

ОЦІНКА ПОКАЗНИКІВ ЯКОСТІ NGN-МЕРЕЖ З УРАХУВАННЯМ ФРАКТАЛЬНОСТІ ВХІДНОГО ТРАФІКУ

Воропаєва В.Я.

Донецький національний технічний університет, м. Донецьк
кафедра автоматики і телекомунікацій

E-mail: voropayeva@meta.ua

Abstract

Voropaeva V.Y. Estimation of NGN-network QoS indexes for Fractal Input Traffic. The author analyzes main NGN-network QoS indexes taking into account fractal nature of input traffic. The self-similar (fractal) random process is defined, a model of system with self-similar input traffic is described, a possibility of fractal input traffic prognostication is grounded and the connection between input traffic prognostication and NGN-network QoS indexes is found.

Загальна постановка проблеми. З точки зору сучасних тенденцій розвитку телекомунікацій актуальною задачею є побудова багатофункціональної конвергентної мережі на платформі NGN (Next Generation Network) – мультисервісної мережі наступного покоління. NGN має забезпечувати необмежений набір послуг, надавати гнучкі можливості з управління, персоналізації та створення нових видів сервісу за рахунок уніфікації мережевих рішень. Останнє потребує реалізації універсальної транспортної мережі з розподіленою комутацією, де взаємодія між пристроями та додатками здійснюється за допомогою створення віртуальних з'єднань, на управління якими помітно впливають особливості стохастичної динаміки процесів пакетної комутації.

Однією з проблем розрахунку імовірісно-часових характеристик мереж, що активно досліджується останнім часом, є необхідність адекватного врахування реального характеру мережевого трафіку. Класичні підходи теорії телетрафіку базуються на припущенні, що вхідні потоки є стаціонарними пуасонівськими, тобто являють собою суперпозицію великої кількості незалежних стаціонарних ординарних потоків без післядії рівномірно малої інтенсивності [1]. Але аналіз результатів останніх теоретичних досліджень та експериментів з вимірювання характеристик трафіку інформаційних мереж демонструє наявність кореляції довжин інтервалів часу як для моментів надходження запитів в систему, так і для моментів закінчення обслуговування, що є ознакою наявності післядії. Таким чином, для створення адекватних математичних моделей функціонування пакетних мереж необхідно враховувати не пуасонівський характер випадкових процесів.

Ці проблеми досліджувалися багатьма фахівцями. Зокрема Вальтер Вілінгер, Роберт Шерман та Даніель Вільсон шляхом статистичного аналізу локальних мереж довели наявність властивості самоподібності (або фрактальності) трафіку Ethernet-мереж [2]. Російські дослідники в галузі інформаційних мереж В. С. Заборовський та О. Я. Городецький прогнозували стан віртуальних з'єднань протоколу TCP/IP, використовуючи апарат фрактального броуновського руху [3] для моделювання процесів обслуговування TCP-пакетів. Але проблема впливу фрактальних властивостей вхідного трафіку на показники якості мультисервісних NGN мереж лишилась не вирішеною.

Постановка задач дослідження. Для оцінки залежності показників якості мультисервісних NGN мереж від впливу фрактальних властивостей вхідного трафіку, необхідно вирішити наступні основні задачі:

- виділити основні показники якості NGN мереж для різних типів вхідного трафіку;
- проаналізувати фрактальну природу вхідного трафіку NGN мереж;

- дослідити вплив фрактальних властивостей вхідного трафіку на основні показники якості NGN мереж;
- проаналізувати можливість прогнозування самоподібного трафіку.

