

МОДЕЛІ ТА МЕТОДИ ПРОГНОЗУВАННЯ МЕРЕЖЕВОГО ТРАФІКУ В РЕАЛЬНОМУ ЧАСІ**Н. Л. Сохін, М. І. Гученко, А. О. Кирса**

Кременчуцький національний університет імені Михайла Остроградського

вул. Першотравнева, 20, м. Кременчук, 39600, Україна. E-mail: work2301@gmail.com

Проаналізовано причини виникнення перевантажень в комп'ютерних мережах та методи запобігання перевантажень, показано їх зв'язок з керуванням трафіком і, зокрема, забезпеченням QoS. Обґрунтовано необхідність застосування прогнозування трафіка для покращення роботи методів запобігання перевантажень. Виконано аналіз моделей та методів, які можуть застосовуватися для прогнозування трафіка комп'ютерних мереж. Наведено переваги та недоліки використання методів аналізу часових рядів та штучних нейронних мереж в задачі прогнозування мережевого трафіка. Наведено методіку використання алгоритму побудови локальної моделі керованого процесу для отримання прогнозу мережевого трафіку. Показано, що дана методика дає змогу прогнозувати мережевий трафік в реальному часі, без необхідності попереднього накопичення, обробки та аналізу великої кількості статистичних даних, не вимагає здійснення вибору моделей та оцінки їх придатності. Отримані результати можуть бути використані при вирішенні практичних завдань, що виникають при дослідженні функціонування різних типів мереж передачі даних, а в перспективі – поширені на вирішення як задач боротьби з перевантаженнями, так і задач виявлення аномалій і загроз в комп'ютерних мережах.

Ключові слова: прогнозування в реальному часі, мережевий трафік, QoS, авторегресійні моделі, нейронні мережі, локальна модель керованого процесу.

МОДЕЛИ И МЕТОДЫ ПРОГНОЗИРОВАНИЯ СЕТЕВОГО ТРАФИКА В РЕАЛЬНОМ ВРЕМЕНИ**Н. Л. Сохин, Н. И. Гученко, А. О. Кирса**

Кременчугский национальный университет имени Михаила Остроградского

ул. Первомайская, 20, 39600, г. Кременчуг, Украина. E-mail: work2301@gmail.com

Проанализированы причины возникновения перегрузок в компьютерных сетях и методы предотвращения перегрузок, показана их связь с управлением трафиком и, в частности, с обеспечением QoS. Обоснована необходимость применения прогнозирования трафика для улучшения работы методов предотвращения перегрузок. Выполнен анализ моделей и методов, которые могут применяться для прогнозирования трафика компьютерных сетей. Приведены преимущества и недостатки использования методов анализа временных рядов и искусственных нейронных сетей в задаче прогнозирования сетевого трафика. Приведена методика использования алгоритма построения локальной модели управляемого процесса для получения прогноза сетевого трафика. Показано, что данная методика позволяет прогнозировать сетевой трафик в реальном времени, без необходимости предварительного накопления, обработки и анализа большого количества статистических данных, не требует осуществления выбора моделей и оценки их пригодности. Полученные результаты могут быть использованы при решении практических задач, возникающих при исследовании функционирования различных типов сетей передачи данных, а в перспективе – распространены на решение как задач борьбы с перегрузками, так и задач обнаружения аномалий и угроз в компьютерных сетях.

Ключевые слова: прогнозирование в реальном времени, сетевой трафик, QoS, авторегрессионные модели, нейронные сети, локальная модель управляемого процесса.

АКТУАЛЬНІСТЬ РОБОТИ. В сучасних мульти-сервісних мережах задіяна велика кількість різних сервісів і додатків, що працюють за різними протоколами, по-різному використовують ресурси мережі та мають різні характеристики (затримку, інтенсивність, розмір пакетів тощо). Тому в мережевій інфраструктурі присутні водночас кілька типів трафіку, що робить особливо актуальною задачу розподілу пропускну здатності мережі [1]. Її вирішення можливе за рахунок розроблення та реалізації інтелектуальних методів організації керування трафіком.

Методи керування трафіком є одними з ключових при забезпеченні якості обслуговування (QoS – Quality of Service). Важливою частиною задачі забезпечення QoS є боротьба з перевантаженнями в комп'ютерній мережі (КМ). Ефективні алгоритми боротьби з перевантаженнями дозволяють підвищити не тільки надійність, але і корисну пропускну здатність мережі [2]. Тому задача дослідження моделей та методів прогнозування трафіка, що можуть бути застосовані в методах боротьби з перевантаженнями, є актуальною.

Метою роботи є дослідження моделей та методів прогнозування мережевого трафіка з точки зору їх придатності для отримання прогнозу трафіку в реальному часі.

МАТЕРІАЛ І РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ. Боротьба з перевантаженнями в КМ є важливою частиною задачі забезпечення QoS і керування трафіком зокрема [3]. Використання ефективних алгоритмів боротьби з перевантаженнями дає можливість підвищити як надійність, так і корисну пропускну здатність мережі.

При перевантаженні продуктивність мережі (число оброблених пакетів) прагне до нуля, а час затримки – до нескінченності. Як правило, перевантаження може викликатися флуктуаціями потоків трафіку або виходом з ладу будь-якого елемента мережі. Така ситуація може призвести як до недотримання зобов'язань мережі щодо забезпечення якості обслуговування існуючих з'єднань, так і до неможливості встановлення нового з'єднання з потрібною якістю обслуговування. У роботі [4] зазнача-

ється, що в умовах перевантаження продуктивність мережі може бути меншою, ніж продуктивність мережі, в якій діють механізми динамічного керування доступом до ресурсів і контролю за їх параметрами.

Керування потоками даних передбачає використання системних методів і алгоритмів керування трафіком (TE, traffic engineering), пов'язаних з оптимізацією робочих характеристик мереж [5]. TE включає набір взаємопов'язаних мережевих елементів, систему моніторингу стану мережі та набір засобів керування конфігурацією, що дозволяє превентивно, використовуючи прогнозування стану і тенденцій розвитку трафіку, виконувати дії для запобігання небажаним майбутнім станам.

