

УДК 621.3

С.Г. Семенов¹, Р.В. Королев², А.В. Петров²¹ *Национальный технический университет «ХПИ», Харьков*² *Харьковский университет Воздушных Сил им. Ивана Кожедуба, Харьков*

ОЦЕНКА СТАТИСТИЧЕСКИХ СВОЙСТВ ИНФОРМАЦИОННОГО ТРАФИКА НА ОСНОВЕ МЕТОДА НОРМИРОВАННОГО РАЗМАХА

Проведен анализ методик оценки статистических характеристик информационно-телекоммуникационных систем. Оценены статистические свойства числовых последовательностей с помощью метода нормированного размаха. Построены графики зависимости показателя нормированного размаха случайной величины от числа значений временного ряда. Проведена оценка статистических особенностей информационного трафика в информационно-телекоммуникационных системах. Сделаны выводы о наличии долгосрочных зависимостей в большинстве примеров информационного трафика.

Ключевые слова: информационный трафик, фрактальные свойства, показатель Херста.

Введение

Постановка проблемы. В современных условиях развития информационных технологий появляются и быстро эволюционируют новые подходы к математическому моделированию информационно-телекоммуникационных систем (ИТКС). Это дает основание разработчикам непрерывно улучшать и внедрять в существующую информационно-телекоммуникационную инфраструктуру новые аппаратные и программные решения.

В настоящее время в связи с расширением спектра подходов к математическому моделированию, появлением новых математических моделей ИТКС, зачастую противоречащих известным методам и методикам, вопрос оценки предлагаемых подходов к моделированию и адекватности разработанных математических моделей приобрел особую актуальность. Одной из составляющих процесса моделирования ИТКС является анализ статистических особенностей поведения

ее отдельных подсистем в различных режимах функционирования. Проведенные исследования показали, что оценку статистических особенностей процессов, протекающих в ИТКС, следует проводить на начальном этапе моделирования, что позволит повысить достоверность получаемых результатов.

Анализ литературы [1 – 4, 6, 8] показал, что существует множество методик оценки статистических характеристик объектов управления. До недавнего времени большинство из них строились на основе центральной предельной теоремы [5, 6] и предположении о том, что изучаемая система является или случайной или детерминированной. Однако, как показали исследования ряда авторов [2, 7], существуют системы, в которых протекающие процессы обладают одновременно свойствами случайности и детерминизма. При определенных условиях (использование гетерогенных сетевых технологий, циркуляция мультисервисного трафика и др.) эта особенность может наблюдаться и в ИТКС.

Для изучения таких систем и создания более общей аналитической структуры целесообразнее использовать подходы, основанные на непараметрической статистике. В одном из таких подходов ряд авторов [2] предлагает использовать метод нормированного размаха (R/S анализ).

Изложение основного материала

1. Оценка статистических свойств числовых последовательностей с помощью метода нормированного размаха. Проведенный анализ показал, что основная особенность данного метода заключается в выдвижении гипотезы о зависимости поведения отдельных подсистем во времени. В этом случае гипотеза, выдвинутая Энштейном в 1908 году [5], которая заключается в том, что расстояние R, которое проходит случайная частица, увеличивается пропорционально квадратному корню от времени T, используемому для его измерения:

$$R = T^{0,50}, \tag{1}$$

претерпевает ряд изменений.

Эти изменения подчеркнул в своих трудах гидролог Херст. Он отметил, что для применения концепции Энштейна к временному ряду, сформированному системой, не находящейся в броуновском движении, необходимо обобщить уравнение (1) и учитывать постулат о зависимости многих внутренних процессов в ряде сложных систем.

Результатом данной гипотезы стало уравнение

$$(R/S)_t = c * t^H, \tag{2}$$

где R/S – нормированный размах случайной величины; t – число значений временного ряда; c – константа; H – показатель Херста.

