

УДК 629.735.051

Муранов О.С.

АЛГОРИТМ ПРОГНОЗУВАННЯ ПУЛЬСАЦІЙ ПАКЕТНОГО ТРАФІКА ЗА ПРОЦЕДУРОЮ БОКСА-ДЖЕНКІНСА

Інститут інформаційно-діагностичних систем
Національного авіаційного університету

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

Останнім часом у задачах керування обладнанням телекомунікаційних мереж дедалі ширше використовуються адаптивні методи настроювання ширини смуги портів комутаторів на прогнозі значення оцінок інтенсивності потоків протокольних блоків даних, що просуваються через ці порти у реальному часі [1]. 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.

Показники якості прогнозування

Задачу прогнозування пульсацій потоків пакетів в багатьох випадках доцільно звести до задачі прогнозування "поведінки" дискретного часового ряду. З погляду загальної теорії прогнозування [2] часовий ряд, що моделює трафік потоку пакетів, можна представити у вигляді

$$x(k) = \chi(k) + \varepsilon(k), \quad (1)$$

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:

$$\hat{x}(n) = \hat{\chi}(n) + \hat{\varepsilon}(n). \quad (2)$$

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

У цій роботі задача прогнозування мережного трафіка на портах пакетного комутатора розглядається як складова задача 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 $\hat{C}(n)$ пропускнуої здатності комутатора, для якого призначений трафік $x(k)$, на n -ому кроці прогнозування будемо розраховувати як мінімальну суму ширин смуг його портів, що перевищує оцінку трафіка на n -ому кроці $\hat{x}(n)$ (тобто, на деякому кроці можуть бути задіяні не всі порти комутатора). На практиці звичайно прагнуть зменшити абсолютну помилку

$$e(k) = x(k) - C(k), \quad (3)$$

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 званім відношенням сигнал/шум:

$$SN = \frac{M[x(k)^2]}{M[e(k)^2]} = \frac{\sum x(k)^2}{\sum e(k)^2}, \quad (4)$$

або різними його модифікаціями [3, 4] (зокрема, зворотним 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 [5]

$$SNR^{-1} = \frac{M[e(k)^2]}{M[x(k) - M[x(k)]]^2}, \quad (5)$$

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

$$e^+(k) = \begin{cases} e(k), & \text{якщо } e(k) \geq 0 \\ 0, & \text{якщо } e(k) < 0 \end{cases}, \quad (6)$$

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

$$e^-(k) = \begin{cases} |e(k)|, & \text{якщо } e(k) < 0 \\ 0, & \text{якщо } e(k) \geq 0 \end{cases}, \quad (7)$$

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

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

$$D^+ = \frac{M[e^+(k)]}{M[x(k)]} = \frac{\sum e^+(k)}{x(k)}, \quad (8)$$

і коефіцієнт переоцінки (коефіцієнт недовикористання)

$$D^- = \frac{M[e^-(k)]}{M[x(k)]} = \frac{\sum e^-(k)}{x(k)}. \quad (9)$$

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, уведемо 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:

$$Ec = \frac{k \cdot \sum_{0 < i < n} W_i - \sum_{i=0}^k Wp_i}{k \cdot \sum_{0 < i < n} W} \cdot 100\%, \quad (10)$$

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 долю трафіка, що економиться завдяки застосуванню того чи іншого алгоритму прогнозування.

Коефіцієнт помилки:

$$Er = \frac{k_{error}}{k} \cdot 100\%, \quad (11)$$

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

Ядро обчислювальної процедури

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 наступного алгоритму.

1) У вихідному дискретному часовому ряді $x(k)$, котрий 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 i т.д.

Тренувальна ділянка для прогнозу пульсацій пакетного трафіку показана на рис. 1, а блок-схема базового алгоритму прогнозування – на рис. 2.

