0 < \alpha < 1$ середнє значення такого розподілу також нескінченне.

Найбільш істотною особливістю випадкової величини, для якої є притаманним розподіл з вагомим хвостом, є її надзвичайна мінливість. З імовірністю, що не є малою, у вибірці може бути наявна деяка кількість «дуже великих» значень (тобто, сплесків). Для таких розподілів є характерною невисока точність статистичних оцінок; наприклад, якщо об'єм вибірки для

оцінювання характеристик PBX задавати загальноприйнятими методами, то це призведе до заниженої оцінки його середнього й дисперсії.

Наявність PBX у зовнішніх по відношенню до розглянутих процесів явища є однією із причин виникнення самоподоби у відповідних стохастичних моделях.

Дослідження кореляційної структури пакетного трафіку

Графічне подання залежностей обчислених значень коефіцієнта кореляції від параметра k на усій множині отриманих вибірок досліджуваного трафіку пакетів дозволяє візуально переконатися в тому, що цьому трафіку притаманна властивість довгострокової залежності.

Зокрема, на рис.1 в логарифмічному масштабі наведено графік залежності коефіцієнта кореляції від параметра k для процесу, що відповідає досліджуваному зрізу трафіка. Очевидно, що точки на рис.1 у цілому групуються навколо прямої лінії, кутовий коефіцієнт якої може бути визначений шляхом лінійної регресії. Якщо процес самоподібний, то відповідно до виразу (3) кутовий коефіцієнт має визначатися так: $\beta = 2 - 2H$.

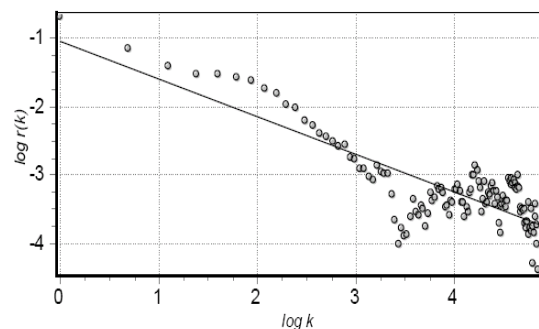


Рис. 1. Експериментально отримана вибірка значень коефіцієнта кореляції досліджуваного трафіка (логарифмічний масштаб)

На виборці, що відображена на рис.1, отримаємо значення $\beta = 0,55004$, а параметр Херста для цих умов виявляється дорівнює 0,724.

На рис. 2, а для порівняння наведена крива, що відповідає значенням коефіцієнта кореляції ідеального самоподібного процесу згідно з виразом (2). Очевидно, що отримані вибіркові оцінки досить точно відповідають «ідеальній» кривій, особливо при збільшенні значень аргументу k .

Значний вплив довгострокової залежності на характеристики досліджуваного трафіку можна виявити й на основі аналізу, так званого, «змішаного» процесу. Такий процес утворюється із вихідного процесу шляхом перестановки у ньому часових інтервалів надходження пакетів. У таким чином отриманому процесі (рис. 2, б) кореляція спадає значно швидше, ніж у вихідному, досить швидко наближаючись до нуля.

Рівень самоподоби у стохастичному процесі зазвичай виявляється шляхом визначення параметра Херста H . Якщо виявлено факт, що $0,5 < H < 1$, тобто значення параметра Херста відмінне від 0,5, то це вважається достатньою підставою для визнання процесу самоподібним (принаймні, асимптотично). Але слід мати на увазі, що значення H , яке є близьким до одиниці, може означати також, що йдеться про детермінований, а не випадковий процес, оскільки для ряду детермінованих процесів характерна структура, що регулярно повторюється на будь-якому масштабі (це приводить до одиничного значення параметра Херста).

