

УДК 621.39:006.91 (02)

МЕТОДИ І ПРОЦЕДУРИ ПЕРВИННОЇ ОБРОБКИ ЗНАЧЕНЬ ПАРАМЕТРІВ ТЕЛЕКОМУНІКАЦІЙНОЇ МЕРЕЖІ

В. О. Бреславський

Державний університет телекомунікацій

вул. Солом'янська, 7-а, м. Київ, 03110, Україна. E-mail: slaava@i.ua

Для вирішення завдань розробки наукових основ дослідження загальних властивостей і принципів функціонування телекомунікаційних систем і комп'ютерних мереж, а також із метою розробки принципово нових методів аналізу, спрямованих на поліпшення їх технічних характеристик, а також із метою забезпечення контролю і діагностики розглянуто методи математичного моделювання, що ґрунтуються на статистичних методах аналізу багатовимірних даних. Зокрема, розглянуто методи первинної обробки значень параметрів телекомунікаційної мережі, процедури збору інформації про роботу телекомунікаційної мережі. Також приділено увагу визначенню особливостей процесу статистичного моніторингу характеристик телекомунікаційної мережі. Проаналізовано методи аналізу сезонності та аналізу тренду. Розглянута методологія збору первинної інформації про телекомунікаційної мережі та процедури перетворення первинних значень характеристик телекомунікаційної мережі у часовий ряд і випадкову послідовність, що дозволяють здійснювати безперервний моніторинг характеристик телекомунікаційної мережі статистичними методами.

Ключові слова: модель АРПСС, методи первинної обробки значень параметрів телекомунікаційної мережі, процедури збору інформації про роботу телекомунікаційної мережі, аналіз сезонності, аналіз тренду, оцінювання параметрів.

МЕТОДЫ И ПРОЦЕДУРЫ ПЕРВИЧНОЙ ОБРАБОТКИ ЗНАЧЕНИЙ ПАРАМЕТРОВ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ

В. О. Бреславский

Государственный университет телекоммуникаций

ул. Соломенская, 7-а, г. Киев, 03110, Украина. E-mail: slaava@i.ua

Для решения задач разработки научных основ исследования общих свойств и принципов функционирования телекоммуникационных систем и компьютерных сетей, а также с целью разработки принципиально новых методов анализа, направленных на улучшение их технических характеристик, а также с целью обеспечения контроля и диагностики, рассмотрены методы математического моделирования, основанные на статистических методах анализа многомерных данных. В частности, рассмотрены методы первичной обработки значений параметров телекоммуникационной сети, процедуры сбора информации о работе телекоммуникационной сети. Также уделено внимание определению особенностей процесса статистического мониторинга характеристик телекоммуникационной сети. Проанализированы методы анализа сезонности и анализа тренда. Рассмотрена методология сбора первичной информации о телекоммуникационной сети и процедуры преобразования первичных значений характеристик телекоммуникационной сети во временной ряд и случайную последовательность, позволяющие осуществлять непрерывный мониторинг характеристик телекоммуникационной сети статистическими методами.

Ключевые слова: модель АРПСС, методы первичной обработки значений параметров телекоммуникационной сети, процедуры сбора информации о работе телекоммуникационной сети, анализ сезонности, анализ тренда, оценки параметров.

АКТУАЛЬНІСТЬ РОБОТИ. Однією з важливих задач моніторингу телекомунікаційної мережі є діагностика характеристик мережевого трафіку з метою аналізу стану всієї мережі. У свою чергу, аналіз інформації про те, в якому стані знаходиться мережа в цілому, які поточні процеси, та, можливо, аномалії відбуваються при роботі каналів зв'язку і яким законам вони підпорядковуються, дозволяє перейти до практичної модифікації телекомунікаційної мережі, розподілу навантаження по каналах, запобіганню збоїв у телекомунікаційній мережі. На даний час подібні завдання вирішуються за допомогою різних протоколів для управління телекомунікаційним обладнанням з використанням суб'єктивних методів системного адміністрування.

Мета роботи – розробка методів і процедур первинної обробки значень параметрів телекомунікаційної мережі.

МАТЕРІАЛ І РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ.
Методи первинної обробки значень параметрів

телекомунікаційної мережі. На практиці застосовуються два сімейства стандартів управління мережами стандарти Internet, побудовані на основі протоколу SNMP (Simple Network Management Protocol – простий протокол управління мережею) та міжнародні стандарти ISO/ITU-T, що використовують як протокол у взаємодії агентів і менеджерів протокол CMIP (Common Management Information Protocol – протокол загальної керуючої інформації) [1–3].