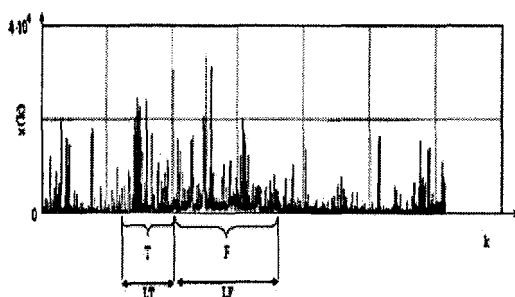


Рис. 1. Тренувальна ділянка для прогнозу пульсацій трафіка на портах пакетного комутатора

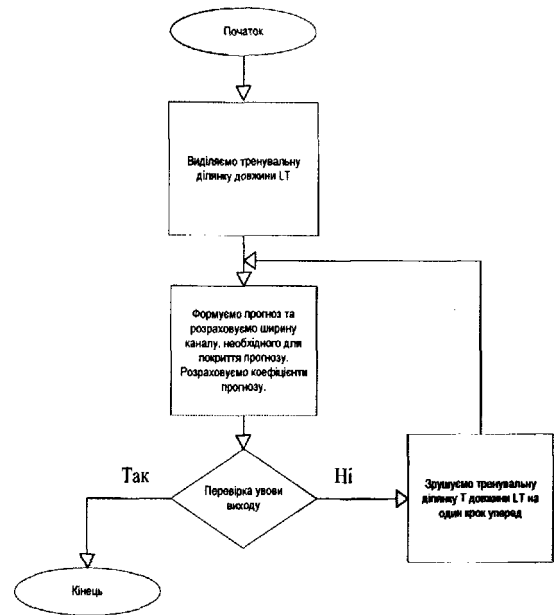


Рис. 2. Блок-схема базового алгоритму прогнозування трафіка

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 $\hat{\Theta}_{\infty}$ 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

$$zt = \Phi 1 * zt - 1 + \dots + \Phi p ** zt - p * + \Theta 0 0 + at - \Theta 1 * at - 1 - \dots - \Theta Q * at - Q *$$

де невідомі параметри замінені їхніми оцінками. Для одержання Φ_i^* спочатку знаходять параметри $\Phi 1', \dots, \Phi p'$ оператора

$$\Phi'(B) = \Phi(Bs) \phi(B) = 1 - \Phi 1' B - \dots - \Phi p' B^p,$$

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 $j = 0, \dots, p'$ і всіма $i = 0, \dots, p$. Константи

ϕ і Φ приймаються рівними -1. Далі, параметри оператора

$$\Phi^*(B) = \delta \Delta^D \Phi'(B) = 1 - \Phi 1^* B - \dots - \Phi p^* B^{p^*}$$

з $p^* = p' + d + s$ виходять у результаті наступної трьох етапної 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 раз 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 раз етап 1 з $s=1$ до значень, отриманих на етапі 1, для знаходження остаточних оцінок коефіцієнтів $\Phi_1^*, \dots, \Phi_{p^*}^*$.

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

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

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

Генерування прогнозів

Для трансформованого ряду z'_t прогнози $\hat{z}_{N-b}^{(L)}$ обчислюються за формулою

$$\hat{z}_{N-b}^{(L)} = \Theta 0;$$

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

де

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

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

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

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 , $t < 0$, нулями.

Висновки

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. *Кочергін Ю.А.* Дослідження шляхів підвищення завантаженості обладнання авіаційних телекомунікаційних мереж на основі використання механізмів адаптивного керування потоками даних. Дисертаційна робота на здобуття наукового ступеню кандидата технічних наук. К.: НАУ, 2005. – 145 с.

2. *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.

3. *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>.

4. *Sadec N., Khotanzad A., Chen T.* ATM Dynamic Bandwidth Allocation Using FARIMA Prediction Model. <http://snoopy.seas.smu.edu/papers/icccn03.pdf>.

5. *Петров В.В.* Структура телетрафика и алгоритм обеспечения качества обслуживания при влиянии эффекта самоподобия. – М.: 2004. – 101 с.