Основною функцією, яку виконує TE, є ефективне керування пропускнуою здатністю. В даний час в телекомунікаційних мережах використовуються різні методи TE. Більшість з них передбачає можливість зовнішньої параметризації, тобто передачі параметрів трафіку безпосередньо використовуваним алгоритмам керування. Деякі з методів, як, наприклад, метод мультипротокольної комутації пакетів за мітками (MPLS), що дозволяє інкапсулювати різні протоколи передачі даних і є незалежним від будь-яких протоколів механізмів передачі даних, допускають модифікацію або заміну алгоритмів керування, що входять в реалізовану технологію керування [6].

Для ефективного керування ресурсами мережі використовуються такі методи [7]:

- метод визначення профілю навантаження ланки телекомунікаційної мережі на основі аналізу пропускнуої здатності ланки;

- метод прогнозування фрактального трафіку, який використовує оцінки статистичних характеристик другого порядку і властивість масштабної інваріантності трафіку, що дозволяє на основі даних про відліки, отриманих до фіксованого моменту часу, отримати оцінки відліків і можливого числа сингулярностей в поведінці трафіку на інтервалі прогнозування при виборі кратно-масштабних коефіцієнтів кореляції відліків;

- метод динамічного керування розподілом навантаження віртуальних з'єднань, що враховує при прогнозуванні фрактальний характер створюваного трафіку;

- методи керування перерозподілом пропускнуої здатності віртуального з'єднання з урахуванням пріоритетів і конкуренції між інтегральними потоками даних при динамічному резервуванні пропускнуої здатності.

Більшість методів керування трафіком можна розділити на алгоритми зі зворотним зв'язком (ЗЗ) і без зворотного зв'язку. Методики без ЗЗ не враховують поточного стану мережі, приймаючи рішення про відкидання пакетів і складаючи розклад роботи окремих ділянок мережі.

Алгоритми із ЗЗ враховують поточний стан мережі, ведуть моніторинг виникнення перевантажень, виконують оповіщення інших ділянок мережі і приймають заходи по усуненню перевантажень. Моніторинг мережі полягає в реєстрації наступних показників [8]:

- відсоток пакетів, які не прийняті до обробки через переповнення буфера;
- середня довжина черги;
- відсоток пакетів, переданих повторно через відсутність підтвердження про отримання;
- середнє (середньоквадратичне відхилення) часу затримки пакетів.

Збільшення значень наведених показників говорить про наростаюче навантаження або про наявність перевантаження. Алгоритми з явним ЗЗ можуть бути бінарними, тобто прямо вказувати на наявність перевантаження, або надавати інформацію про конкретний рівень навантаження.

В результаті проведеного аналізу методів керування трафіком та забезпечення QoS в TCP/IP-мережах можна зазначити, що головним питанням щодо забезпечення керування КМ є ефективне керування пропускнуою здатністю, в тому числі через поділ смуги пропускання. Алгоритми керування трафіком і боротьби з перевантаженнями без ЗЗ спрямовані на запобігання перевантажень, а виявлення та керування перевантаженнями здійснюється алгоритмами із ЗЗ, які враховують, наскільки це можливо, поточний стан мережі. Тому застосування методів прогнозування мережевого трафіка в реальному часі в алгоритмах керування чергами може покращити їх роботу.

Проаналізуємо методи та моделі, що найчастіше використовуються для прогнозування трафіка, та оцінимо їх придатність для здійснення прогнозу в реальному часі.

Метод прогнозування являє собою сукупність операцій та прийомів, з допомогою яких можна передбачати майбутній розвиток процесу. Тобто метод прогнозування – це послідовність дій, необхідних для отримання прогнозу [9].

Моделлю прогнозування називають функціональне уявлення, яке адекватно описує процес, що досліджується, та є основою для отримання його майбутніх значень.

На сьогодні, за оцінками спеціалістів, існує близько 150 різних методів прогнозування, хоча на практиці найчастіше використовують 15-20 [10]. Найбільш загальною класифікацією методів прогнозування є поділ їх на кількісні (формалізовані) та якісні (інтуїтивні).

Кількісні методи використовують інформацію, отриману на основі спостережень за зміною параметрів чи за наявності статистично достовірних залежностей, що характеризують процес (до таких методів відносять аналіз часових рядів, причинно-наслідкове моделювання і т.д.). Такі методи використовуються в разі, якщо інформація про об'єкт прогнозування носить в основному кількісний характер, а вплив різних чинників можна описати за допомогою математичних формул.

Якісні методи ґрунтуються на експертних оцінках фахівців (прикладом таких методів є методи експертних оцінок, моделі очікувань споживача і т.п.). Дані методи застосовують коли інформація кількісного характеру про об'єкт прогнозування відсутня або носить в основному якісний характер і вплив факторів неможливо описати математично.

Вищенаведені дві групи методів можна поділити за загальним принципом діяльності та способом отримання прогнозованої інформації. Кількісні методи поділяються на методи екстраполяції та методи моделювання.

До методів екстраполяції відносяться методи ковзного середнього, експоненціального згладжування, найменших квадратів. Ці методи досить широко застосовуються на практиці, оскільки їх перевагами є простота, дешевизна та відсутність необхідності у великій статистичній базі для розрахунків. Застосування даних методів передбачає, що основні впливаючі на процес чинники зберуть свій вплив у майбутньому, а досліджуване явище розвивається за плавною траєкторією, яку можна описати математично.

До методів моделювання відносять: методи інформаційного моделювання (патентний, публікаційний), статистичне моделювання, логічне моделювання (прогнозна аналогія, «дерево цілей»).

Дані щодо мережевого трафіка можна вважати часовим рядом, оскільки часовий ряд – це послідовність значень, отриманих через рівні проміжки, що описують процес, який перебігає у часі. Проте методи прогнозування часових рядів по відношенню до мережевого трафіка [2, 11] дослідники стали застосовувати відносно недавно.

Задача прогнозування значень часового ряду не є новою, але дослідження в даній області не втрачають актуальності і постійно виникають нові алгоритми і методики прогнозування [12].

В роботах, присвячених цим питанням, найчастіше досліджують авторегресійні моделі прогнозування, рідше – на основі нейромереж. Крім того, алгоритмів прогнозування часових рядів і їх варіацій досить багато, а дослідження кожного з них досить трудомістке завдання. Таким чином, задача керування трафіком на основі прогнозів стану мережі залишається вельми актуальною.