Завданням адміністрування сучасної великої телекомунікаційної мережі є не тільки забезпечення «потреб» користувачів, підключення нових робочих станцій і заповнення таблиць маршрутизації, але й одержання інформації про мережу на глобальному рівні з метою накопичення статистики та управління трафіком. Для успішного адміністрування мережі необхідно знати стан кожного її елемента і мати можливість змінювати параметри його функціонування. Зазвичай мережа складається з пристроїв різних виробників і керувати нею було б неможливо.

во, якби кожне з мережевих пристроїв розуміло тільки свою систему команд. Внаслідок цього виникла необхідність у використанні єдиної мови управління мережевими ресурсами, яку б розуміли всі пристрої, і яка, в силу цього, використовувався б усіма пакетами управління мережею для взаємодії з конкретними пристроями.

Подібною мовою стала SNMP-Simple Network Management Protocol, що розроблена для систем, орієнтованих під операційну систему UNIX. Вона стала фактично загальноприйнятим стандартом мережевих систем управління та підтримується переважною більшістю виробників мережевого устаткування в своїх продуктах.

У силу своєї назви – Простий Протокол Мережевого Управління – вона максимально проста в реалізації. Це протокол, що включає мінімальний набір команд, але дозволяє виконувати практично весь спектр завдань управління мережевими пристроями – від отримання інформації про місцезнаходження конкретного пристрою, до можливості виробляти його тестування. Основною концепцією протоколу є те, що вся необхідна для управління пристроєм інформація зберігається на самому пристрої – будь то сервер, модем або маршрутизатор у так званій Адміністративній Базі Даних (MIB – Management Information Base). MIB є набором змінних, що характеризують стан об'єкта управління. Ці змінні можуть відображати такі параметри, як кількість пакетів, оброблених пристроєм, стан його інтерфейсів, час функціонування пристрою і т.п. Кожен виробник мережевого устаткування, крім стандартних змінних, включає в MIB які-небудь параметри, специфічні для даного пристрою. Однак при цьому не порушується принцип представлення та доступу до адміністративної інформації – всі вони будуть змінними в MIB. Тому SNMP як безпосередньо мережевий протокол надає тільки набір команд для роботи з перемінними MIB.

Для того, щоб проконтролювати роботу деякого пристрою мережі, необхідно просто отримати доступ до його MIB, яка постійно оновлюється самим пристроєм, і проаналізувати значення деяких змінних.

Важливою особливістю протоколу SNMP є те, що в ньому не містяться конкретні команди управління пристроєм. Замість визначення всього можливого спектру таких команд, які безумовно переважали б сам протокол, який вважається все-таки простим, визначені змінні MIB, перемикання яких сприймається пристроєм як вказівка виконати деяку команду.

Таким чином, удається зберегти простоту протоколу, але разом з цим зробити його досить потужним засобом, що дає можливість стандартним чином задавати набори команд управління мережевими пристроями. Завдання забезпечення виконання команд складається, таким чином, в реєстрації спеціальних змінних MIB і реакції пристрою на їх зміни.

За своєю структурою MIB є деревом, зображеним на рис. 1.

Кожному елементу відповідає чисельний і символічний ідентифікатор. До імені змінної включається повний шлях до неї від кореневого елемента root.

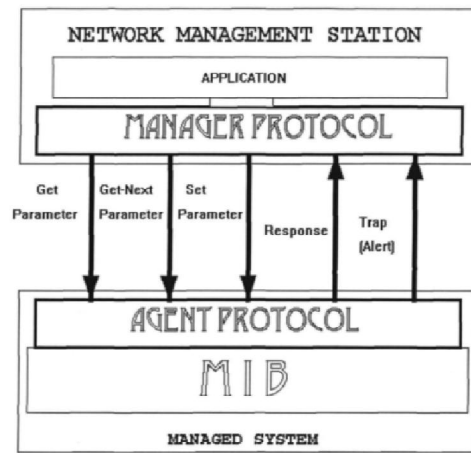


Рисунок 1 – Структура MIB

Наприклад, час роботи пристрою з моменту перезавантаження зберігається у змінній, що знаходиться в розділі system під номером 3 і має назву sysUpTime. Відповідно, ім'я змінної включатиме весь шлях: iso (1) .org (3) .dod (6) .internet (1) .mgmt (2) .mib- 2 (1) .system (1) .sysUpTime (3); або мовою чисел: 1.3.6.1.2.1.1.3. Слід зауважити, що при цьому вузли дерева поділяються точками.

Процедури збору інформації про роботу телекомунікаційної мережі. Для аналізу мережевого трафіку з листопада 2000 року на сервері spb.runnet.ru системним адміністратором мережі RUNNet здійснювався збір даних за допомогою протоколу SNMP. На сервері використовувалися такі технології передачі даних як: ATM 1/0, FastEthernet1 / 0, FastEthernet 4/0 (1 і 4 – умовні номери двох різних каналів, що істотно відрізняються об'ємом завантажень).

