

## Висновки

Запропонований метод розрахунку транспортної мережі дає можливість розрахувати трафік між вузлами мережі при зміні навантаження і розподілу потоків між вузлами мережі. Аналіз матриці розподілення трафіка (Тс) дає змогу визначити доцільність оптимізації топології мережі. Оптимізація топології ґрунтується на дослідженні матриці розподілення трафіка (Тс), корегування топології та розподілу трафіка між вузлами мережі в такий спосіб, щоб досягти приблизно однакового завантаження лінійних трактів, вираженого числом цифрових потоків Е1.

1. *Вербовецкий А. А. Основы проектирования цифровых оптоэлектронных систем связи. – М., 2000.* 2. *Сергеева Т.М., Баркова И.В. Оптимизация проектирования сетей SDH с резервированием / Вестник “Связи”. – 2003.* 5. *Климаш М.М., Романчук В.І. Розрахунок ефективності використання пропускної здатності каналу для різних видів трафіка та мережевих технологій / Вісник НУ”Львівська політехніка” “Ком’ютерні системи проектування: теорія і практика”. – 2003.*

УДК 621.382.33:681

**Маркіян Павликевич, Орест Костів, Ольга Шаталова, Мар’ян Кирик**  
Національний університет “Львівська політехніка”,  
кафедра телекомунікацій

## АНАЛІЗ, ПРОГНОЗУВАННЯ ТА УПРАВЛІННЯ ТРАФІКОМ ЗОВНІШНЬОГО КАНАЛУ WEB-ВУЗЛА

© Павликевич Маркіян, Костів Орест, Шаталова Ольга, Кирик Мар’ян, 2004

**Наведено результати аналізу та дослідження трафіка зовнішнього каналу Інтернет-провайдера. Здійснено спробу спрогнозувати цей трафік на основі популярних аналітичних методів прогнозування часових рядів.**

**There are given results of the analysis and investigation of the Internet-provider external channel traffic. The attempt to predict given traffic on the base of popular analytical procedures of time series prediction has been made.**

Стрімкий розвиток інформаційних та телекомунікаційних технологій, що викликає, своєю чергою, зростання попиту на інфокомунікаційні послуги, зокрема Інтернет-послуги, вимагає розробки нових алгоритмів збільшення пропускної здатності каналів зв’язку. Одним із шляхів вирішення цієї проблеми є оптимізація проходження тих чи інших інформаційних застосувань залежно від їх пріоритетності та розміру.

Ефективність надання послуг кінцевим користувачам визначається кількома чинниками, одним з яких є завантаженість каналу Інтернет-провайдера. Тому дослідження розподілу трафіка Інтернет-провайдера за популярними протоколами, прогнозування навантаження у його зовнішньому каналі відносно тижня та годин доби, а також управління трафіком кожного з популярних Інтернет-протоколів на основі такого прогнозу є актуальною темою наукових досліджень на сучасному етапі [1].

### Аналіз трафіка зовнішнього каналу

Такі дослідження протягом двох років проводяться спільно кафедрою “Телекомунікації” та Центром інформаційних та телекомунікаційних технологій Національного університету “Львівська політехніка” на основі статистичних даних завантаження зовнішнього каналу Інтернет-вузла.

Схема організації Інтернет-вузла зображена на рис. 1. Збір даних проводився як на сервері баз даних, на якому зберігаються лог-файли усіх користувачів Інтернету, яких обслуговує цей Інтернет-провайдер (користувачі, які працюють через локальну мережу, комутовані та некомутовані лінії), так і на кеш-сервері інтернет-вузла.