В роботі [13] моделі прогнозування поділяються на статистичні та структурні. До статистичних моделей належать регресійні та авторегресійні, а також моделі згладжування. Моделі прогнозування на основі нейромереж, ланцюгів Маркова та класифікаційних дерев відносяться до структурних моделей. Класифікація моделей часових рядів наведена на рис. 1.



Рисунок 1 – Загальна класифікація моделей прогнозування на основі часових рядів

Формально задача прогнозування часового ряду $X(t) = X(1), X(2), \dots, X(T)$ в момент часу T полягає у визначенні значень ряду $X(t)$ в моменти часу $T+1, T+2, \dots, T+T'$, де T' – час попередження або горизонт прогнозу.

Тобто визначення майбутнього стану процесу ґрунтується на відомих значеннях ряду, а прогноз будується з використанням моделі, що відображає функціональну залежність між майбутніми та минулими значеннями ряду.

Різні алгоритми згладжування часового ряду або розрахунку середнього значення можуть використовуватися в якості простої і наочної прогнозуючої моделі [14].

Найпростіший алгоритм змінного середнього полягає в обчисленні наступного значення часового ряду на підставі середнього за період:

$$X_{t+1} = \frac{X_t + X_{t-1} + X_{t-2} + \dots + X_{t-N+1}}{N}, \quad (1)$$

де X_{t+1} – прогнозоване значення ряду на момент $t+1$; X_{t-1} – значення ряду в момент $t-1$; N – період згладжування.

В даному випадку важливо підібрати відповідне значення N . Також логічно припустити, що значення ряду X_t мусить сильніше впливати на прогноз, ніж X_{t-N} . Алгоритм експоненціального згладжування враховує цей недолік через коефіцієнт:

$$X'_{t+1} = \alpha X_t + (1 - \alpha) X'_t, \quad (2)$$

де α – ваговий коефіцієнт, $\alpha \in [0, 1]$; X_t – реальне значення ряду в момент t ; X'_t – прогнозоване значення ряду.

Таким чином, ми отримуємо зважене ковзне середнє, де прогнозоване значення формується з прогнозованого і реального поточного значення ряду. При цьому коефіцієнт підбирається відповідно до характеру ряду – чим сильніше змінюються значення ряду, тим більшим має бути коефіцієнт, щоб швидше «приспосуватися» до великих змін. А чим вища шумова складова ряду, тим меншим має бути значення коефіцієнта, щоб згладжувати флуктуації.

Для вирішення завдання підбору коефіцієнта α розроблений алгоритм адаптивного згладжування.

Існує безліч варіацій алгоритму адаптивного згладжування, в найпростішому випадку коефіцієнт α підбирається таким чином:

$$\alpha_{t+2} = \left| \frac{X'_{t+1} - X_{t+1}}{X_{t+1}} \right|. \quad (3)$$

При цьому, якщо значення коефіцієнта виходить рівним або більшим за 1, то приймається $\alpha = 0,99(0,999)$, а якщо рівним 0, то $\alpha = 0,01(0,001)$. Наведені алгоритми згладжування часових рядів не враховують наявності у ряду тренда і сезонної складової. Алгоритм експоненційного згладжування з урахуванням тренда при згладжуванні враховує нахил ряду:

$$L_t = \alpha X_t + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad (4)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, \quad (5)$$

де L – рівень тренда; T – тренд; α – ваговий коефіцієнт, $\alpha \in [0, 1]$; β – ваговий коефіцієнт, $\beta \in [0, 1]$.

Тоді прогнозоване значення ряду на m кроків:

$$X_{t+m} = L_t + T_t m. \quad (6)$$

З урахуванням сезонної складової ряду прогноз будеться наступним чином:

$$L_t = \alpha \frac{X_t}{SA_{t-C}} + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad (7)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, \quad (8)$$

$$SA_t = \gamma \frac{X_t}{L_t} + (1 - \gamma)(L_{t-1} + SA_{t-C}), \quad (9)$$

де L – рівень тренда; T – тренд; SA_t – сезонна складова ряду на момент часу t ; C – розмір циклу для сезонної складової; α – ваговий коефіцієнт, $\alpha \in [0, 1]$; β – ваговий коефіцієнт, $\beta \in [0, 1]$; γ – ваговий коефіцієнт, $\gamma \in [0, 1]$;

Тоді прогнозоване значення ряду на m кроків вперед:

$$X_{t+m} = (L_t + T_t m) SA_{t-C+m}. \quad (10)$$

Хоча дана модель намагається не просто усереднити ряд ковзаючим вікном, але й врахувати його сезонну, трендову і циклічну складові, вона, як і решта згладжуючих моделей, відфільтровує досить багато інформації про реальний процес. Разом з шумами усувається інформація, яка могла б бути використана при прогнозуванні ряду.

Далі розглянемо регресійні і авторегресійні моделі, що відрізняються від алгоритмів згладжування більшою точністю та враховують більше факторів, що впливають на розподіл ряду.

У статистичній обробці даних регресійний аналіз застосовується для визначення залежності між початковою (вхідною) залежною змінною і зовнішніми незалежними змінними, т.зв. регресорами. Найпростіша регресійна модель називається лінійною і передбачає наявність єдиного зовнішнього фактора, що впливає на процес з лінійним зв'язком:

$$X_t = \beta_0 + \beta_1 Y_t + \varepsilon_t, \quad (11)$$

де Y_t – незалежна змінна; β_0 і β_1 – коефіцієнти регресії; ε_t – помилка моделі.

У разі множинної регресійної моделі зовнішніх факторів, що впливають на значення ряду – кілька:

$$X_t = \beta_0 + \beta_1 Y_t^1 + \beta_2 Y_t^2 + \dots + \beta_n Y_t^n + \varepsilon_t. \quad (12)$$

Залежність між вихідним процесом і регресорами не обов'язково має бути лінійною, можливий випадок, при якому залежність описується довільною функцією.