Розглянемо технології збору первинної інформації про функціонування телекомунікаційної мережі докладніше [3–5].

Дані про функціонування телекомунікаційної мережі реєструвалися за допомогою чотирьохбайтового лічильника з інтервалом 5 хв. Для аналізу були визначені наступні характеристики:

- завантаження каналу на вході і виході (байт); число пакетів на вході і виході; число помилок в їх реєстрації; загрузка процесора (% від 100);
- обсяг вільної пам'яті процесора і системи введення–виведення для маршрутизатора (байт).

Збір і реєстрація параметрів телекомунікаційної мережі здійснювалися протягом тривалого періоду часу за допомогою чотирьохбайтового лічильника, при переповненні лічильників відбувалося їх обнулення (або скидання). Це призводило до пилкоподібний в уявленні значень параметрів телекомунікаційної мережі і не дозволяло безпосередньо використовувати відомі методи статистичної обробки інформації.

У табл. 1 приведений фрагмент отриманих первинних значень для різних телекомунікаційних технологій.

Таблиця 1 – Фрагмент первинних значень

Час, с	Завантаження каналу інтерфейсу в байтах		
	ATM 1/0	FastEthernet 0/0	FastEthernet 0/0 4/0
972679200	0	0	0
972679500	2295814400	1591956794	4292426841
972679800	2845509618	1847131537	23553383
972680100	3362825219	2065629362	56449704
972680400	3861342029	2248183765	92505895
972680700	44955162	2454219679	127426756
972681000	504745715	2603268793	160571793
972681300	991229953	2768885858	191229552
972681600	1467868509	2943932422	221035042
972681900	1934793959	3118191498	252665332
972682200	2405304536	3277795544	283981824
972682500	2866122697	3445428635	316374364
972682800	3351797421	3634525866	351033268
972683100	3796589541	3820092939	387842305

Пропонуються дві процедури обробки даних, що дозволяють усунути це обмеження. Подальший вибір процедури визначатиметься набором кінцевих характеристик, що відбивають поведінку аналізованих випадкових послідовностей і спостерігається частоти «обнулення». Як видно із наведеної табл. 1, перший стовпець є часом, представленим лічильником і виражений в секундах. Для подальшого аналізу найкраще прив'язати час лічильника до реального часу для з'ясування та адміністратором що відбуваються і подальшої інтерпретації отриманих результатів. Для зручності введемо поняття дискретного моменту часу $t_j = j \Delta$, де Δ інтервал між окремими спостереженнями (у даному випадку Δ одно 5 хв.), а ціле число J змінюється від 1 до N .

Визначення особливостей процесу статистичного моніторингу характеристик телекомунікаційної мережі. За допомогою описаної вище процедури перетворення даних із накопичувального типу у випадкову послідовність були представлені інші характеристики мережевого трафіку. Нижче показані графіки (рис. 2–5) надходження інформації для таких характеристик, як завантаження каналів на вході інтерфейсів ATM, FastEthernet 1/0, FastEthernet 4/0 та завантаження процесора.

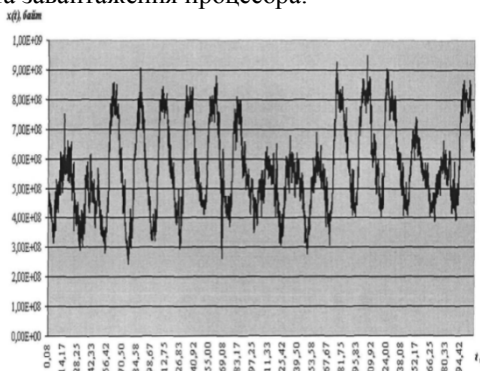


Рисунок 2 – Динаміка завантаження каналу на вході інтерфейсу ATM (кількість інформації, яка прийшла по 5 хв.)

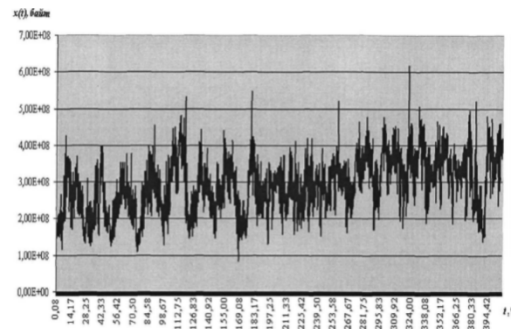


Рисунок 3 – Динаміка завантаження каналу на вході інтерфейсу FastEthernet 1/0 (кількість інформації, яка прийшла по 5 хв.)