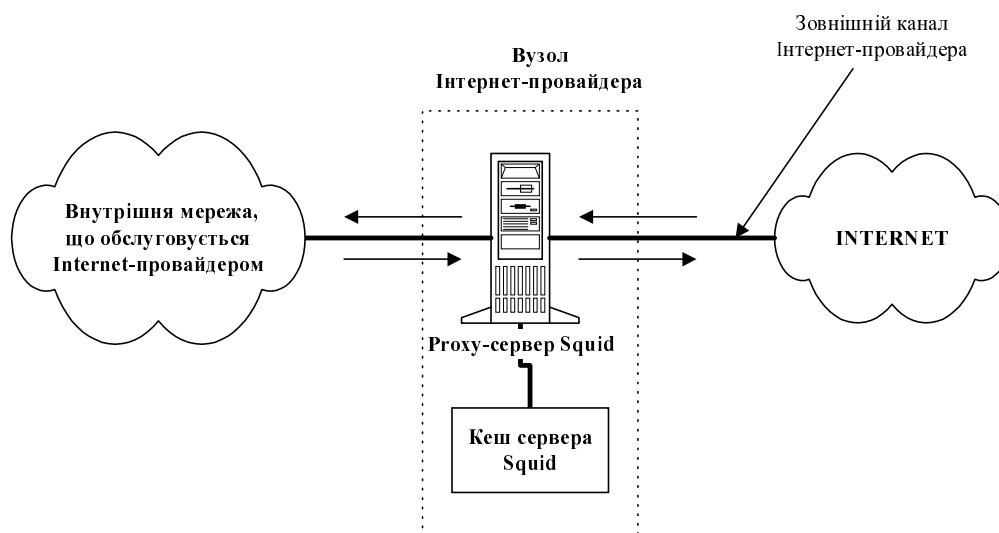


Рис. 1. Схема організації Інтернет-вузла

### Аналіз трафіка зовнішнього каналу

Для дослідження трафіка зовнішнього каналу було використано спектральний аналіз Фур'є для всього періоду спостережень.

$$x(t) = \sum_{n=1}^{\infty} c_n e^{i n \omega_0 t}, c_n = \frac{1}{T} \int_T x(t) \cdot e^{-i n \omega_0 t} dt. \quad (1)$$

Для наших досліджень спектр являє собою інтерес з часовим кроком  $1/T$ . Дискретність часового ряду приводить до періодичності його спектра. У цій роботі нас цікавить тільки  $P_n = |c_n|^2$  – спектр потужності. Як видно з рис. 2, максимум спектра потужності виявлено на 7-му дні, що відповідає тижневій періодичності.