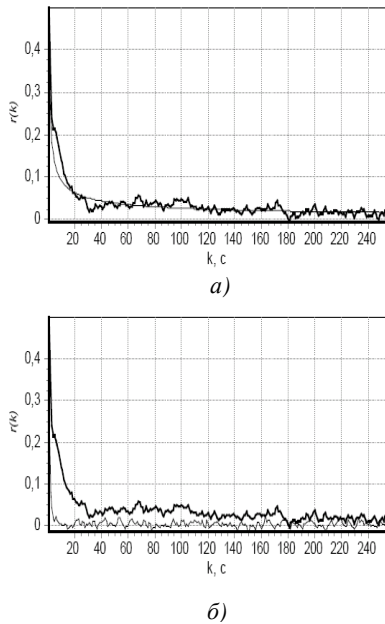


Рис. 2. Графіки залежності коефіцієнта кореляції від інтервалу кореляції у досліджуваному процесі (поток пакетів) у порівнянні з кривою коефіцієнта кореляції: а – ідеального самоподібного процесу при $\beta=0,55$; б – процесу зі «змішаними» випадковим чином інтервалами кореляції

Зокрема, на рис. 3 наведені значення $\ln[F(T)-1]$ залежно від $\ln T$. Відповідно до виразу (7), точки утворюють пряму лінію, що дозволяє однозначно визначити параметр Херста. Як бачимо, пряма, що отримана шляхом лінійної регресії, має кутовий коефіцієнт 0.459. Це означає, що $H=0,729$.

Таким чином, констатуємо: для розглянутої вибірки досліджуваного трафіку значення параметра Херста, що отримані виходячи з вигляду кривої коефіцієнта кореляції (рис. 3) та шляхом аналізу IDC (0.724 й 0.729 відповідно), практично збігаються і є такими, що дозволяє говорити про істотно виражені фрактальні властивості досліджуваного трафіку.

Для процесу зі «змішаними» випадковим чином інтервалами кореляції значення параметра Херста, отримане шляхом аналізу IDC, виявляється істотно меншим: $H=0,563$. Цей факт, а також результати аналізу вигляду кореляційної структури вихідного й

«змішаного» процесів дозволяє стверджувати, що для розглянутих вибірок пакетного трафіку властивість самоподоби пов'язана не стільки з «вагомістю» розподілу часових інтервалів, скільки з довгостроковою залежністю – угрупованням коротких інтервалів у пачки, тобто «пачковістю».

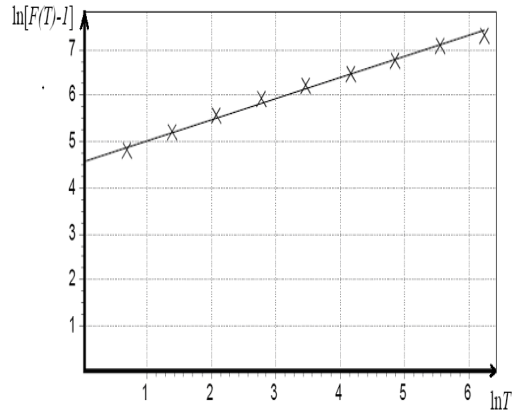


Рис. 3. Ілюстрація визначення параметра Херста H за значеннями IDC

Дослідження можливостей прогнозування самоподібного трафіка

Пакетний трафік мереж передачі даних, зведений до еквідистантної форми за координатою часу (наприклад, за допомогою процедури агрегування), являє собою певний часовий ряд (тобто, реалізацію дискретного випадкового процесу), миттєві значення якого можна намагатися прогнозувати за допомогою наявних методик прогнозування.

Числовою характеристикою випадкового процесу, що на пряму пов'язана зі ступенем його передбачуваності, є *інтервал кореляції* τ_k , обумовлений виразом

$$\tau_k = \int_0^{\infty} |r(\tau)| d\tau, \quad (8)$$

де $r(\tau)$ – нормована функція кореляції випадкового процесу.