Рішення задач та результати дослідження. В рекомендаціях МСЕ-Т Е.800 якість послуг визначається як сумарний ефект показників якості послуги, який визначає ступінь задоволеності користувачів послуги [4]. Користувачі потребують таких рівнів показників якості, які в поєднанні з їхніми хостами, кінцевим обладнанням та іншими пристроями забезпечують задовільну підтримку їхніх додатків. Ключові параметри QoS – наскрізна затримка, варіація затримки (джиттер) та рівень втрат. Під наскрізною затримкою розуміють суму затримок у різних мережних пристроях і мережних каналах, через які проходить мультимедійний трафік. Варіація затримки (джиттер) - це зміна часу прибуття між пакетами. Для усунення тремтіння потрібно збирати пакети в буферах і втримувати їх протягом достатнього часу, для того щоб самі повільні пакети прибули вчасно й відтворилися в правильній послідовності. Потоків додатки можуть використовувати інформацію про загальне значення варіації затримки для усунення незаповнення й переповнення буфера. Збільшення варіації затримки може привести до того, що протокол TCP збільшить значення таймера очікування відповіді, повтор передачі пакетів буде відкладений або ж навпаки відбудеться непотрібна ретрансляція пакетів.

Рівень втрат безпосередньо впливає на якість. Втрата інформації не обмежується бітовими помилками і втратами пакетів при передачі, але також включає погіршення, що виникли при кодуванні медіа компонентів (наприклад, при використанні низькошвидкісних кодеків для голосу).

На усі зазначені параметри суттєво впливає алгоритм використання спільних ресурсів, наприклад пропускної здатності магістральних каналів. В мережах з розділенням каналів за кожним додатком закріплюється певна частина такого ресурсу – частота або час залежно від того, яка система мультиплексування використовується. В пакетних мережах використовується статистичний розподіл пропускної здатності.

Залежно від співвідношень між трьома розглянутими параметрами розрізняють п'ять класів обслуговування [4]. Найбільш вимогливий до показників якості 0 клас, він використовується для трафіку реального часу, чутливого до тремтіння, з підвищеним ступенем взаємодії (наприклад VoIP, VTC) і містить наступні вимоги:

- верхнє обмеження значення затримки $TD < 100$ мс;
- верхнє обмеження варіації затримки $DV < 50$ мс;
- верхнє обмеження для імовірності втрати пакета $LR < 1 \cdot 10^{-3}$;
- верхнє обмеження для імовірності помилки $ER < 1 \cdot 10^{-4}$;

Інші типи додатків (наприклад, передача даних) мають менш жорсткі вимоги до зазначених показників. Тому досить ефективним засобом забезпечення якості для одночасного обслуговування багатьох додатків такого типу виявляється алгоритм статистичного мультиплексування, який широко використовується в мережах пакетної комутації для економного витрачання пропускної здатності магістральних каналів.

Є два основних способи управління інтенсивністю та згладжування профілю мережевого трафіку при статистичному мультиплексуванні: шейпінг (Traffic Shaping) і полісінг (Traffic Policing). При алгоритмі шейпінга забезпечується постійна інтенсивність транспортування трафіку шляхом постановки в чергу (буферизації) пакетів, інтенсивність генерації яких перевищує задане значення. Тоді як полісінг передбачає відкидання пакетів, інтенсивність яких вище узгодженої. Звичайно, шейпінг виглядає більш привабливим, оскільки не відкидає пакети, отже задовольняє вимозі стосовно LR. Але додаткове внесення затримок, пов'язаних з буферизацією, може негативно вплинути на параметри TD і DV. Алгоритм полісінга може вимагати значно підвищити пропускну здатність каналу з метою досягнення прийнятних значень LR для трафіка з високою пачечністю. Нагадаємо, що коефіцієнт пачечності дорівнює

відношенню пікової інтенсивності надходження заявок на обслуговування до її середнього значення.

NGN як мережа з комутацією пакетів використовує моделі систем з очікуванням для опису мережних пристроїв. При цьому основні розрахунки виконані для класичної моделі M/M/v/W [1]. Ця модель має низку припущень та спрощень: по-перше, довжина черги не обмежується, що звісно не виконується для будь-якого мережевого обладнання. По-друге, час обслуговування пакетів пристроями комутації або маршрутизації в зазначеній моделі вважається експоненціально-розподіленим. Як показано в [5], при широкому спектрі діючих в мережі протоколів слід приймати довільний закон розподілу часу обслуговування, тобто використовувати модель M/G/v/W.