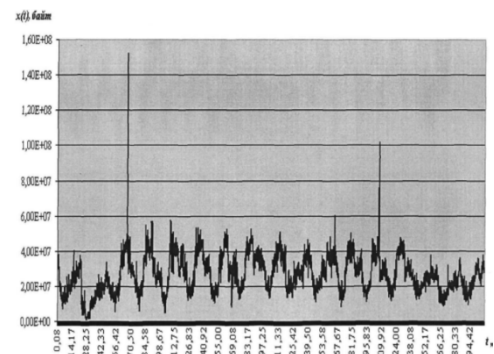


Рисунок 4 – Динаміка завантаження каналу на вході інтерфейсу FastEthernet 4/0 (кількість інформації, яка прийшла по 5 хв.)

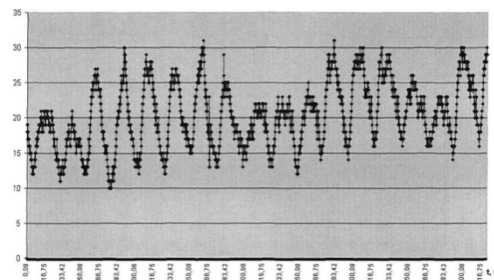


Рисунок 5 – Завантаження процесора

Як видно з приведених графіків (рис. 2–5), можна виділити основні особливості статистичного моніторингу телекомунікаційної мережі – *нестационарність*; *неоднорідність*; *періодичність* (нерівномірність завантаження каналів); складна форма періодичного сигналу.

Форми сигналів ближчі до трапецієподібних з явно вираженим «плато» в області максимальних завантажень; кількість шумів (явно випадуючих точок) – більша при максимальних завантаженнях.

З вищевказаних характеристик телекомунікаційної мережі особливо можна виділити найбільш інформативний параметр – завантаження каналу. Завантаження каналів безпосередньо впливає на стабільність і надійність каналів зв'язку. Надалі під аналізом мережевого трафіку, будемо розуміти обсяг інформації проходить по каналу зв'язку, тобто завантаження каналу за певну одиницю часу завантаження каналів (байт).

Відповідно до технології збору інформації та

природи вимірюваних характеристик, реєстровані значення утворюють випадкові часові ряди з інтервалом дискретизації, рівним 5 хв. Нижче буде показано, що адекватним математичним апаратам для аналізу параметрів функціонування телекомунікаційної мережі є статистичний аналіз часових рядів. Однак через перераховані вище особливостей не можна безпосередньо використовувати класичні методи аналізу часових рядів. Важлива умова застосування цих методів – стаціонарність часового ряду. Як відомо, процес називається стаціонарним, якщо його властивості не змінюються в часі. Зокрема, він має постійне математичне очікування, постійну дисперсію, визначальний розмах його коливань щодо середнього значення [7–9]. Для виконання даної умови необхідно скористатися евристичними методами аналізу.

Моделювання часових рядів. Моделювання різних складових, що характеризують роботу мережі, таких, як обсяг трафіку, кількість утрачених пакетів і ін. [10] у вигляді часових рядів, має ряд очевидних переваг порівняно з вищеописаними способами. При побудові моделі часових рядів використовується експериментальна інформація (отримана в реально функціонуючій мережі), потрібно менше припущень і, отже, більш адекватно відображається реальний об'єкт, тобто телекомунікаційна мережа. Математична модель описує потік інформації залежно від моменту T .

Низка моделей, побудованих на сучасних частково евристичних методах і запропонованих у даній роботі для дослідження та аналізу функціонування телекомунікаційних мереж, буде розглянуто нижче.

Статистичні моделі телекомунікаційних мереж у вигляді часових рядів, на наш погляд, найбільш достовірні, оскільки засновані на великому числі експериментальних даних, а, отже, є найбільш інформативними для прогнозу стану мережі.

Аналіз тренду. Не існує «автоматичного» способу виявлення тренда в часовому ряді. Однак якщо тренд є монотонним (стійко зростає або стійко убыває), то аналізувати такий ряд звичайно неважко. Якщо часові ряди містять значну помилку, то першим кроком виділення тренда є згладжування.

Згладжування завжди включає деякий спосіб локального усереднення даних, при якому несистематичні компоненти взаємно погашають один одного. Найбільший загальним є метод згладжування – ковзне середнє, в якому кожен член ряду замінюється простим або зваженим середнім n сусідніх членів, де n – ширина «вікна» (Бокс і Дженкінс, 1976; Velleman and Hoaglin, 1981). Замість середнього можна використовувати медіану значень, що потрапили до вікна. Основна перевага медіанного згладжування, порівняно зі згладжуванням ковзним середнім, полягає в тому, що результати стають більш стійкими до викидів (наявним всередині вікна). Таким чином, якщо в даних є викиди (пов'язані, наприклад, із помилками вимірів), то згладжування медіаною зазвичай призводить до більш гладких або, принаймні, більш «надійних» кривих, порівняно з ковзаючим середнім із тим же самим вікном. Основний недолік медіанного згладжування в тому, що за відсутності

явних викидів, він призводить до більш «зубчастим» кривих (ніж згладжування ковзним середнім) і не дозволяє використовувати ваги.