Якщо відома «поведінка» реалізації процесу в минулому, то є можливим імовірнісний прогноз випадкового процесу на проміжок часу, тривалістю τ_k [2]. Однак прогнозування «поведінки» процесу на більший час виявляється безрезультатним, оскільки далеко віддалені один від одного відліки процесу є статистично незв'язаними (некорельованими). Щоб усвідомити фізичний зміст поняття інтервалу кореляції, досить глянути на рис. 4.

Аналізуючи його, можна помітити, що інтервал кореляції чисельно дорівнює площі під графіком нормованої кореляційної функції випадкового процесу. Згадаємо, що всі визначення самоподібного процесу даються за умов гіперболічно спадаючої кореляційної функції:

$$r(\tau) \sim \tau^{-\beta}, \tau \rightarrow \infty, 0 \leq \beta \leq 1. \quad (9)$$

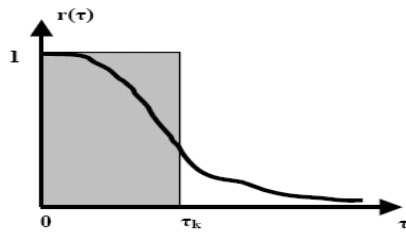


Рис. 4. До визначення інтервалу кореляції (площа під кривою дорівнює площі заштрихованого прямокутника)

Підставивши (9) у (8), одержимо

$$\tau_k = \int_0^{+\infty} |\tau^{-\beta}| d\tau.$$

Розв'язуючи останній вираз в загальному випадку для будь-якого β , отримаємо:

Якщо $\beta \neq 1$, то для будь-якого $R > 0$

$$\lim_{R \rightarrow +\infty} \int_0^R \frac{d\tau}{\tau^\beta} = \lim_{R \rightarrow +\infty} \left. \frac{\tau^{-\beta}}{1-\beta} \right|_0^R = \lim_{R \rightarrow +\infty} \frac{R^{-\beta}}{1-\beta} = \begin{cases} 0, & \text{при } \beta > 1; \\ \infty, & \text{при } \beta < 1; \end{cases}$$

Якщо $\beta = 1$, то для будь-якого $R > 0$

$$\lim_{R \rightarrow +\infty} \int_0^R \frac{d\tau}{\tau} = \lim_{R \rightarrow +\infty} \ln \tau \Big|_0^R = \infty.$$

Таким чином, для випадку самоподібного процесу, що характеризується $0 < \beta < 1$, маємо нескінченно тривалий інтервал кореляції:

$$\tau_k \rightarrow \infty. \quad (10)$$

Вищезазначене забезпечує собою теоретично принципову можливість прогнозування самоподібного процесу. Звичайно, у реальній ситуації (при дослідженні реального самоподібного мережного трафіка) складно вимагати жорсткого виконання умов згідно з виразом (9), однак у більшості випадків щодо пакетного трафіка можна чекати, що τ_k приймає досить великі значення.

Звісно, що з поняттям самоподібний мережний трафік тісно зв'язані такі поняття, як *персистентність* і *тривала пам'ять*, які відбивають ступінь статистичного зв'язку між послідовними відліками процесу за координатою часу, який моделює цей трафік.

Персистентні стохастичні процеси виявляють чітко виражені тенденції до збереження напрямку змін при відносно малому шумі. Наприклад, якщо протягом деякого часу в минулому спостерігалися позитивні прирости процесу, тобто відбувалося збільшення, то й надалі в середньому буде відбуватися збільшення. Інакше кажучи, імовірність того, що процес на $i+1$ часовому інтервалі відхиляється від середнього в тім же напрямку, що й на i інтервалі, настільки велика, наскільки значення параметра Херста H наближено до 1.

Ще один аспект на користь задовільної прогнозованості мережного трафіка – наявність у

його розподілі вагомих хвостів. Вище було показано: вагомні хвости розподілу випадкового процесу можуть бути причиною його задовільної передбачуваності.