По-третє, вхідний потік вважається найпростішим з властивостями стаціонарності, ординарності та відсутності післядії. Зазначені властивості забезпечуються, якщо потік формується безкінечною кількістю джерел, причому в кожний момент часу активується не більше одного джерела, а інтервал між двома сусідніми пакетами потоку є випадкова величина з експоненціальним розподілом. Такий підхід справедливий для трафіку даних, який характеризувався певним рівнем пачечності, але не вимагав мінімальних затримок, на відміну від мультимедійного трафіку, пачечність якого значно вища на фоні вимог до обслуговування в режимі реального часу.

Ці припущення дозволяли використовувати марківські випадкові процеси для моделювання таких систем. Але вимірювання та дослідження мультимедійного трафіку на мережах пакетної комутації [6] свідчать про самоподібну його природу, яка суттєво погіршує показники якості обслуговування порівняно з пуасонівськими потоками. На якісному рівні самоподібність проявляється в наявності повільно убуваючої залежності між параметрами трафіку (зокрема, дисперсії довжини кадрів або коефіцієнта автокореляції) в різні моменти часу. Мультимедійний трафік подібно трафіку даних має високий рівень пачечності, причому ці пачки даних статистично подібні в широкому діапазоні зміни масштабу часу.

Дамо визначення самоподібного випадкового процесу. Беручи до уваги пакетний характер аналізованих мереж, будемо розглядати процеси з дискретним часом та дискретними станами.

Нехай $X = (X_1, X_1, \dots)$ – напівбезкінечний відрізок стаціонарного дійсного випадкового процесу дискретного аргументу (часу) $t = \{1, 2, \dots\}$. Введемо додатково 2 процеси, пов'язані з X : $X^{(m)} = (X_1^{(m)}, X_2^{(m)}, \dots)$ – усереднений за блоками довжиною m агрегований процес X та $\hat{X}^{(m)} = (\hat{X}_1^{(m)}, \hat{X}_2^{(m)}, \dots)$ – нормований агрегований процес. Компоненти цих процесів визначаються співвідношеннями:

$$\begin{aligned} X_t^{(m)} &= \frac{1}{m} (X_{t-m+1} + \dots + X_t), \quad m, t \in N; \\ \hat{X}_t^{(m)} &= \frac{1}{m^H} (X_{t-m+1} + \dots + X_t), \quad H = 1 - \frac{\beta}{2}, \quad 0 < \beta < 1. \end{aligned} \quad (1)$$

Позначимо через $\mu, \sigma^2, r(k), b(k)$ та $\mu_m, \sigma_m^2, r_m(k), b_m(k)$ середнє, дисперсію, коефіцієнт автокореляції та автоковаріації процесів X і $X^{(m)}$ відповідно. Тобто

$$\begin{aligned} r(k) &= \frac{M(X_{t+k} - \mu)(X_t - \mu)}{\sigma^2}, \\ b(k) &= \sigma^2 r(k), \quad k = \{0, 1, 2, \dots\}. \end{aligned} \quad (2)$$

Середнє μ , дисперсія $\sigma^2 = b(0)$, коефіцієнт автокореляції $r(k)$ та автоковаріації $b(k)$ не залежать від часу t і $r(k) = r(-k)$, $b(k) = b(-k)$. Процес X називається самоподібним з параметром $H = 1 - (\beta/2)$, $0 < \beta < 1$, якщо його коефіцієнт автокореляції

$$r(k) = \frac{1}{2}[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}] = g(k), k \in N, \quad (3)$$

де функція $g(k) = \frac{1}{2}\delta^2(k^{2-\beta})$ виражена через центральний різнісний оператор 2-го порядку $\delta^2(f(x))$, що діє на функцію $f(x) = x^{2-\beta}$ таким чином:

$$\delta(f(x)) = f(x + \frac{1}{2}) - f(x - \frac{1}{2}). \quad (4)$$

Самоподібність проявляється в тому, що для процесу X , який задовольняє умові (3) виконується рівність $r_m(k) = r(k)$, тобто в такому процесі не змінюється коефіцієнт автокореляції після усереднення за блоками будь-якої довжини m . Отже, для самоподібного процесу статистичні характеристики другого порядку нормованого агрегованого процесу $X^{(m)}$ не відрізняються від характеристик вихідного процесу X при значному інтервалі змін m [6].