Часто при дослідженні процесу складно виділити зовнішні впливаючі фактори, при цьому регресійний аналіз ґрунтується лише на самих значеннях ряду, і такі моделі називають авторегресійними. Одна з найбільш відомих авторегресійних моделей – ARMA (Autoregressive Moving Average) – являє собою комбінацію авторегресійної моделі та ковзного середнього. Авторегресійна (AR (p)) модель може бути виражена таким чином:

$$X_t = c + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + \varepsilon_t, \quad (13)$$

де X_t – поточне значення ряду; ε_t – випадкова помилка; φ_i , ($i = 1, 2, \dots, p$) – вагові коефіцієнти; c – константа.

У записі AR (p), p – це порядок моделі. Наведемо видозмінений запис наведеної вище моделі ковзного середнього MA (q):

$$X_t = \mu + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t, \quad (14)$$

де μ – середнє значення вибірки; θ_j ($j = 1, 2, \dots, q$) – параметри моделі; q – порядок моделі.

Тоді ARMA (p, q) модель має вигляд:

$$X_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j}. \quad (15)$$

ARMA-модель використовує припущення, що ряд стаціонарний. Тобто необхідна операція усунення тренда, виконання якої в найкращому випадку призведе до стаціонарності ряду або хоча б усуне сам тренд.

Якщо в якості вихідних даних використовується не сам часовий ряд, а різниця його компонентів, то в такому випадку модель називається ARIMA (p, d, q) – Autoregression integrated moving average, де d – порядок різниці між значеннями ряду.

Ще одним широкоживим інструментом для прогнозування трафіка є штучні нейронні мережі (ШНМ) [15]. Вони часто використовуються в задачах прогнозування часових рядів, оскільки дозволяють імітувати процеси з множинним зовнішнім впливом. Особливо ефективними ШНМ є в задачах класифікації об'єктів і розпізнавання образів. Структурою нейрон ШНМ схожий на людський нейрон, а його схема представлена на рис. 2.

На вхід штучного нейрона надходить вектор вхідних сигналів $X_1, X_2, X_3, \dots, X_n$, для яких вираховується середнє значення z , яке використовується в активаційній функції F для розрахунку значення виходу Y .

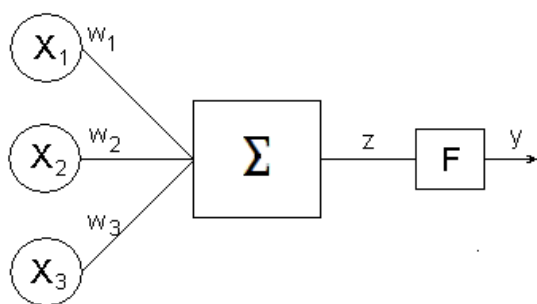


Рисунок 2 – Схема нейрона ШНМ

Активаційна функція може бути функцією одиничного стрибка, лінійною функцією і т.д.

Приклад ШНМ, яка має три вихідні і два приховані нейрони, наведено на рис. 3. Вхідний шар служить для розподілу вхідного вектора значень і не містить нейронів. Не всі нейрони обов'язково повинні бути пов'язані між собою, тобто для деяких зв'язків ваговий коефіцієнт w може бути нульовим.

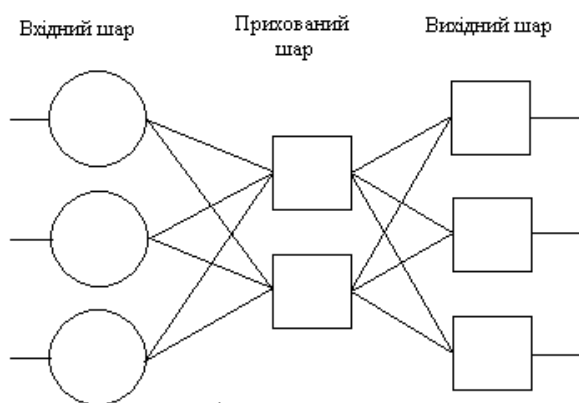


Рисунок 3 – Приклад структурної схеми нейронної мережі

Також ШНМ може мати зворотні зв'язки для переналаштування вагових коефіцієнтів. Вагові коефіцієнти зв'язків призначаються в процесі навчання, а також можуть перераховуватися в залежності від результатів роботи мережі.

Згідно з [16] можна виділити основні особливості ШНМ:

- ШНМ можуть моделювати нелінійні залежності між вхідними та вихідними даними;
- ШНМ навчається на вхідних даних, тобто дані визначають модель між вхідними та вихідними даними;
- ШНМ можуть узагальнювати вхідні дані, що не призводить до зниження ефективності роботи мережі при зміні характеру даних;
- ШНМ, на відміну від статистичних моделей, можуть вносити припущення про розподіл вхідних даних.

В ході використання ШНМ виникає ряд труднощів:

- складність підбору оптимальної комбінації параметрів ШНМ, таких як швидкість навчання, число прихованих шарів, число прихованих нейронів у кожному шарі;

– складність вибору алгоритму навчання і суворі вимоги до навчальної вибірки;

– навчена ШНМ є певною мірою «чорним ящиком», тобто досить складно виділити сукупність правил, за якими вона приймає рішення.

Алгоритм пошуку оптимальної ШНМ для числового ряду наступний:

- випадковим чином генерується вихідна популяція (набір хромосом) ШНМ;
- розраховується фенотип (архітектура ШНМ) і значення оптимізаційної функції для кожної «особини» популяції;
- для формування нової популяції застосовуються методи генетичних алгоритмів – елітизм, відбір вдалих особин, кросовер і мутація;
- попередні два кроки повторюються до настання ефекту перенавчання мережі.

В роботі [17] авторами пропонується використання для задач прогнозування алгоритму побудови локальної моделі керованого процесу (ЛМКП).

ЛМКП, концепція якої запропонована в [18], синтезується в реальному часі через спостереження за поведінкою об'єкта керування, та дозволяє здійснювати прогнозування, спираючись на пробні та синтезуючі впливи. Даний алгоритм призначений для використання в задачах керування при невідомій динаміці об'єкта керування і зовнішніх впливах. Метод активно здобуває інформацію про об'єкт керування і зовнішні впливи шляхом спостереження і пробних керувальних впливів. Будується сигнал керування, еквівалентний дії зовнішніх впливів, після чого задача прогнозування зводиться до побудови регресійної залежності та її апроксимації та екстраполяції отриманого керувального впливу.