Результати аналізу досліджуваного пакетного трафіка показують, що поряд з постійною складовою процесу, що моделює цей трафік (тобто, математичним сподіванням процесу), та стохастичною компонентою цього процесу чітко проявляються також регулярні (детерміновані) компоненти, котрі відповідають гармонійним коливанням з певними частотами.

Показники якості прогнозування

Останнім часом у задачах керування обладнанням телекомунікаційних мереж дедалі ширше використовуються адаптивні методи настроювання ширини смуги портів комутаторів на прогнозні значення оцінок інтенсивності потоків протокольних блоків даних, що просуваються через ці порти у реальному часі [3]. Враховуючи високі динамічні характеристики пульсацій трафіку, для отримання їх прогнозних оцінок бажано застосовувати відповідні алгоритми прогнозування. Один із таких алгоритмів, що базується на звісній процедурі прогнозування Бокса-Дженкінса, розглянуто у цій статті.

Задачу прогнозування пульсацій потоків пакетів в багатьох випадках доцільно звести до задачі прогнозування "поведінки" дискретного часового ряду. З погляду загальної теорії прогнозування [4] часовий ряд, що моделює трафік потоку пакетів, можна подати у вигляді

$$x(k) = \chi(k) + \varepsilon(k), \quad (11)$$

де $\chi(k)$ – не випадковий компонент (що включає в себе в загальному випадку тренд, циклічні й сезонні коливання); $\varepsilon(k)$ – випадкові (нерегулярні) коливання потоку пакетів.

Тоді прогноз n -го значення тимчасового ряду $x(k)$ складається із прогностичної оцінки випадкового й не випадкового компонентів:

$$\hat{x}(n) = \hat{\chi}(n) + \hat{\varepsilon}(n). \quad (12)$$

Зрозуміло, що якість прогнозу ряду $x(k)$ тим вища, чим якісніші оцінки його компонентів.

У цій роботі задача прогнозування мережного трафіка на портах пакетного комутатора розглядається як складова задачі динамічного керування смугами портів пакетного комутатора. Тому бажано не тільки зробити висновки щодо прогностичності мережного трафіка, але й оцінити виграш від застосування адаптивного алгоритму динамічного керування смугами портів комутатора у порівнянні із класичною схемою статичного розподілу пропускнуої здатності комутатора між його портами. Це важливо з точки зору оцінки доцільності застосування адаптивних методів

керування на практиці. У цьому зв'язку оцінку $\hat{C}(n)$ пропускної здатності комутатора, для якого призначений трафік $x(k)$, на n -му кроці прогнозування будемо розраховувати як мінімальну суму ширин смуг його портів, що перевищує оцінку трафіка на n -му кроці $\hat{x}(n)$ (тобто, на деякому кроці можуть бути задіяні не всі порти комутатора). На практиці зазвичай прагнуть зменшити абсолютну помилку

$$e(k) = x(k) - C(k), \quad (13)$$

що фактично є нев'язкою між реальним значенням інтенсивності трафіка і розрахованими значеннями прогнозу щодо цього трафіка.

У якості параметра, що оцінює якість алгоритму прогнозування, часто користуються так званим відношенням сигнал/шум:

$$SN = \frac{M[x(k)^2]}{M[e(k)^2]} = \frac{\sum x(k)^2}{\sum e(k)^2}, \quad (14)$$

або різними його модифікаціями [5; 6] (зокрема, оберненим відношенням цього параметра, що вимірюється у децибелах й ін.). Однак така оцінка якості прогнозування залежить від величини середнього значення ряду $x(k)$, що не дозволяє порівнювати якість прогнозу рядів з різним середнім.