Параметр H , або коефіцієнт Херста, має принципове значення в теорії самоподібних процесів. Він є індикатором ступеню самоподібності процесу, а також свідчить про наявність в нього таких властивостей як персистентність/антиперсистентність і тривала пам'ять. Для марківських процесів (без післядії) коефіцієнт Херста дорівнює 0,5. При $0 < H < 0,5$ процесу властива антиперсистентність: високі значення процесу йдуть за низькими і навпаки. Тобто, імовірність того, що на $k+1$ кроці процес відхилиться від середнього в протилежному напрямі (відносно відхилення на k кроці), настільки велика, наскільки параметр H близький до 0.

У випадку $0,5 < H < 1$ процес є персистентним або з тривалою пам'яттю: якщо протягом якогось часу в минулому спостерігалось збільшення параметрів процесу, то і в майбутньому в середньому буде відбуватися їх зростання. Інакше кажучи, імовірність того, що на $k+1$ кроці процес відхилиться від середнього в тому ж напрямі, що і на k кроці, настільки велика, наскільки параметр H близький до 1. Персистентні стохастичні процеси виявляють чітко виражені тенденції змін при відносно малому "шумі". За даними більшості експериментальних досліджень значення параметру H для мережевого трафіку знаходиться в інтервалі $(0,5; 1)$ [6].

Зі зростанням k функція $g(k)$ з (3) асимптотично убуває за ступеневим законом:

$$g(k) \rightarrow \frac{1}{2}(2-\beta)(1-\beta)k^{-\beta}, k \rightarrow \infty, \quad (5)$$

тобто для самоподібних процесів наявне більш повільне убування автокореляційної залежності порівняно з експоненціальним убуванням, властивим марківським процесам.

Не адекватність експоненціального розподілу для трафіку сучасних телекомунікаційних мереж доведена численими дослідженнями. Зокрема встановлено [6], що для опису вхідного потоку кадрів бездротової технології доступу стандарту IEEE 802.11b доцільно використовувати розподіли "з важким хвостом" (РВХ), для яких виконується $P(Z < x) = 1 - cx^{-a}$, $x \rightarrow \infty$, де $0 < a < 2$ – параметр форми, c – деяка додатня константа. На відміну від розподілів "з легкими хвостами" (РЛХ), наприклад, експоненціального чи гаусівського, які мають швидке убування хвоста за експонентою, хвости РВХ зменшуються за ступеневим (гіперболічним) законом.

Основна властивість випадкової величини (наприклад, кількості пакетів Z , що надходять на мережевий пристрій за певний відрізок часу), що розподілена за РВХ, полягає у її високій мінливості. Тобто вибірка з РВХ являє собою здебільшого відносно невеликі значення, але також

містить і суттєву кількість дуже великих значень. При $a \rightarrow 1$ вплив важкого хвоста відбивається на вибірці спадом сходимості вибіркового середнього до математичного очікування, отже, статистичні оцінки втрачають властивість незміщеності.

Найчастіше серед РВХ для мережевих характеристик використовується розподіл Парето:

$$P(Z \leq x) = 1 - \left(\frac{b}{x}\right)^a, b \leq x. \tag{6}$$

Для розподілу (6) вище зазначені особливості призводять до значної недооцінки математичного очікування по вибірковому середньому. Фактично модуль помилки

$$\left| \bar{Z}_m - M(Z) \right| \rightarrow m^{1/a-1}, \tag{7}$$

де m – обсяг вибірки,

\bar{Z}_m – вибіркоче середнє випадкової величини Z , розраховане за вибіркою обсягом m ,

$M(Z)$ – математичне очікування Z .