Використання ЛМКП дає змогу замінити задачу прогнозування при відомому вхідному і вихідному значеннях і невідомих зовнішніх впливах на прогнозовану величину, на задачу прогнозування при відомих вхідному і вихідному значеннях і відсутності зовнішніх впливів незалежно від їх кількості.

Тобто використання ЛМКП в задачах прогнозування не вимагає накопичення і обробки значної кількості статистичної інформації щодо об'єкта.

В роботі [19] запропоновано методіку короткострокового прогнозування трафіка на основі ЛМКП. Прогнозування здійснюється за наступним алгоритмом:

1) Отримати відомості про кількість пакетів, які надійшли за перші 2 відліки часу.

2) Обчислити значення df_i , y_{c_i} , K , δ_i , y_{up_i} :

– першу похідну df_i (на основі кількості пакетів, що надійшли за 2 попередні відліки часу):

$$df_i = \frac{y_{u_i} - y_{u_{i-2}}}{2 \cdot h} \quad (16)$$

– величину y_{c_i} :

$$y_{c_i} = (f_i + \delta_i) \cdot K \quad (17)$$

– масштабний коефіцієнт K (обчислюється лише 1 раз на 2-му кроці побудови ЛМКП):

$$K = \frac{y_{c_i} - y_{up_i}}{\delta_i} \quad (18)$$

- коефіцієнт корекції δ_i :

$$\delta_i = \begin{cases} \frac{y_{up_i}}{K}, & i \text{ не парне} \\ \frac{i-1}{i-2} \cdot \delta_{i-1}, & i \text{ парне} \end{cases} \quad (19)$$

- прогнозоване значення кількості пакетів y_{up_i} :

$$y_{up_i} = \begin{cases} 2 \cdot y_{up_{i-1}} - y_{up_{i-2}} + df_i, & i \text{ не парне} \\ y_{c_i} - \delta_i \cdot K + df_i, & i \text{ парне} \end{cases} \quad (20)$$

3) Продовжувати виконання кроку 2, крім (18), доки здійснюється прогнозування.

Блок-схема алгоритму прогнозування на основі ЛМКП наведена на рис. 4. Прогнозування здійснюється за припущення, що у часовому проміжку побудови ЛМКП динаміка мережі являє собою масштабу ланку.

Для перевірки працездатності алгоритму було взято трафік локальної мережі, побудованої за технологією Ethernet. Період спостереження – 115 с.

Графік результатів моделювання трафіку з допомогою ЛМКП наведено на рис. 5. Графік результатів прогнозування трафіка для тих же даних з використанням ARIMA наведено на рис. 6.