Тому на практиці зазвичай використовується модифікована оцінка [7]

$$SNR^{-1} = \frac{M[e(k)^2]}{M[x(k) - M[x(k)]]^2}, \quad (15)$$

яка не залежить від середнього значення прогнозованої ділянки ряду й відбиває ступінь поліпшення прогнозу трафіка у порівнянні із прогнозуванням за середнім значенням. Іншими словами, оцінка $SNR^{-1} = 1$, якщо якість прогнозу така ж, як і у випадку прогнозу за середнім значенням прогнозованої ділянки ряду. При $SNR^{-1} < 1$ якість прогнозу покращується. Завдяки такій властивості, оцінка SNR^{-1} дозволяє порівнювати якість прогнозу різних часових рядів між собою.

Крім того, поряд з абсолютною погрішністю прогнозу $e(k)$ будемо розглядати окремо помилки недооцінки ряду

$$e^+(k) = \begin{cases} e(k), & \text{якщо } e(k) \geq 0 \\ 0, & \text{якщо } e(k) < 0 \end{cases}, \quad (16)$$

які в задачі адаптивного керування пропускними здатностями портів комутатора пов'язані з кількістю загубленої інформації, і помилки переоцінки ряду

$$e^-(k) = \begin{cases} |e(k)|, & \text{якщо } e(k) < 0 \\ 0, & \text{якщо } e(k) \geq 0 \end{cases}, \quad (17)$$

які, у свою чергу, відображають кількість недовикористаних ресурсів комутатора.

У якості оцінки якості прогнозу будемо також розглядати коефіцієнт недооцінки (коефіцієнт втрат)

$$D^+ = \frac{M[e^+(k)]}{M[x(k)]} = \frac{\sum e^+(k)}{x(k)}, \quad (18)$$

і коефіцієнт переоцінки (коефіцієнт недовикористання)

$$D^- = \frac{M[e^-(k)]}{M[x(k)]} = \frac{\sum e^-(k)}{x(k)}. \quad (19)$$

Ці коефіцієнти, хоч і залежать від середнього значення прогнозованої ділянки ряду, однак мають чітку фізичну інтерпретацію: при заданому способі оцінки пропускної здатності порту коефіцієнт недооцінки D^+ виражає відношення кількості загубленої інформації до загальної кількості інформації, яку потрібно було обробити. Відповідно, коефіцієнт переоцінки D^- відбиває кількість невикористаної пропускної здатності комутатора відносно загальної кількості інформації, яку можливо було пропустити через його порти. Очевидно, чим ближче прогностичні оцінки до дійсних значень $x(k)$, тим ближче до нуля коефіцієнти D^+ й D^- .

Окрім наведених вище коефіцієнтів оцінювання якості прогнозу, введемо певні коефіцієнти оцінювання ширин смуг портів комутатора за прогнозними значеннями, які спрямовані на підвищення ефективності функціонування механізмів керування смугами портів цього комутатора. Саме ці коефіцієнти будуть оцінювати якість обраного алгоритму прогнозування (прогностичний механізм може доволі часто помилятися, але якщо при цьому вибрано правильно ширину смуги порту, то помилок в роботі адаптивного механізму керування шириною смуг портів не буде).

Коефіцієнт економії:

$$Ec = \frac{k \cdot \sum_{0 < i < n} W_i - \sum_{i=0}^k Wp_i}{k \cdot \sum_{0 < i < n} W} \cdot 100\%, \quad (20)$$

де k – номер точки прогнозу; n – загальна кількість каналів; W_i – ширина смуги i -го порту комутатора; Wp_i – ширина смуги за прогнозом. Коефіцієнт Ec відображає частину трафіка, що економиться завдяки застосуванню того чи іншого алгоритму прогнозування.

Коефіцієнт помилки:

$$Er = \frac{k_{error}}{k} \cdot 100\%, \quad (21)$$

де k – номер точки прогнозу; k_{error} – кількість точок, в яких прогнозована ширина смуги порту була недостатньою. Коефіцієнт Er відображає частину помилок, що виникають під час застосування того чи іншого алгоритму прогнозування.

Експеримент із дослідження можливостей механізму динамічного (тобто, на підставі прогнозу)

керування пропускною здатністю пакетного комутатора (точніше, керування величинами смуг пропускання його портів) проведемо із використанням такого алгоритму.