Коли помилка виміру дуже велика, використовується метод згладжування методом найменших квадратів, зважених щодо відстані або метод негативного експоненційно зваженого згладжування. Всі ці методи відфільтровують шум і перетворюють дані у відносно гладку криву. Ряди з відносно невеликою кількістю спостережень і систематичним розташуванням точок можуть бути згладженими за допомогою бікубичних сплайнів.

Багато монотонні часові ряди можна добре наблизити лінійною функцією. Якщо ж мається явна монотонна нелінійна компонента, то дані спочатку слід перетворити, щоб усунути нелінійність. Зазвичай для цього використовують логарифмічне, експоненціальне або (менш часто) поліноміальне перетворення даних.

Аналіз сезонності. Періодична і сезонна залежність (сезонність) є іншим загальним типом компонент часового ряду. Можна легко бачити, що кожне спостереження дуже схоже на сусіднє; додатково, мається повторювана сезонна складова. Це означає, що кожне спостереження також схоже на спостереження, що були в тому ж самому місяці рік тому. Загалом, періодична залежність може бути формально визначена як кореляційна залежність порядку k між кожним i -м елементом ряду і $(i-k)$ -м елементом (Kendall, 1976). Її можна виміряти за допомогою автокореляції (тобто кореляції між самими членами ряду); k зазвичай називають лагом (іноді використовують еквівалентні терміни: зрушення, запізнювання). Якщо помилка виміру не надто велика, то сезонність можна визначити візуально, розглядаючи поведінку членів ряду через кожні k тимчасових одиниць.

Сезонні складові часового ряду можуть бути знайдені за допомогою корелограм (автокорелограма), що показує чисельно і графічно автокореляційну функцію (АКФ), іншими словами коефіцієнти автокореляції (та їх стандартні помилки) для послідовності лагів із певного діапазону (наприклад, від 1 до 30). На корелограмі звичайно відзначається діапазон в розмірі двох стандартних помилок на кожному лазі, проте зазвичай величина автокореляції більш цікава, ніж її надійність, тому що інтерес в основному становлять дуже сильні (а, отже, високо значущі) автокореляції.

При вивченні корелограм слід пам'ятати, що автокореляції послідовних лагів формально залежні між собою. Розглянемо наступний приклад. Якщо перший член ряду тісно пов'язаний з другим, а другий з третім, то перший елемент повинен також якимось чином залежати від третього і т.д. Це призводить до того, що періодична залежність може істотно змінитися після видалення автокореляцій першого порядку, тобто після взяття різниці з лагом 1.

Інший корисний метод дослідження періодичності полягає в дослідженні приватної автокореляційної функції (ПАКФ), що становить поглиблення поняття звичайної автокореляційної функції. У ПАКФ усувається залежність між проміжними спостереженнями

(спостереженнями всередині лага). Іншими словами, приватна автокореляція на даному лагу аналогічна звичайній автокореляції, за винятком того, що при обчисленні з неї віддаляється вплив автокореляцій з меншими лагами (Бокс і Дженкінс, 1976; McDowall, McCleary, Meidinger, and Hay, 1980). На лагу 1 (коли немає проміжних елементів усередині лага), приватна автокореляція дорівнює, очевидно, звичайній автокореляції. Насправді, приватна автокореляція дає більш «чисту» картину періодичних залежностей.

Як зазначалося вище, періодична складова для даного лага k може бути видалена взяттям різниці відповідного порядку. Це означає, що з кожного g -го елемента ряду віднімається $(j-k)$ -й елемент. Є два доводи на користь таких перетворень.

По-перше, таким чином можна визначити приховані періодичні складові ряду. Нагадаємо, що автокореляції на послідовних лагах залежні. Тому видалення деяких автокореляцій змінить інші автокореляції, які, можливо, домінували над ними, і зробить деякі інші сезонні складові більш помітними.

По-друге, видалення сезонних складових робить ряд стаціонарним, що необхідно для застосування АРПСС та інших методів, наприклад, спектрального аналізу.

Модель АРПСС. Загальна модель, запропонована Боксом і Дженкінсом (1976) включає як параметри авторегресії, так і параметри змінного середнього. Саме, є три типи параметрів моделі: параметри авторегресії (p), порядок різниці (d), параметри змінного середнього (q). У позначеннях Боксу і Дженкінса модель записується як АРПСС (p, d, q). Як зазначено раніше, для моделі АРПСС необхідно, щоб ряд був стаціонарним, це означає, що його середнє постійно, а вибіркові дисперсія і автокореляція не змінюються в часі.