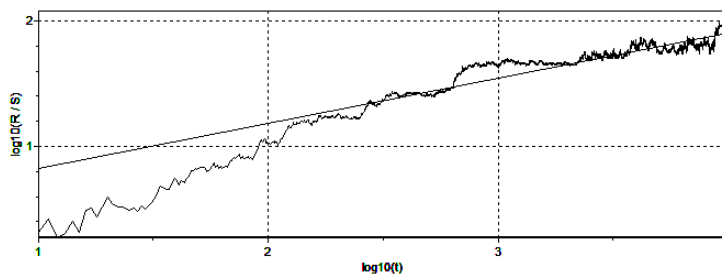
Как показали исследования [7], показатель Херста может быть приближен посредством построения зависимости $\log(R/S)_t$ от $\log(t)$ и вычисления наклона через простую регрессию методом наименьших квадратов [2]:

$$\log(R/S)_t = \log(c) + H \log(t). \tag{3}$$

Оценим данную методику на примерах различных псевдослучайных числовых последовательностей, полученных с помощью известных алгоритмов генерирования (RC-4, ГОСТ 28147-89, Галуа, суммирования). Из [6] (методика NIST STS) известно, что числовые последовательности, полученные генераторами RC-4 и ГОСТ 28147-89, обладают лучшими автокорреляционными свойствами по сравнению с последовательностями, полученными на основе генераторов Галуа и суммирования.

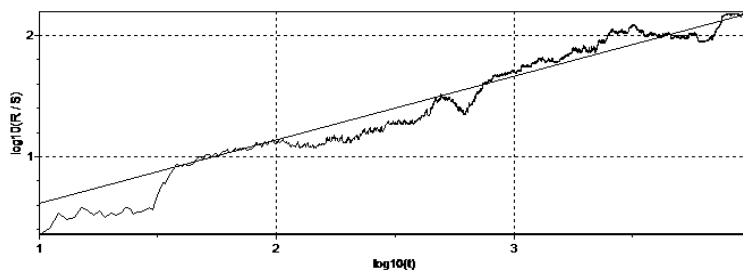
Оценим фрактальные свойства данных числовых последовательностей и проведем их общую сравнительную оценку.

На рис. 1 – 4 представлены результаты исследований и расчетов показателя Херста, а также графики зависимости $\log(R/S)_t$ от $\log(t)$ для различных числовых последовательностей. Анализ результатов моделирования и расчетов показал явление персистентности Херста ($H > 0,5$) для временных рядов, сформированных с помощью алгоритма Галуа и суммирования, а также антиперсистентности Херста ($H < 0,5$) для временных рядов, сформированных с помощью алгоритма RC-4. Кроме того, результаты моделирования показали, что значение показателя Херста для временных рядов, сформированных с помощью алгоритма ГОСТ 28147-89, близко к 0,5. Это говорит о независимости внутренних процессов в подобных системах.



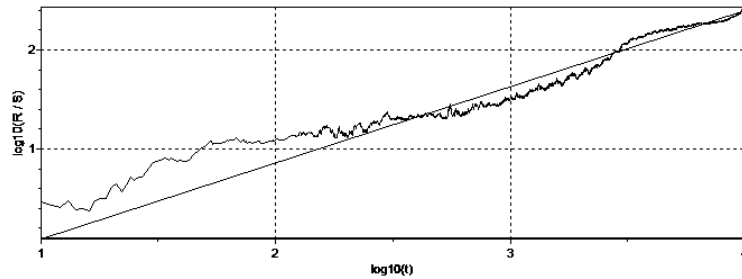
Показатель Херста $\hat{H} = 0,3582 \pm 0,1575$, Фрактальная размерность $D = 1,6418 \pm 0,1575$

Рис. 1. Кривая зависимости $\log(R/S)_n$ от $\log(n)$ для числового ряда, сформированного с помощью алгоритма RC-4



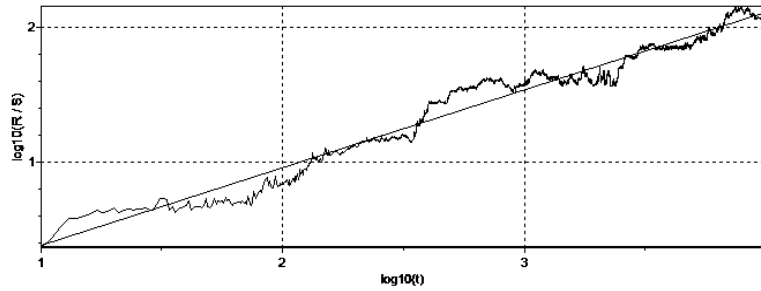
Показатель Херста $\hat{H} = 0,5279 \pm 0,1776$, Фрактальная размерность $D = 1,4721 \pm 0,1776$

Рис. 2. Кривая зависимости $\log(R/S)_n$ от $\log(n)$ для числового ряда, сформированного с помощью алгоритма ГОСТ 28147-89