Таким чином, при значенні параметру форми $a \rightarrow 1$ слід враховувати (7), роблячи висновки про параметри якості NGN мереж. Експериментальне свідчення суттєвого впливу фрактальних властивостей вхідного трафіку на середній час очікування пакету в черзі отримано в роботі [7]. Це один з основних показників якості системи з очікуванням, що визначає в результаті наскрізний час затримки пакета TD при проходженні через послідовність мережних пристроїв, кожний з яких моделюється системою M/G/1/W. Проведене порівняння систем M/G/1/W та M/M/1/W на основі імітаційного моделювання трафіку з різними значеннями коефіцієнту Херста свідчить про значне зниження реального часу очікування у випадку використання моделі M/M/1/W (рис.1). На графіку наведена залежність середнього часу затримки в черзі пакетів на вході маршрутизатора (в логарифмічному масштабі) від завантаженості системи.



Рисунок 1. – Залежність середнього часу очікування пакету для систем M/G/1/W та M/M/1/W від коефіцієнту використання каналу

Як видно з отриманого графіка, у випадку самоподібного трафіку середній час очікування в черзі більше порівняно з випадком найпростішого вхідного потоку, причому ця різниця стає кардинальною у випадку наближення завантаженості системи до 1. Саме в цій зоні суттєво впливає на час очікування і коефіцієнт Херста: при збільшенні H суттєво зростає середній час очікування.

Слід зазначити, що система з очікуванням має “гірші” показники якості порівняно із системою з втратами навіть у випадку найпростішого вхідного потоку. Так, для забезпечення однакового рівня якості (наприклад, рівності імовірності втрати та імовірності очікування) системі М/М/ν/∞ потрібно більше каналів, ніж системі М/М/ν/∞. Тобто, при однаковому вхідному навантаженню і кількості каналів, перша формула Ерланга (імовірність втрат) дає менші значення порівняно з другою (імовірність очікування). Це пояснюється тим, що в системі з втратами втрачені виклики не впливають на подальше обслуговування нових викликів. А в системі М/М/ν/∞ виклики, що потрапили в чергу, уповільнюють обслуговування нових. У випадку ж самоподібного трафіку, що має велике значення пачечності, такий вплив стає значно суттєвішим.

Ці особливості вхідного потоку для сучасних мереж з одного боку вимагають врахування фрактальних властивостей, а з іншого дозволяють використовувати ці властивості для прогнозу трафіку і відповідного налаштування параметрів мережних пристроїв.

Розглянемо прогнозованість деякої випадкової величини з РВХ. Нехай Z - випадкова величина з РВХ, що інтерпретується як час життя (тривалість) мережевого з'єднання (ТСР-сесії, наприклад). Припустимо, що з'єднання було активне протягом деякого часу $\tau > 0$. Будемо вважати час дискретним та введемо індикаторну функцію $A(t)$ таку, що $A(t)=1$ тільки при $Z(t) \geq t$. Тоді умовна імовірність існування в майбутньому з'єднання, що підтримувало активність на інтервалі часу $1 \leq t \leq \tau$, буде оцінюватися як:

$$P\{A(\tau+1)=1 | A(t)=1\} = 1 - \frac{P\{Z = \tau\}}{P\{Z \geq \tau\}}. \tag{8}$$

Обчислимо спочатку (8) для РЛХ, зокрема для експоненціального розподілу $P\{Z < x\} = 1 - c_1 e^{-c_2 x}$, де $c_1, c_2 > 0$ - константи. Тоді другий член у (8) має вигляд:

$$\frac{P\{Z = \tau\}}{P\{Z \geq \tau\}} = \frac{c_1 e^{-c_2 \tau} - c_1 e^{-c_2(\tau+1)}}{c_1 e^{-c_2 \tau}} = 1 - e^{-c_2}. \tag{9}$$