Тому зазвичай необхідно брати різниці ряду доти, поки він не стане стаціонарним (часто також застосовують логарифмічне перетворення для стабілізації дисперсії). Число різниць, які були взяті, щоб досягти стаціонарності, визначаються параметром d . Для того, щоб визначити необхідний порядок різниці, потрібно досліджувати графік ряду і автокорелограму.

Сильні зміни рівня (сильні скачки вгору або вниз) зазвичай вимагають взяття несезонного різниці першого порядку (лаг=1). Сильні зміни нахилу вимагають взяття різниці другого порядку.

Сезонна складова вимагає взяття відповідної сезонної різниці (см. нижче). Якщо мається повільне спадання вибірових коефіцієнтів автокореляції залежно від лага, зазвичай беруть різницю першого порядку. Однак слід пам'ятати, що для деяких часових рядів потрібно брати різниці невеликого порядку або зовсім не брати їх. Зауважимо, що надмірна кількість взятих різниць призводить до менш стабільним оцінками коефіцієнтів.

На цьому етапі (який зазвичай називають ідентифікацією порядку моделі) необхідно вирішити, як багато параметрів авторегресії (p) і змінного середнього (q) повинна бути присутньою в ефективній і економній моделі процесу. (Економність моделі означає, що в ній є найменше число параметрів і найбільше число ступенів свободи серед всіх моде-

лей, які підганяються до даних). На практиці дуже рідко буває, що число параметрів p або q більше 2.

Наступний, після ідентифікації, крок (Оцінювання) полягає в оцінюванні параметрів моделі (для чого використовуються процедури мінімізації функції втрат. Отримані оцінки параметрів використовуються на останньому етапі {Прогноз} для того, щоб обчислити нові значення ряду і побудувати довірчий інтервал для прогнозу. Процес оцінювання проводиться за перетвореними даними (підданим застосуванню різного оператора). До побудови прогнозу потрібно виконати зворотню операцію (інтегрувати дані). Таким чином, прогноз методології буде порівнюватися з відповідними вихідними даними. На інтегрування даних вказує буква I у загальній назві моделі (АРПСС = авторегресійне проінтегроване змінено середнє).

До того, як почати оцінювання, необхідно вирішити, який тип моделі буде підбиратися до даних, і яка кількість параметрів присутній в моделі, іншими словами, потрібно ідентифікувати модель АРПСС. Основними інструментами ідентифікації порядку моделі є графіки, автокореляційна функція (АКФ), приватна автокореляційна функція (ЧАКФ). Це рішення не є простим, і потрібно ґрунтовно проекспериментувати з альтернативними моделями. Тим не менш, більшість зустрічаються на практиці часових рядів, тому можна з достатнім ступенем точності апроксимувати однією із п'яти основних моделей, які можна ідентифікувати по виду автокореляційної (АКФ) і приватної автокореляційної функції (ЧАКФ). Додаткові практичні методики надано в Hoff (1983), McCleary and Hay (1980), McDowall, McCleary, Meidinger, and Hay (1980), and Vandaele (1983).

Мультиплікативна сезонна АРПСС представляє природний розвиток і узагальнення звичайної моделі АРПСС на ряди, в яких мається періодична сезонна компонента. У доповненні до несезонним параметрам, в модель вводяться сезонні параметри для певного лага (встановлюваного на етапі ідентифікації порядку моделі). Аналогічно параметрам простої моделі АРПСС, ці параметри називаються: сезонна авторегресія (ps), сезонна різниця (ds) і сезонне ковзне середнє (qs). Таким чином, повна сезонна АРПСС може бути записана як АРПСС (p, d, q) (ps, ds, qs). Загальні рекомендації щодо вибору звичайних параметрів (за допомогою АКФ і ЧАКФ) повністю застосовні до сезонних моделями. Основна відмінність полягає в тому, що в сезонних рядах АКФ і ЧАКФ мають суттєві значення на лагах, кратних сезонному лагу (в доповненні до характерного поведінки цих функцій, що описують регулярну (несезон) компоненту АРПСС).

Оцінювання параметрів. Існують різні методи оцінювання параметрів, які дають дуже схожі оцінки, але для даної моделі одні оцінки можуть бути ефективніші, а інші менш ефективні. Загалом, під час оцінювання порядку моделі використовується так званий квазіньютонівський алгоритм максимізації правдоподібності (вірогідності) спостереження значень ряду за значеннями параметрів. Практично це

вимагає обчислення (умовних) сум квадратів залишків моделі. Є різні способи обчислення суми квадратів залишків SS'' ; можна вибрати: (1) наближений метод максимальної правдоподібності Маклеода і Сейлз (1983), (2) наближений метод максимальної правдоподібності з ітераціями назад, (3) точний метод максимального правдоподібності за Мелардом (1984).