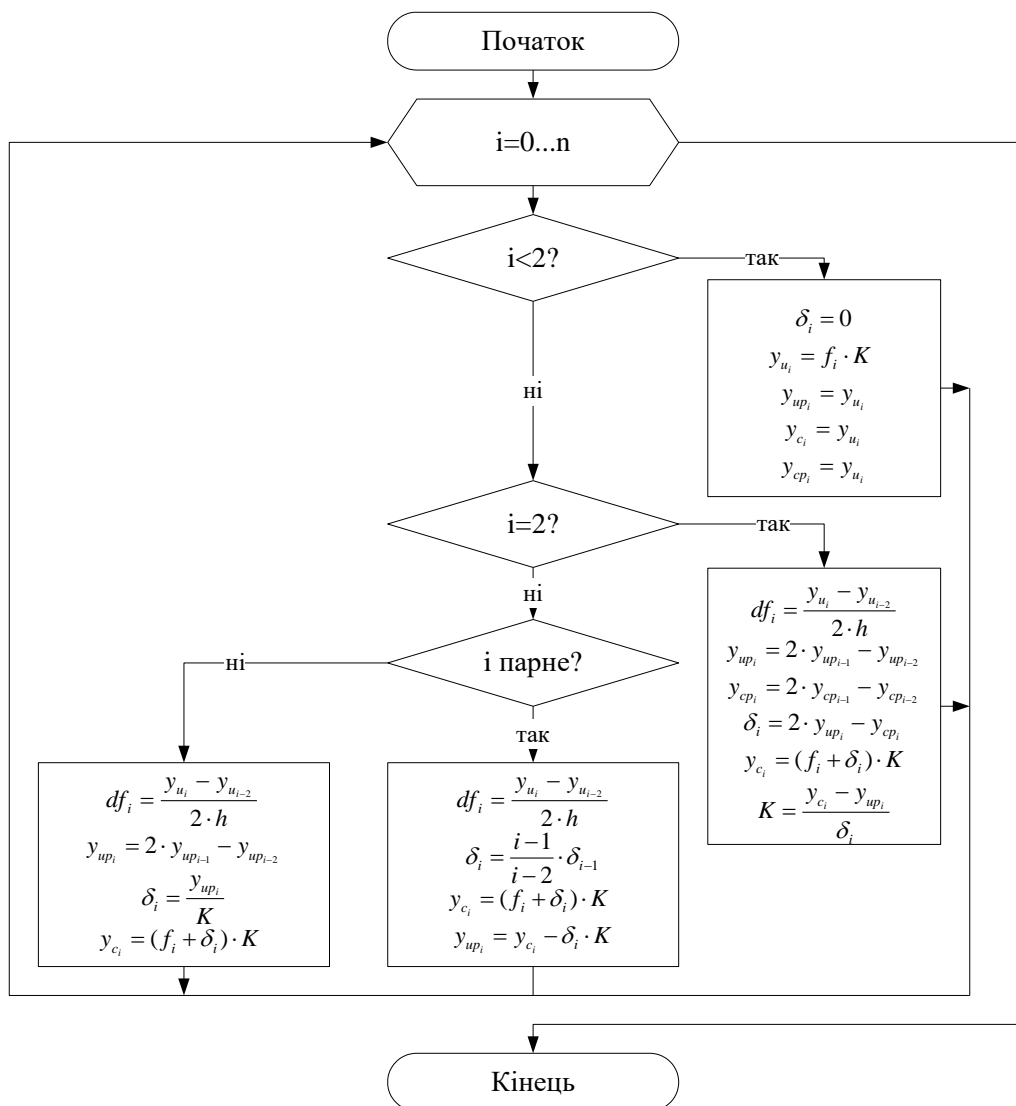


Рисунок 4 – Блок-схема алгоритму прогнозування на основі ЛМКП

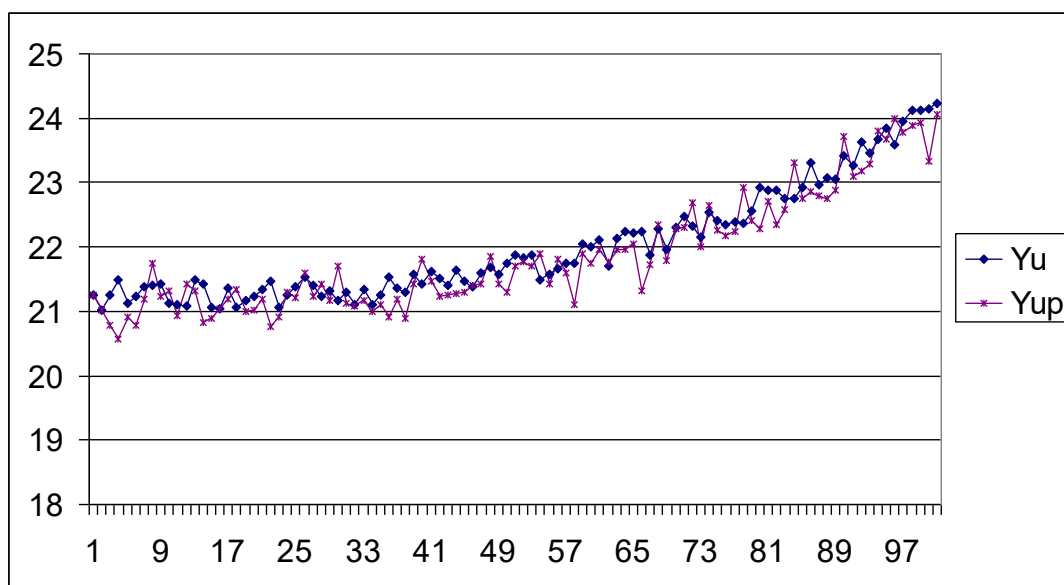


Рисунок 5 – Результати моделювання трафіка з використанням алгоритму побудови ЛМКП
 Y_u – вхідний трафік, Y_{up} – прогнозоване значення

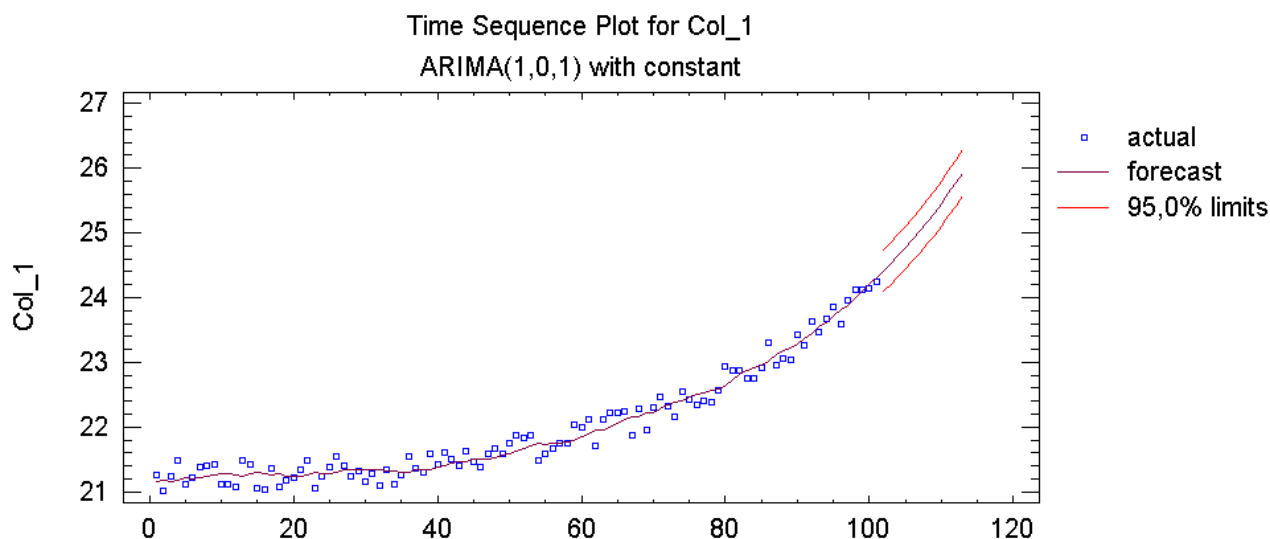


Рисунок 6 – Результати прогнозування трафіка з використанням ARIMA

Як видно з наведених графіків результатів прогнозування, ARIMA відслідковує лише загальний тренд (графік ніби «усереднює» значення), а ЛМКП дозволяє відслідковувати поточні тренди, що в зв'язку з великою пачечністю мережевого трафіку дозволяє більш адекватно здійснювати прогнозування в реальному часі.

Згідно з результатами досліджень [20], прогнозування з використанням алгоритму побудови ЛМКП дає значення середньої абсолютної похибки в межах, отримуваних при використанні стандартних методів прогнозування часових рядів. Однією з головних переваг даної моделі є малий інтервал спостережень, необхідний для початку створення ЛМКП (лише 2 значення ряду), що дозволяє здійснювати прогнозування трафіка в реальному часі. Інші його переваги – простий математичний апарат та можливість виконувати прогнозування в умовах невизначеності.

ВИСНОВКИ. У статті проаналізовано моделі та методи прогнозування, що можуть бути використані для прогнозування мережевого трафіку в реальному часі, наведено їх переваги та недоліки. Практичне значення роботи полягає у здійсненні аналізу та логічного узагальнення відомих результатів наукових досліджень щодо прогнозування трафіка.

Недоліками методів аналізу часових рядів є необхідність накопичення статистичної інформації для попереднього аналізу та вибору моделей, що призводить до необхідності повторного проведення даної процедури при зміні параметрів КМ, трафік якої прогнозується. Недоліками ШНМ є необхідність в досить тривалому процесі підбору оптимальної структури мережі та тривалому навчанні. Даних недоліків позбавлений метод прогнозування на основі алгоритму побудови ЛМКП, що не вимагає накопичення статистичної інформації, вибору моделей чи тривалого навчання, що робить його більш

універсальним та нечутливим до змін параметрів КМ. Крім того, більшість методів прогнозування працюють з умовою, що чинники, які впливають на процес, залишаються незмінними. Прогнозування на основі ЛМКП вільне від цього обмеження, оскільки в кожному циклі прогнозування модель будується заново, з урахуванням поточних діючих чинників.