Показатель Херста $\hat{I} = 0,7707 \pm 0,1643$, Фрактальная размерность $D = 1,2293 \pm 0,1643$

Рис. 3. Кривая зависимости $\log(R/S)_n$ от $\log(n)$ для числового ряда, сформированного с помощью алгоритма Галуа



Показатель Херста $\hat{I} = 0,5776 \pm 0,1482$, Фрактальная размерность $D = 1,4224 \pm 0,1482$

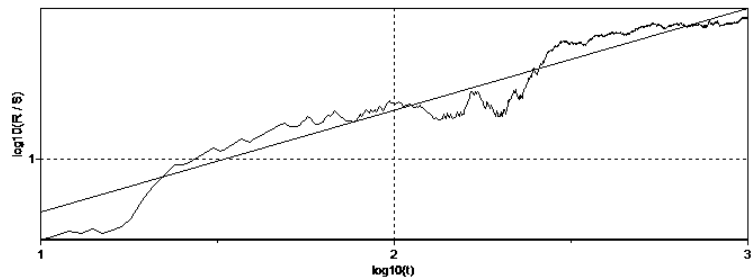
Рис. 4. Кривая зависимости $\log(R/S)_n$ от $\log(n)$ для числового ряда, сформированного с помощью алгоритма суммирования

В целом результаты исследования фрактальных свойств исследуемых временных рядов подтвердили сведения, полученные с помощью методики статистического тестирования NIST STS. Результаты данного примера позволяют сделать вывод о возможности и, в ряде случаев, целесообразности использования метода нормированного размаха для оценки статистических особенностей в поведении ИТКС.

2. Оценка статистических особенностей информационного трафика в ИТКС. Для оценки

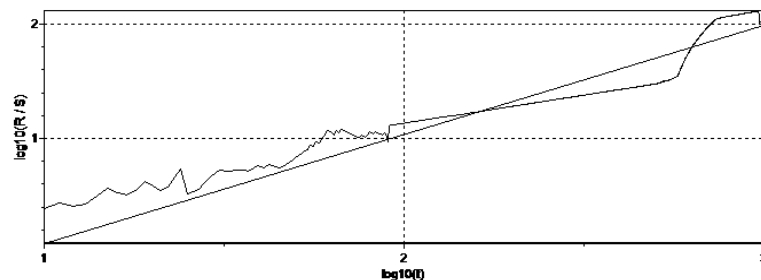
статистических свойств информационного трафика воспользуемся данными эксперимента, проведенного с помощью анализатора трафика Wireshark, описанного в работах [1, 3].

Результаты расчетов показателя Херста, а также кривые зависимости $\log(R/S)_t$ от $\log(t)$ оценки информационного трафика различных сетевых услуг (IP-телефонии, FTP-трафика, мультисервисного трафика) представлены на рис. 5 – 7.



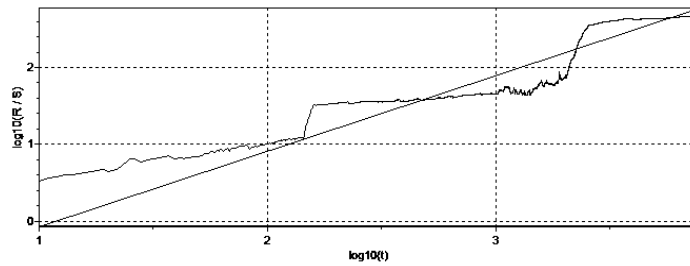
Показатель Херста $\hat{I} = 0,6311 \pm 0,1769$, Фрактальная размерность $D = 1,3689 \pm 0,1769$

Рис. 5. Кривая зависимости $\log(R/S)_n$ от $\log(n)$ оценки информационного трафика IP-телефонии



Показатель Херста $\hat{I} = 0,9543 \pm 0,3412$, Фрактальная размерность $D = 1,0457 \pm 0,3412$

Рис. 6. Кривая зависимости $\log(R/S)_n$ от $\log(n)$ оценки FTP-трафика



Показатель Херста $\hat{I} = 0,9819 \pm 0,4016$, Фрактальная размерность $D = 1,0181 \pm 0,4016$

Рис. 7. Кривая зависимости $\log(R/S)_n$ от $\log(n)$ оценки мультисервисного трафика

Анализ графиков рис. 5 – 7 показывает наличие долговременных зависимостей в статистических данных информационного трафика практически всех сетевых сервисов. Так, показатель Херста для всех служб $H > 0,5$. А наибольшее значение показателя Херста достигает в статистических данных мультисервисного трафика.