Для всіх оцінок параметрів обчислюються так звані асимптотичні стандартні помилки, для обчислення яких використовується матриця приватних похідних другого порядку, що апроксимується кінцевими різницями.

Процедура оцінювання мінімізує (умовну) суму квадратів залишків моделі. Якщо модель не є адекватною, може трапитися так, що оцінки параметрів на якомусь кроці стануть неприйнятними – дуже великими (наприклад, не задовольняють умові стаціонарності). У такому випадку, SS буде приписано дуже велике значення (штрафне значення). Зазвичай це «змушує» ітераційний процес видалити параметри з неприпустимою областю. Проте в деяких випадках і ця стратегія може виявитися невдалою, і ви все одно побачите на екрані (під час процедури оцінювання) дуже великі значення SS на серії ітерацій. В таких випадках слід з обережністю оцінювати придатність моделі. Якщо модель містить багато параметрів і, можливо, має інтервенцію, то слід кілька разів випробувати процес оцінювання з різними початковими даними.

Оцінювання моделі. Якщо значення обчислюється t статистики не значимі, відповідні параметри в більшості випадків видаляються з моделі без збитку підгонки. Іншої звичайної заходом надійності моделі є порівняння прогнозу, побудованого за урізаною рядом з «відомими (вихідними) даними».

Однак якісна модель повинна не лише давати досить точний прогноз, але бути економною і мати незалежні залишки, що містять тільки шум без систематичних компонент (зокрема, АКФ залишків не повинна мати будь-якої періодичності). Тому необхідний всебічний аналіз залишків. Нормальною моделі є: (а) графік залишків та вивчення їх трендів, (б) перевірка АКФ залишків (на графіку АКФ зазвичай чітко видна періодичність).

Якщо залишки систематично розподілені (наприклад, негативні в першій частині ряду і приблизно рівні нулю у другій) або включають деяку періодичну компоненту, то це свідчить про неадекватність моделі. Аналіз залишків надзвичайно важливий і необхідний при аналізі часових рядів. Процедура оцінювання передбачає, що залишки не корельовані і нормально розподілені.

Слід нагадати, що модель АРПСС є підходящою лише для рядів, які є стаціонарними (середнє, дисперсія і автокореляція приблизно постійні в часі); для нестационарних рядів слід брати різниці. Рекомендується мати, як мінімум, 50 спостережень у файлі вихідних даних. Також передбачається, що параметри

моделі постійні, тобто не змінюються в часі.

ВИСНОВКИ.

1. Як інструмент первинного збору статистичної інформації про функціонування телекомунікаційної мережі обрана процедура збору інформації за допомогою протоколу управління SNMP. Сучасні мережні протоколи та засоби управління мережею, доступні її адміністратору, дозволяють акумулювати масиви статистичної інформації з будь-яким заданим інтервалом часу по практично необмеженому числу показників.

2. На основі проведених експериментів зроблено припущення, що для аналізу параметрів функціонування телекомунікаційної мережі доцільно використовувати статистичний аналіз часових рядів.

3. Статистичні моделі телекомунікаційних мереж у вигляді часових рядів, на наш погляд, найбільш достовірні, оскільки ґрунтуються на великому числі експериментальних даних, а, отже, є і найбільш інформативними для прогнозу стану мережі.

ЛІТЕРАТУРА

1. Бугай А.И., Калинина Э.В., Ретинская И.В., Скуратов А.К. Моделирование некоторых характеристик научно-образовательных сетей методами математической статистики // Труды Всероссийской научно-методической конференции «Телематика' 2002», 03–06 июня 2002 г. – Санкт-Петербург: Изд-во СПбГТУ, 2002. – С. 213–214.
2. Бугай А.В., Калинина Е.В., Ретинская В.В., Скуратов А.К. Статистичний аналіз інформаційних потоків у глобальних мережах // Інформаційні технології. – 2002. – № 1. – С. 11–15.
3. Васильев В.Н., Гугель Ю.В., Гуров И.П., Шалаев М.П. Анализ характеристик информационного трафика в компьютерных сетях на основе моделей Марковских процессов // Известия Вузов. Приборостроение. – 2003. – Т. 46, № 8. – С. 19–24.
4. Васильев П.М., Иванов В.В., Кореньков В.В. и др. Система сбора, анализа и управления сетевым трафиком фрагмента сети ОИЯД на примере подсети университета «Дубна». – Дубна, 2001. – 11 с.
5. Вегешна Ш. Качество обслуживания в сетях IP. – М.: Издательство «Вильямс», 2003. – 368 с.
6. Вильям Столингс. Структурная организация и архитектура компьютерных систем. – М.: Издательство «Вильямс», 2002. – 368 с.
7. Дженкинс Р., Ваттс Д. Спектральный анализ и его применение. – М.: Мир, 1972. – Вып. 2. – 287 с.
8. Домарев В.В. Безопасность информационных технологий: Методология создания систем защиты. – К.: ООО «Тид «ДС», 2001. – 688 с.
9. Домрачев В.Р., Безрукавный Д.С., Калинина Е.В. и др. Нечеткие методы в задачах мониторинга сетевого трафика // Інформаційні технології. – 2006. – № 3. – С. 2–10.
10. Ефремова М.Р., Петрова Е.В., Румянцев С.Н. Общая теория статистики. – М.: ИНФА–М, 1998. – 416 с.