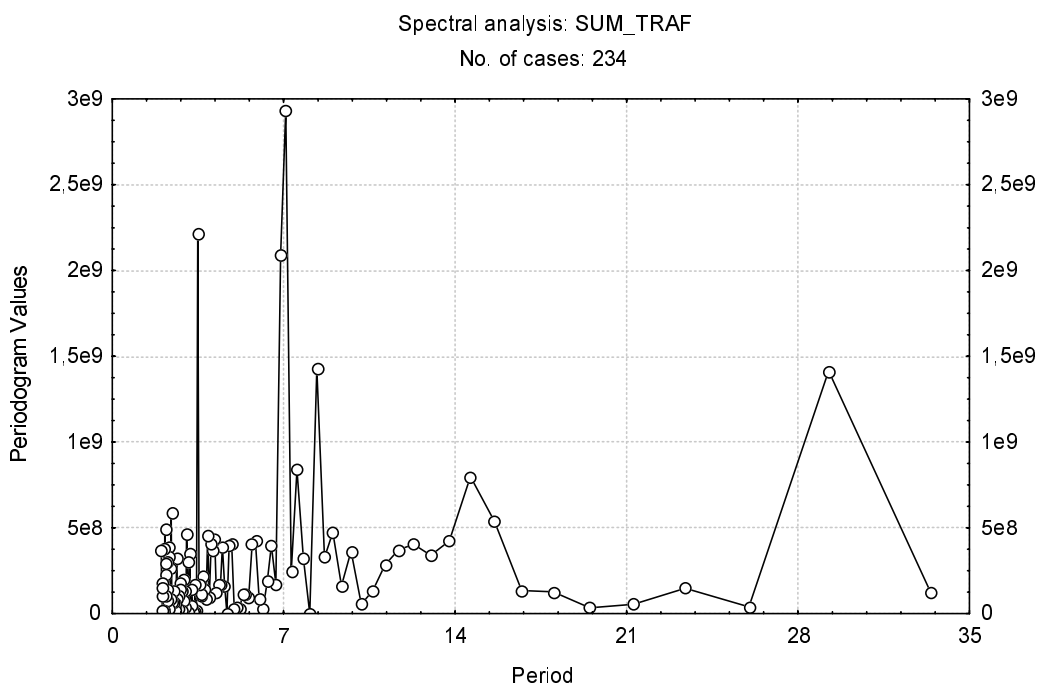


Рис. 2. Періодограма спектрального аналізу трафіка Інтернет-провайдера

### Прогнозування трафіка зовнішнього каналу

Отже, задачу прогнозування трафіка у зовнішньому каналі Інтернет-провайдера можна звести до задачі аналізу та прогнозування часових рядів, що ґрунтується на припущенні, що періодичні закономірності у файлі даних спостерігаються через рівні проміжки часу [2], тобто, маючи послідовність часового ряду

$$X=(x_1, x_2, x_3, \dots, x_n), \quad (2)$$

можна з високою ймовірністю побудувати його продовження

$$X'=(x_{n+1}, x_{n+2}, x_{n+3}, \dots). \quad (3)$$

Результати прогнозу трафіка, здійсненого популярними аналітичними методами [3] та за допомогою нейромережевих технологій [4], наведені у табл. 1.

Таблиця 1

#### Результати точності прогнозу трафіка зовнішнього каналу

Метод	Точність прогнозу (MAPE)
<b>Метод авторегресії</b>	18,85 %
Метод авторегресії та проінтегрованого ковзного середнього	15,495 %
Метод сезонної декомпозиції (Census I)	23,859 %
Багат шаровий перцептрон з однією вхідною змінною	25,8 %
Багат шаровий перцептрон з двома вхідними змінними	18,8 %
Багат шаровий перцептрон з трьома вхідними змінними	7,5 %

### Управління трафіка на основі прогнозу

У результаті аналізу було статистично виведено такі співвідношення для трафіка WEB-вузла (табл. 2), а також встановлено, що вхідний трафік перевищує вихідний у 6,5 рази.

Таблиця 2

#### Статистичні співвідношення для трафіка WEB-вузла

Протокол	HTTP	SMTP	FTP	SSH
Вхідний трафік	79 %	10 %	11 %	-
Вихідний трафік	33 %	62 %	1 %	4 %
Сумарний трафік	73 %	17 %	9 %	1 %

Отже, як видно з табл. 2, основними протоколами, які використовуються при роботі з Інтернет, є протоколи *HTTP* та *SMTP*, трафік яких займає 90 % сумарного трафіка WEB-вузла. Таку закономірність можна пояснити як об'єктивними умовами (скажімо, не кожний з користувачів має право запустити FTP-сесію з віддаленого терміналу), так і суб'єктивними, коли більшість користувачів при доступі до Інтернет-серверів використовують тільки HTTP-протокол, навіть якщо є альтернатива здійснити ці дії за допомогою інших Інтернет-протоколів.

З одержаних числових рядів чітко проглядається періодичність активності користувачів мережею Internet – в будні дні тижня відбувається активна робота з мережею, а у вихідні – спад майже до нуля. Якщо для протоколу HTTP така періодичність є закономірною – у вихідні дні зменшується кількість включених машин, то для SMTP-протоколу така закономірність не є зовсім очевидною. Вона може бути пояснена двома чинниками: тим, що зменшується доступ до комп'ютерів, які відсилають листи на цей Інтернет-вузол, та тим, що кількість людей, які проводять час за комп'ютером у вихідні, є меншою, ніж у робочі дні тижня. Аналізуючи роботу HTTP-протоколу, було одержано залежність обсягу HTML-сторінок, які приймає вузол, від їх кількості. Очевидним є факт, що існує якась область, у яку з великою ймовірністю може попасти середньостатистична HTML-сторінка, а сторінки, які перевищують певний обсяг, зустрічаються рідко (рис. 3).