Выводы

Таким образом, анализ методики оценки статистических характеристик временных рядов на основе метода нормированного размаха показал возможность и целесообразность ее использования для оценки статистических свойств информационного трафика. Проведенные исследования статистических свойств трафика показали наличие долговременных зависимостей в большинстве из них. Дальнейшие исследования ИТКС должны быть направлены на анализ и оценку возможных закономерностей в поведении информационного трафика, а также поиска возможных артефактов в условиях внешних воздействий на телекоммуникационную систему.

Список литературы

1. Кузнецов А.А. Метод структурной идентификации информационных потоков в телекоммуникационных сетях на основе BDS-тестирования / А.А. Кузнецов, С.Г. Семенов, С.Н. Симоненко, Е.В. Мелешко // *Наука і техніка Повітряних Сил Збройних Сил України: науково-технічний журнал*. – 2010. – № 2 (4). – С. 131-137.

2. Петерс Э. Фрактальный анализ финансовых рынков: Применение теории хаоса в инвестициях и экономике / Э. Петерс. – М.: Интернет-Трейдинг, 2004. – 304 с.

3. Семенов С.Г. Процедура идентификации информационного трафика на основе N-мерного шкалирования / С.Г. Семенов, И.В. Ильина, Е.В. Мелешко // *Системи управління, навігації та зв'язку*. – К.: ДП «ЦНДІ навігації і управління» 2011. – Вип. 1(17). – С. 198-201.

4. Шелухин О.И. Фрактальные процессы в телекоммуникациях / О.И. Шелухин, А.М. Тенякиев, А.В. Осин. – М.: Радиотехника, 2003. – 480 с.

5. Эйнштейн А. Собрание научных трудов в четырех томах (Академия наук СССР. "Классики естествознания"), под ред. И.Е. Тамма, Я.А. Смородинского, Б.Г. Кузнецова. Том I. Работы по теории относительности 1905-1920 / А. Эйнштейн. – М.: Наука, 1965. – 700 с.

6. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. NIST Special Publication 800-22. Technology Administration U.S. Department of Commerce. – Washington: National Institute of Standards and Technology, 2000. – 164 p.

7. Hurst H.E. Long-term Storage of Reservoirs / H.E. Hurst. *Transactions of the American Society of Civil Engineers* 116, 1951.

8. Brock W. A test for independence based on correlation dimension / W. Brock, W. Dechert, J. Scheinkman, B. LeBaron // *Econometric Reviews* 15: 197-235, 1996.

Поступила в редколлегию 8.11.2011

Рецензент: д-р техн. наук, проф. В.А. Краснобаев, Полтавский национальный технический университет им. Юрия Кондратюка, Полтава.

ОЦІНКА СТАТИСТИЧНИХ ВЛАСТИВОСТЕЙ ІНФОРМАЦІЙНОГО ТРАФІКУ НА ОСНОВІ МЕТОДУ НОРМОВАНОГО РОЗМАХУ

С.Г. Семенов, Р.В. Королев, О.В. Петров

Проведений аналіз методик оцінки статистичних характеристик інформаційно-телекомунікаційних систем. Оцінені статистичні властивості числових послідовностей за допомогою методу нормованого розмаху. Побудовані графіки залежності показника нормованого розмаху випадкової величини від числа значень часового ряду. Проведена оцінка статистичних особливостей інформаційного трафіку в інформаційно-телекомунікаційних системах. Зроблені висновки про наявність довготривалих залежностей в більшості прикладів інформаційного трафіку.

Ключові слова: інформаційний трафік, фрактальні властивості, показник Херста.

ESTIMATION OF STATISTICAL PROPERTIES OF INFORMATIVE TRAFFIC ON BASIS OF METHOD OF THE RATIONED SCOPE

S.G. Semenov, R.V. Korolyov, A.V. Petrov

The analysis of methods of estimation of statistical descriptions of the informatively-telecommunication systems is conducted. Statistical properties of numerical sequences are appraised by the method of the rationed scope. The graphs of dependence of index of the rationed scope of casual size are built from the number of values of temporal row. The estimation of statistical features of informative traffic is conducted in the informatively-telecommunication systems. Conclusions are done about the presence of long duration dependences in most examples of informative traffic.

Keywords: informative traffic, fractal properties, Kherst index.