**METHODS AND PROCEDURES OF THE PRIMARY VALUES PROCESSING
OF THE TELECOMMUNICATION NETWORK PARAMETERS**

V. Breslavskiy

State University of Telecommunications

vul. Solomenskaya, 7-a, Kyiv, 03110, Ukraine. E-mail: slaava@i.ua

Paper reviews the methods of mathematical modeling, which are based on statistical analysis of multidimensional data, to solve the problems of development of scientific bases for research of common properties and principles of functioning of telecommunication systems and computer networks, and to develop a fundamentally new analysis methods aimed at improving their performance, and to ensure monitoring and diagnostics. In particular, they were considered: methods of the primary processing of telecommunications network parameter values, the procedures of data collection from the telecommunication network. Also, it is described the definition of the statistical monitoring of the telecommunications network characteristics. The methods of seasonality analysis and trend analysis are analyzed. It is researched the methodology of collecting primary information and procedures of the conversion of primary values of telecommunication network in time series and random sequence, which allow to make the continuous monitoring of telecommunications network characteristics by statistical methods.

Key words: Model of ARPSS, primary processing, parameters, telecommunications network, data collection, analysis of seasonality, trend analysis, parameters estimation.

REFERENCES

1. Bugay, A.V., Kalinina, E.V., Retinskaya, V.V., Skuratov, A.K. (2002), "Modeling of some characteristics of scientific and educational networks, methods of mathematical statistics", *Pratsi Vserosiys'koyi naukovo - metodychnoyi konferentsiyi "Telematyka-2002"* [Proceedings of the All-Russian scientific-technical conference "Telematics 2002"], Saint – Petersburg, June 3-6, 2002, pp. 213–214.
2. Bugay, A.V., Kalinina, E.V., Retinskaya, V.V., Skuratov, A.K. (2002), "Statistical analysis of the data flows in global networks", *Information Technologies*, no. 1, pp. 11–15.
3. Vasilev, V.N., Gugel, Yu.V., Gurov, V.P., Shalaev, M.P. (2003), "Analysis of characteristics of traffic information in computer networks based on models of Markov processes", *Pratsi Vserosiys'koyi naukovo - metodychnoyi konferentsiyi "Telematyka-2003"* [Proceedings of the All-Russian scientific-technical conference "Telematics 2003"], Saint – Petersburg, April 14-17, 2003, pp. 19–24.
4. Vasiliev, P.M., Ivanov, V.V., Korenkov, V.V. (2001), *Systema zboru, analizu ta upravlinnya merezhevym trafikom frahmenta merezhi OIYaD na prykladi pidmerzhi universytetu "Dubna"* [Data collection, analysis and management of network traffic of JINR fragment on the example university subnet], Dubna, Russia.
5. Vegashna, Sh. (2003), *Yakist obsluhovuvannya v merezhakh IP* [Quality of service in IP networks], Vilyams, Moscow, Russia.
6. Stolings, V. (2002), *Strukturna orhanizatsiya ta arkhitektura kompyuternykh system. Yzd.5* [Structural organization and architecture of computer systems. 5th Ed.], Vilyams, Moscow, Russia.
7. Jenkins, R., Vatta, D. (1972), *Spektralnyy analiz ta yoho zastosuvannya* [Spectral analysis and its application]. Svit, Moscow, USSR.
8. Domarev, S. (2001), *Bezpeka informatsiynykh tekhnolohiy. Metodolohiya stvorenniya system zakhystu* [Safety of information technology. Methodology of creation of protection systems], OOO «Tid «DS», Kyev, Ukraine.
9. Domrachev, V., R., Bezrukavniy, D.S., Kalinina, E.V., Retinska, V.V., Skuratov, A.K. (2006), "Fuzzy methods in problems of monitoring network traffic", *Information Technologies*, no. 3, pp. 2–10.
10. Efremova, M.R., Petrova, E.V., Rummyantsev, S.N. (2000), *Zahalna teoriya statystyky* [General Theory of Statistics], INFA–M, Moscow, Russia.

Стаття надійшла 20.03.2015.