Тому прогнозування на основі ЛМКП є найбільш перспективним з перерахованих методів для прогнозування поведінки мережевого трафіка в реальному часі та застосування в методах керування трафіком і забезпечення QoS.

ЛІТЕРАТУРА

1. Кузьмин В. В. Классификация и идентификация трафика в мультисервисной сети оператора связи. *Электронный журнал «Современные проблемы науки и образования»*. 2014. № 5. URL: <http://www.science-education.ru/article/view?id=15039> (дата звернення: 10.02.2019)
2. Методика краткосрочного прогнозирования трафика телекоммуникационных сетей. Е. Г. Игнатенко и др. *Збірник наукових праць ДонІЗТ. Вип. 28*. Донецьк: ДонІЗТ, 2011. С. 102–108.
3. Славко О. Г. Інформаційна технологія керування перевантаженнями в мультисервісних телекомунікаційних мережах. *Вісник Кременчуцького національного університету імені Михайла Остроградського*. Кременчук: КрНУ, 2011. Вип. 2 (67), част. 1. С. 29–34.
4. Кучерявый Е. А. Управление трафиком и качество обслуживания в сети Интернет. СПб.: Наука и Техника, 2004. 336 с.
5. Luc De Ghein. MPLS Fundamentals. Cisco Press, 2006. URL: <https://doc.lagout.org/network/Cisco/CCIE/CCIE%20SP/CiscoPress%20-%20MPLS%20Fundamentals.pdf> (дата звернення: 21.01.2019)
6. Фомін М. М., Житник І. В. Обгрунтування впровадження технології мультипротокольної комутації по міткам як основи транспортної мережі зв'язку. *Сучасні інформаційні технології у сфері безпеки та оборони*. К.: НУОУ ім. І. Черняхівського, 2014. № 3 (21). С. 64–68.
7. Поштаренко В. М., Кравченко В. С. Методы борьбы с перегрузками на критических участках сети. *Вісник НТУ «ХПІ». Серія: Техніка та електрофізика високих напруг*. Х.: НТУ «ХПІ», 2013. № 27 (1000). С. 129–138.
8. Шелухин О. И., Осин А. В., Смольский С. М. Самоподобие и фракталы. ФИЗМАЛИТ, 2008. 368 с.
9. Чучуева И. А. Модель прогнозирования временных рядов по выборке максимального подобия,

дисс. ... канд. тех. наук: 05.13.18 / Московский государственный технический университет им. Н. Э. Баумана. Москва, 2012. 155 с.

10. Грабовецкий Б. С. Основы економічного прогнозування: Навчальний посібник. Вінниця: ВФ ТАНГ, 2000. 209 с.

11. Гребенников А. В., Крюков Ю. А., Чернягин Д. В. Моделирование сетевого трафика и прогнозирование с помощью модели ARIMA. *Электронный журнал «Системный анализ в науке и образовании»*. 2011. № 1. URL: www.sanse.ru/download/79 (дата звернення: 17.03.2019)

12. Nozian M., Maiza H. A., Zuhaimy I. Improving Short Term Load Forecasting Using Double Seasonal Arima Model. *World Applied Sciences Journal* V. 15 (2). 2011. P. 223–231.

13. Yang J. Power System Short-term Load Forecasting: Thesis for Ph.d degree. Germany, Darmstadt, Elektrotechnik und Informationstechnik der Technischen Universitat, 2006. 139 p.

14. Ганчук А. А., Соловійов В. М., Чабаненко Д. М. Методи прогнозування: навч. посіб. / Черкаси: Брама-Україна, 2012. 140 с.

15. Хлапонін Ю. І., Жиров Г. Б., Нікітчин О. М. Застосування нейронних мереж в статистичній системі аналізу і моніторингу телекомунікаційних мереж. *Технологічний аудит і резерви виробництва*, т. 5, № 2, 2016. С. 35–41.

16. Томашевський О. М., Цегелик Г. Г., Вітер М. Б., Дубук В. І. Інформаційні технології та моделювання бізнес-процесів: навч. посіб. К.: Видво «Центр учбової літератури», 2012. 296 с.

17. Сохін Н. Л., Гученко М. І. Використання локальної моделі керованого процесу в задачах прогнозування. *VI Всеукраїнська наукова конференція «Математичне моделювання та математична фізика»*, 4–6 жовтня 2018 р. Кременчук: КрНУ, 2018. С. 11–13.

18. Гученко М. І. Активно-резонансний алгоритм стабілізації. *Нові технології: науковий вісник Інституту економіки та нових технологій ім. Ю. І. Кравченка*. 2003. № 1 (2). С. 57–61.

19. Сохін Н. Л., Гученко М. І. Методика короткострокового прогнозування трафіка. *Міжнародна науково-практична інтернет-конференція «Актуальні проблеми і перспективи інноваційного розвитку економіки та техніки в умовах інтеграції України в Європейський науково-виробничий простір»*, 14–15 травня 2019 р. Кременчук: ДП «УкрНДІВ», 2019, 147 с. С. 16–18.