Тобто для експоненціального розподілу тривалість активності τ не впливає на прогноз. Для РВХ аналогічні виклидки приводять до

$$\frac{P\{Z = \tau\}}{P\{Z \geq \tau\}} = \frac{c \tau^{-a} - c(\tau+1)^{-a}}{c \tau^{-a}} = 1 - \left(\frac{\tau}{\tau+1}\right)^a. \tag{10}$$

З (10) видно, що при $\tau \rightarrow \infty$ умовна імовірність існування в майбутньому з'єднання прагне до 1, тобто збільшення періоду спостережаної активності підвищує імовірність існування з'єднання в майбутньому. Таким чином, процес є персистентним і при достатньо великих τ помилка прогнозу може бути наскільки завгодно малою.

Прогнозованість трафіку обумовлена залежністю середньої ентропії трафіку від довжини інтервалу часу, на якому він досліджується. За даною залежністю можна оцінити час, необхідний для спостереження трафіку, перед тим, як зробити висновок про його майбутнє з прийнятною достовірністю.

Достовірно зроблений прогноз вхідного трафіку дозволить виконати динамічний перерозподіл пропускної здатності спільного ресурсу (магістрального каналу), тим самим уникнув додаткових затримок через буферизацію пакетів (при використанні алгоритму шейпінга) або неприпустимих втрат пакетів (при використанні алгоритму полісінга). Тим самим будуть дотримані основні характеристики якості обслуговування.

Висновки.

1. Доведено, що мультимедійний трафік реального часу є найбільш вимогливим до показників якості NGN мереж серед інших типів вхідного трафіку внаслідок чутливості до тремтіння.
2. Запропонований підхід до динамічного розподілу спільного мережевого ресурсу на базі прогнозу вхідного трафіку з урахуванням його фрактальних особливостей на відміну від алгоритмів статистичного мультиплексування (шейпінг і полінг)
3. Вперше запропоновано використання системи масового обслуговування G/M/v/W для моделювання мережних пристроїв NGN мереж.
4. Дана кількісна оцінка коефіцієнту Херста для різних мережевих технологій. Встановлено, що фрактальний вхідний трафік суттєво погіршує середній час очікування порівняно з найпростішим потоком, особливо при збільшенні коефіцієнту Херста.
5. Проаналізовані фактори, що обумовлюють прогнозованість трафіку, та доведена можливість прогнозування персистентних процесів.

Література

1. Крылов В.В., Самохвалова С.С. Теория телетрафика и ее приложения. – СПб.:БХВ-Петербург, 2005. – 288с.
2. Walter Willinger, Murad S. Taqqu, Robert Sherman and Daniel V. Wilson Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level.: IEEE/ACM Transactions on Networking, Vol. 5, No. 1, 1997.
3. Городецкий А.Я., Заборовский В.С. Информатика. Фрактальные процессы в компьютерных сетях: Учеб. пособие / СПб.: Изд-во СПбГТУ, 2000. 102 с.
4. ITU-T Recommendation Y.1541. Network performance objectives for IPbased services. Geneva: International Telecommunication Union. - 2006.
5. Воропаева В.Я, Верховський Я.М., Федосеева О.С. Порівняльний аналіз характеристик якості пакетних мереж. Матеріали XIV міжнародної конференції з автоматичного управління (Автоматика-2007), м. Севастополь, 10-14 вересня 2007 року. – Ч.2 – Севастополь: СНУЯЄтаП, 2007. – С. 107-108.
6. Структура телетрафика и алгоритм обеспечения качества обслуживания при влиянии эффекта самоподобия. Петров В. В. Диссертация на соискание ученой степени кандидата технических наук. Москва, 2004.
7. Влияние самоподобности речевого трафика на качество обслуживания в телекоммуникационных сетях. Осин А.В. Автореферат диссертации на соискание ученой степени кандидата технических наук. Москва, 2005.