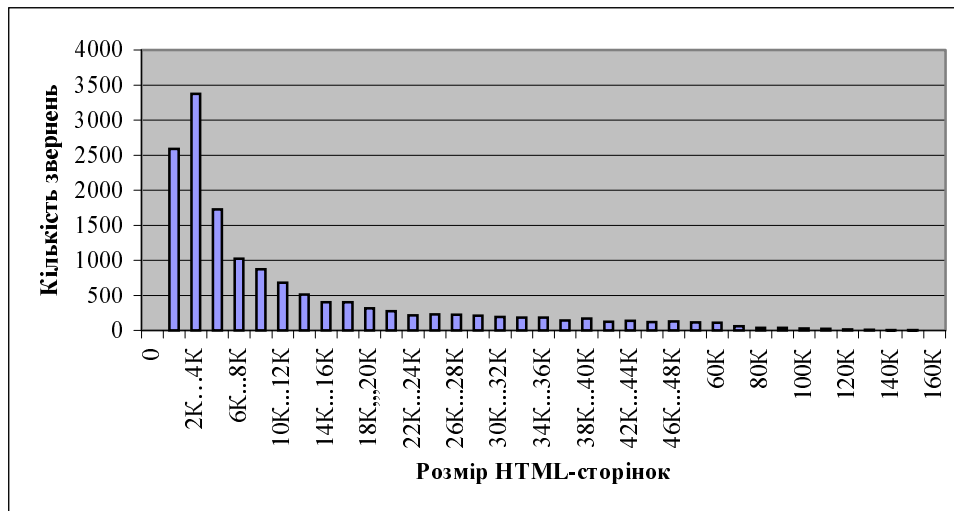


Рис. 3. Гістограма розподілу обсягів HTML-сторінок

Аналогічне твердження буде істинним для електронних листів, які переносяться за допомогою SMTP-протоколу, а також для передавання файлів за допомогою протоколу FTP. Таким чином, можна зробити висновок, що залежність частоти прийому файла Інтернет-сервером залежить в певний спосіб від розміру самого файла. На рис. 3 показано гістограму такої залежності для HTML-сторінок.

Зрозуміло, що обсяг вхідного трафіка, який передається через HTTP-протокол за одиницю часу, можна виразити формулою

$$V_{\text{HTTP}} = \int_0^b F(x) dx, \quad (4)$$

де  $V_{\text{HTTP}}$  – обсяг вхідного трафіка, переданий за допомогою HTTP протоколу за одиницю часу;  $x$  – обсяг HTML сторінки;  $b$  – обсяг максимальної HTML-сторінки.

$V_{\text{HTTP}}$  можна обчислити, застосовуючи, наприклад, формулу прямокутників. Отже, знаючи пропускну здатність зовнішнього каналу зв'язку та його завантаженість для вхідного трафіка протоколом HTTP, можна обмежити розмір HTML-сторінки, яка пропускатиметься Інтернет-провайдером до користувача у час найбільшого завантаження каналу. Цей алгоритм можна застосувати і для інших Інтернет-протоколів, що дасть змогу регулювати вхідний трафік Інтернет-провайдера.

### Висновки

1. Зважаючи на результати досліджень, задачу прогнозування трафіка пропонується здійснювати методами прогнозування часових рядів.
2. Зважаючи на результати прогнозу, запропоновано механізм регулювання трафіка за допомогою обмеження проходження певних інтернет-протоколів та довжини файлів у час найбільшого завантаження каналу.

1. Павликевич М.Й., Костів О.Л., Шаталова О.І. Методи прогнозування трафіку Інтернет-вузлів / Вісник Національного університету "Львівська політехніка" Радіоелектроніка та телекомунікації. – 2000. – № 477. – С. 127–134. 2. Боровиков В. STATISTICA: искусство анализа данных на компьютере. Для профессионалов. – СПб., 2001. 3. Прогнозирование в системе Statistica в среде Windows. – М., 2001. 4. Костів О.Л. Прогнозування трафіку SMTP протоколу за допомогою нейромережевих технологій. Моделювання та інформаційні технології: Збірник наукових праць. – Інститут проблем моделювання в енергетиці НАН України. – 2001. – № 18. – С.212–215.