1) У вихідному дискретному часовому ряді $x(k)$, який відповідає досліджуваному пакетному трафіку, що просувається через порт комутатора, виділимо так звану тренувальну ділянку T фіксованої довжини LT .

2) Вивчаючи характеристики ряду на даній тренувальній ділянці (зокрема, оцінюючи параметри прогностичної моделі), формуємо прогноз $P(i)$ (на один крок уперед) i -го значення ряду $x(k)$, що впливає за кінцем тренувальної ділянки, а також розраховуємо величину смуги порту комутатора і коефіцієнти похибки та економії.

3) Фіксуємо абсолютну помилку прогнозу $e(i)$, яка мала місце.

4) Зрушуємо тренувальну ділянку T довжини LT на один крок уперед, маючи на увазі, що до цього моменту часу вже стало відомим дійсне значення прогнозованого відліку i . Здійснюємо прогноз наступного значення і т.д.

Процедура прогнозування за Боксом-Дженкінсом

1) Спочатку процедури процес типу АРПСС (p, d, q) розглядається як чисто авторегресійний процес АР (p, d). Для оцінки параметрів процесу авторегресії (АР) використаємо систему лінійних рівнянь, які часто називають рівняннями Юла-Уокера:

$$\begin{cases} \rho_1 = \Phi_1 + \Phi_2\rho_1 + \dots + \Phi_p\rho_p - 1 \\ \rho_2 = \Phi_1\rho_2 + \Phi_2 + \dots + \Phi_p\rho_p - 2 \\ \dots \\ \rho_p = \Phi_1\rho_{p-1} + \Phi_2\rho_{p-2} + \dots + \Phi_p \end{cases}$$

де Φ – оцінюваний параметр; ρ_j – автокореляція j -го порядку;

$$\rho_k = c_k / c_0, \quad c_k = \frac{1}{N} \sum_{t=1}^{N+k} (z_t - \hat{z})(z_{t+k} - \hat{z})$$

де \hat{z} – середнє значення часового ряду.

2) З метою уточнення коефіцієнтів процесу ковзного середнього (КС) процес АРПСС(p, d, q), що моделюється, розглядається як процес ковзного середнього КС(d, q). Значення функції автоковаріації процесу КС(d, q) уточнюється з урахуванням обчислених параметрів авторегресії:

$$C_j = \begin{cases} \sum_{i=0}^p \sum_{k=0}^p \Phi_{i0} \Phi_{k0} C_{|j+i-k|}, & p > 0 (\Phi_{00} = -1), \\ C_j, & p = 0 \end{cases}$$

де C_j – автокореляція j -го порядку вихідного ряду.

3) Обчислюються параметри процесу ковзного середнього й дисперсії за допомогою лінійно збіжного ітераційного процесу:

$$\sigma_a^2 = \frac{C_0'}{1 + \Theta_1^2 + \dots + \Theta_q^2};$$

$$\Theta_j = - \left(\frac{C_j}{\sigma_a^2} - \Theta_1 \Theta_{j+1} - \Theta_2 \Theta_{j+2} - \dots - \Theta_{q-j} \Theta_q \right).$$

4) Після завершення ітераційного процесу обчислюється оцінка загальної константи $\hat{\Theta}_{00}$ й дисперсії білого шуму $\hat{\sigma}_a^2$:

$$\hat{\Theta}_{00} = \begin{cases} \hat{Z} \left(1 - \sum_{i=1}^p \hat{\Phi}_{i0} \right), & p > 0 \\ \hat{Z}, & p = 0 \end{cases}$$

$$\hat{\sigma}_a^2 = \begin{cases} \tau_0^2, & q > 0 \\ C_0 - \sum \Phi_i C_i, & q = 0 \end{cases}$$

5) Здійснюється прогнозування.