MODELS AND METHODS OF NETWORK TRAFFIC FORECASTING IN REAL TIME

N. Sokhin, M. Guchenko, A. Kirska

Kremenchuk Mykhailo Ostrohradskyi National University

vul. Pershotravneva, 20, Kremenchuk, 39600, Ukraine. E-mail: work2301@gmail.com

Purpose. The paper analyzes the models and methods which can be used to forecast network traffic in real time. **Methodology.** The authors use analysis, synthesis, comparison, and descriptive method, as well as forecasting methods (time series analysis). **Findings.** In this paper analyzed the reasons of overloads occurrence in computer networks and methods of preventing overloads. Also it shown their connection with traffic management and quality of service (QoS) support. Substantiated the necessity of using network traffic forecasting for improvement of overloads prevention methods. The analysis of models and methods that can be used to forecast the traffic of computer networks in real-time is

given. In this paper reviewed the advantages and disadvantages of using time series analysis methods and artificial neural networks in the task of computer networks traffic forecasting. Also the authors described the method of using the algorithm for synthesis local model of controlled process (LMCP) for the obtaining of computer network traffic forecast. The main scientific results of the article consist in the analysis and synthesis of known results of scientific researches in field real-time computer network traffic forecasting. **Originality.** It is shown that using algorithm for synthesis LMCP allows to forecast computer network traffic in real time, without the need for preliminary accumulation, processing and analysis a large number of statistical data, does not require the choice of models and assess their suitability. **Practical value.** The obtained results can be used in solving practical problems in the computer networks research. They can also be used in network traffic management and QoS improvements. **Conclusions.** The method of forecasting network traffic on the basis of algorithm for synthesis LMCP is devoid of the disadvantages of time series forecasting methods and artificial neural networks. This makes it more universal and insensitive to changes in computer network parameters. References 19, figures 6.

Key words: real-time forecasting, network traffic, QoS, auto-regression models, neural networks, local model of controlled process.

REFERENCES

1. Kuzmyn, V. V. (2014), "Traffic classification and identification in a telecom operator multiservice network", *Elektronnij zhurnal «Sovremennye problemi nauky i obrazovannya»*, no. 5. URL: <https://www.science-education.ru/ru/article/view?id=15039> (accessed: 10, February 2019).
2. Ygnatenko, E. G. (2011), "Methods for short-term forecasting of telecommunication network traffic", *Zbirnyk naukovykh prats DonIIT*, no. 28, pp.102–108.
3. Slavko, O. G. (2011), "Information technology of congestion control in multiservice telecommunication networks", *Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*, no. 2 (67), part. 1, pp. 29–34.
4. Koucheryavy, E. A. (2004), *Upravlenye trafikom i kachestvo obsluzhivaniya v sety Internet* [Traffic control and QoS support in IP-based networks], Nauka i tekhnika, Saint Petersburg, Russia.
5. De Ghein, L. (2006), *MPLS Fundamentals*. Cisco Press. URL: <https://doc.lagout.org/network/Cisco/CCIE/CCIE%20SP/CiscoPress%20-%20MPLS%20Fundamentals.pdf> (accessed 21, January 2019)
6. Fomin, M. M., Zhytnyk, I. V. (2014), "Implementation study of multi-protocol commutation technology on labels as the transport communication network basis", *Suchasni informatsijni tekhnologii u sferi bezpeki ta oboroni*, no. 3 (21), pp. 64–68.
7. Poshtarenko, V. M., Kravchenko, V. S. (2013), "The methods of struggle against congestion in the critical parts of the network", *Visnyk NTU «KPI». Seriya: Tekhnika ta elektrofizyka vysokykh naprug*, no. 27 (1000), pp. 129–138.
8. Shelukhyn, O. Y., Osyn, A. V., Smolskyj, S. M. (2008), *Samopodobye i fraktali* [Self-similarity and fractals], Fyzmalyt, Moscow, Russia.
9. Chuchueva, Y. A. (2012), *Model prognozyrovannya vremennikh ryadov po viborce maksimalnogo podobyia* [Time series prediction model for maximum similarity sample], Moskovskij gosudarstvennij tekhnicheskij unyversytet ym. N. E. Bauman, Moscow, Russia.
10. Grabovetskyj, B. E. (2000), *Osnovy ekonomichnogo prognozuvannya* [Basics of economic forecasting], VF TANG, Vinnytsya, Ukraine.
11. Grebennikov, A., Krukov, Yu., Chernyagin, D. (2011), "Modeling network traffic and forecasting with model ARIMA", *Elektronnij zhurnal «Systemnij analiz v nauke i obrazovannyi»*, no. 1. URL: www.sanse.ru/download/79 (accessed 17, March .2019)
12. Nozian, M., Maiza, H. A., Zuhaimy, I. (2011), "Improving Short Term Load Forecasting Using Double Seasonal Arima Model", *World Applied Sciences Journal* V. 15 (2), pp. 223–231.
13. Yang, J. (2006), *Power system short-term load forecasting*: Thesis for Ph.d degree. Elektrotechnik und Informationstechnik der Technischen Universitat, Darmstadt, Germany.
14. Ganchuk, A. A., Solovjov, V. M., Chabanenko, D. M. (2012), *Metody prognozuvannya* [Forecasting methods], Brama-Ukraina, Cherkasy, Ukraine.
15. Khlaponin, Yu., Zhyrov, G., Nikitchin, A. (2016), "Neural networks application in the statistical system for telecommunication networks analysis and monitoring", *Technology audit and production reserves*, V. 5, no. 2, pp. 35–41.
16. Tomashevskij, O. M., Tsegelyk, G. G., Viter, M. B., Dubuk, V. I. (2012), *Informatsijni tekhnologii ta modelyuvannya biznes-protsesiv* [Information technology and business process modeling], Tsentru uchbovoi literatury, Kyiv, Ukraine.
17. Sokhin, N. L., Guchenko, M. I. (2018), "Using a local model of the controlled process in forecasting tasks", *VI Vseukrainska naukova konferen-tsiya "Matematychni modelyuvannya ta matematychna fizyka"* [VI Ukrainian Scientific Conference "Mathematical Modeling and Mathematical Physics"], Kremenchuk, October 04–06, 2018, pp. 11–13.
18. Guchenko, M. I. (2003), "Active-resonant stabilization algorithm", *Novi tekhnologii: naukovyj visnyk Instytutu ekonomiky ta novykh tekhnologij im. Yu. I. Kravchenka*, no. 1 (2), pp. 57–61.
19. Sokhin, N. L., Guchenko, M. I. (2019), "Method of short-term traffic predicting", *Mizhnarodna naukovo-praktychna internet-konferentsiya «Aktualni problemy i perspektivy innovatsijnogo rozvytku ekonomiky ta tekhniky v umovakh integratsii Ukrainy v Evropejskij naukovo-vyrobnychij prostir»* [Actual problems and prospects of innovative economics and technology development in the conditions of Ukraine's integration into European research and production space], Kremenchuk, May 14-15, 2019, pp. 16–18.

Стаття надійшла 29.05.2019.