Роз'єднання операторів. Для цілей прогнозу сезонна (у загальному випадку) модель представляється у вигляді

$$z_t = \Phi I^* z_{t-1} + \dots + \Phi p^* z_{t-p^*} + \Theta 00 + a_t - \Theta I^* a_{t-1} - \dots - \Theta Q^* a_{t-Q^*},$$

де невідомі параметри замінені їхніми оцінками. Для одержання Φ_i^* спочатку знаходять параметри $\Phi I^*, \dots, \Phi p^*$ оператора $\Phi(B) = \Phi(Bs) \phi(B) = 1 - \Phi I^* B - \dots - \Phi p^* B^{p^*}$, де $p^* = p + s$.

Для цього використовується формула:

$$\Phi^k = - \sum_i \sum_j \Phi_i \Phi_j, \quad j + si = k.$$

Подвійна сума є сумою за всіма $j = j = 0, \dots, p_j$ і всіма $i = 0, \dots, P$. Константи ϕ і Φ приймаються рівними -1. Далі, параметри оператора $\Phi^*(B) = \delta \Delta^D \Phi(B) = 1 - \Phi I^* B - \dots - \Phi p^* B^{p^*}$ з $p^* = p' + d + s$ виходять у результаті наступної триетапної процедури:

Етап 1. Обчислюються

$$\Phi''_j = \begin{cases} \Phi'_j, & 0 \leq j \leq s \\ \Phi'_j - \Phi'_{j-s'}, & s \leq j \leq p' \\ -\Phi'_{j-s}, & p' < j \leq p' + s \end{cases}$$

Етап 2. Етап 1 повторюється D разів із заміною Φ'_j на Φ''_j , p' на $p'' = p' + s$ на кожній ітерації для одержання коефіцієнтів $\Delta^D_s \Phi(B)$.

Етап 3. Аналогічно застосовуємо d разів етап 1 з $s=1$ до значень, отриманих на етапі 1, для знаходження остаточних оцінок коефіцієнтів $\Phi_1^*, \dots, \Phi_{p^*}^*$.

Висновки

Оператори ковзного середнього визначають тим же способом, як й Φ , за формулою

$$\Theta_k^* = -\sum_i \sum_j \Theta_i \theta_j, \quad j + ik = k,$$

де $i = 0, 1, \dots, Q$; $j = 0, 1, \dots, q$; $\Theta_0 = -1$.

Для трансформованого ряду z'_i прогнози

$\hat{z}_{N-b}^{(L)}$ обчислюються за формулою

$$\hat{z}_{N-b}^{(L)} = \Theta_0;$$

$$\sum_{i=1}^{p+sP+d+sD} \Phi[Z'_{N-b-i+L}] - \sum_j \Theta_j^* [a_{N-b-j+L}],$$

$$\text{де } [Z'_{N-b-i+L}] = \begin{cases} Z'_{N-b}(L-i) & , L > i \\ Z'_{N-b-i+L} & , L \leq i \end{cases}$$

$$[a_{N-b-j+L}] = \begin{cases} 0 & , L > j \\ Z'_{N-b-i+L} - Z'_{N-b-i+L-1}(1) & , L \leq j \end{cases}$$

$L = 1, 2, \dots, L$.

Прогнози знаходять для кожного моменту часу $b = 0, 1, \dots, B$. Для обчислення $[a_{N-b-j+L}]$ починаємо процес прогнозування з апроксимації $a_t < 0$, нулями.

Запропоновано алгоритм прогнозування пульсацій пакетного трафіка на портах пакетного комутатора, що базується на використанні процедури Бокса-Дженкінса. Передбачається, що основна галузь використання цього алгоритму пов'язана із розробкою адаптивних механізмів керування смугами пропускання портів пакетного комутатора в задачах інженерії трафіку. Перевагою даного алгоритму є те, що результати прогнозу, які отримані за даним алгоритмом, базуються на достатньо великій кількості даних, значення котрих тим більше впливає на результат, чим ближче ці дані до моменту прогнозування. Саме такі умови є характерними під час динамічного керування пропускнуою здатністю комутатора. Крім того, на кожному кроці прогнозу є можливість не проводити повний перерахунок усіх коефіцієнтів, оскільки їх значення накопичуються на кожному кроці.

Список літератури

1. Городецкий, А. Я. Информатика. Фрактальные процессы в компьютерных сетях / А. Городецкий, В. Забровский. – СПб : Изд-во СПб ГТУ, 2000. – 96 с.
2. Баскаков, С. И. Радиотехнические цепи и сигналы: Учеб. для вызов по спец. Радиотехника. 2-е изд. перераб. и доп. – М. : Высш. шк., 1998. 448 с.
3. Кочергин, Ю. А. Задача авторегулирования перераспределением пропускной способности пакетного коммутатора между его портами // Математичні машини і системи. – Вип. 2. – 2006.
4. Айвазян, С. А. Прикладная статистика. Основы эконометрики: Учебник для вузов: В 2 т. 2-е изд. испр. – Т. 2. Айвазян С.А. Основы эконометрики. – М. : ЮНИТИ-ДАНА, 2001. – 432 с.
5. Ghaderi, M. On the Relevance of Self-Similarity in Network Traffic Prediction, 2003. <http://www.cs.uwaterloo.ca/cs-archive/CS-2003/28/TR-CS-2003>. – 28. pdf.
6. Sadec, N., Khotanzad, A., Chen, T. ATM Dynamic Bandwidth Allocation Using FARIMA Prediction Model. <http://snoopy.seas.smu.edu/papers/icccn03>. pdf.
7. Експлуатація телекомунікаційних систем [текст] : підручник. / Г. Ф. Конахович, В. М. Чуприн, І. О. Мачалін, О. П. Ткаліч. – К. : Центр учбової літератури, 2014. – 372 с.
8. Конахович, Г. Ф., Чуприн, В. М.. Сети передачи пакетных данных. – К. : «МК-Пресс», 2006. – 272 с.
9. ITU-T Recommendation M.3010(02/00). Principles for a Telecommunications management network (Принципи управління телекомунікаційними мережами).

References

1. Horodetskyi, A. Ya., Zabrovskyi, V. (2000). Informatics. Fractal processes in computer networks, St.Petersburg, Russia, SPb HTU, 96.
2. Buskacov, S. I. (1998). Radiotechnic signals, Moscow, Russia, H.Schule, 448.
3. Kocherhin, Yu. A. (2006) Autoregulation task redistribution capacity batch commutator between its ports, Mathematical Machines and Systems, Kyiv, Ukraine: IC ASU, 2, 60-70.
4. Ivazyan, S. A., (2001). Statistics. Moscow, Russia, UNITY-DANA, 432.
5. Ghaderi, M. On the Relevance of Self-Similarity in Network Traffic Prediction, 2003. <http://www.cs.uwaterloo.ca/cs-archive/CS-2003/28/TR-CS-2003>. – 28. pdf.
6. Sadec, N., Khotanzad, A., & Chen, T. ATM Dynamic Bandwidth Allocation Using FARIMA Prediction Model. <http://snoopy.seas.smu.edu/papers/icccn03>. pdf.
7. Konakhovych, H. F., Chupryn, V. M., Machalin, I. O. & Tkalic O. P. (2014). The operation of telecommunications systems, Kyiv, Ukraine: Center educational literature, 372.
8. Konahovich, G. F., & Chupryn, V. M. (2006). Transport Networks of package data. –Kiev, Ukraine, МК-Пресс», 272.
9. ITU-T Recommendation M.3010(02/00). Principles for a Telecommunications management.

Стаття надійшла до редколегії 20.10.2014

Рецензент: д-р техн. наук, проф. Г.Ф. Конахович, Національний авіаційний університет, Київ.