

УДК 621.372.061.3.001.63:681.3

DOI: 10.15587/1729-4061.2014.28026

У роботі пропонуються підхід, що дозволяє провести оцінку поведінки мережевого трафіка високошвидкісних мереж зв'язку, який має самоподібні властивості з погляду методів нелінійної динаміки. Дана методика дає можливість подальшого прогнозування поведінки мережевого трафіка на більш тривалих інтервалах часової осі шляхом його аналізу на відносно невеликих відрізках

Ключові слова: самоподоба, мережевий трафік, хаотичні явища, динамічна система, протокол TCP

В работе предлагается подход, который позволяет провести оценку поведения сетевого трафика высокоскоростных сетей связи, обладающего самоподобными свойствами с точки зрения методов нелинейной динамики. Данная методика дает возможность дальнейшего прогнозирования поведения сетевого трафика на более длительных интервалах временной оси, посредством его анализа на относительно небольших отрезках

Ключевые слова: самоподобие, сетевой трафик, хаотические явления, динамическая система, протокол TCP

ПРИМЕНЕНИЕ МЕТОДОВ НЕЛИНЕЙНОЙ ДИНАМИКИ ДЛЯ ИССЛЕДОВАНИЯ ПОВЕДЕНИЯ СЕТЕВОГО ТРАФИКА ВЫСОКОСРОСТНЫХ СЕТЕЙ

А. В. Карпужин

Кандидат технических наук, доцент, ведущий научный сотрудник
Кафедра прикладной математики

Харьковский национальный университет радиоэлектроники
пр. Ленина, 14, г. Харьков, Украина, 61126

E-mail: kav-102@yandex.ru

Д. И. Грицив

Аспирант

Кафедра информационных технологий
в энергетических системах

Харьковский национальный университет им. В. Н. Каразина
пл. Свободы, 4, г. Харьков, Украина, 61022

E-mail: dgritsiv@gmail.com

А. А. Ткаченко

Заместитель директора по вопросам маркетинга и продажи услуг
ХФ ПАО "Укртелеком"

ул. Иванова, 7/9, г. Харьков, Украина, 61002

E-mail: oatkachenko@ukrelecom.ua

1. Введение

В настоящее время в компьютерных сетях достаточно остро стоят проблемы обеспечения качества обслуживания (QoS), расчета пропускной способности сети и моделирования поведения трафика. Как было показано в многочисленных исследованиях, сетевой трафик компьютерной сети обладает свойством самоподобия. Самоподобный трафик выглядит качественно одинаково при почти любых масштабах временной оси. Использование при проектировании компьютерных сетей классических методов, основанных на модели Пуассона при выборе таких параметров, как пропускная способность каналов связи или емкость памяти буфера промежуточных маршрутизаторов, могут дать неоправданно оптимистичные результаты использования полосы пропускания и привести к недооценке нагрузки, а, следовательно, к заторам (congestions). Предотвращение заторов – в основном проблема предсказания и включает в себя обнаружение того факта, когда затор неизбежен, и выполнение действий, которые его предотвращают. Восстановление – это действие, предпринятое сетью после того,

как ухудшение работы обнаружено для того, чтобы как можно более уменьшить влияние затора. Чтобы знать, какие действия необходимо предпринять для восстановления, важно знать среднее время существования затора и какие виды потерь пакетов (или задержки) возникают во время заторов.

2. Анализ литературных данных и постановка проблемы

Сетевой трафик, обладающий свойством самоподобия [1–4], носит взрывной характер, т. е. при относительно небольшом среднем уровне трафика возможны очень большие выбросы (bursts) [5, 6]. Это явление значительно ухудшает характеристики качества при прохождении самоподобного трафика через сеть [7] даже в тех случаях, когда средняя интенсивность трафика намного ниже потенциально достижимой скорости передачи в данном канале, что приводит к увеличению времени задержки и потере пакетов. Потери пакетов, выражающиеся в долях процента, при больших скоростях передачи данных приводят к значительным

потерям информации. Основной причиной самоподобия трафика является TCP [8–10] (основной транспортный протокол Internet). В приведенных выше работах предлагаются различные способы улучшения производительности работы компьютерной сети. В данной работе предлагается методика исследования поведения информационных сетей связи посредством применения, как статистических методов, так и методов нелинейной динамики или теории хаоса.

3. Цель и задачи исследования

Целью данной работы является исследование методами нелинейной динамики реализаций реального сетевого трафика, направленное на выявление его характерных особенностей.

Для достижения поставленной цели решались следующие задачи:

- Проанализировать сетевой трафик компьютерной сети для прогнозирования его поведения на более длительных интервалах времени, используя метод детрендированного флуктуационного анализа.

- Рассмотреть модельную сеть как динамическую систему и проанализировать её работу по наблюдаемому трафику на наличие нежелательных хаотических явлений, применяя методы нелинейной динамики. Такой подход позволяет дать рекомендации для проектирования или реинжиниринга компьютерных сетей.

4. Метод исследования модельной сети

В данной работе для исследования сетевого трафика была создана модельная сеть TCP/IP (рис. 1), в которой все хосты связаны с маршрутизаторами соединением типа точка-точка. Для соединения хостов с маршрутизаторами (Cisco Catalyst 2960G-48) использовалась витая пара в качестве среды передачи данных. Таким образом, максимальная пропускная способность данного участка сети достигала $C_b = 100$ Mbps. В свою очередь, для соединения маршрутизаторов между собой использовалось оптоволокно. Хост 4 был соединен с маршрутизатором посредством сетевого коммутатора, что добавляет ещё одно потенциально узкое место в модельной сети. Другие хосты были соединены с маршрутизаторами напрямую.

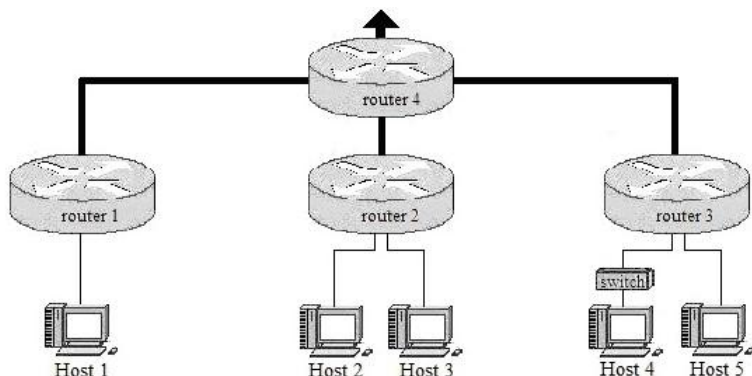


Рис. 1. Топология модельной сети

Для того чтобы проводить исследование сетевых приложений и протоколов, а также чтобы находить проблемы в работе сети, и, что важно, выяснять причины этих проблем, необходим такой инструмент, который бы позволял отображать трафик в детальном виде. Для регистрации сетевого трафика использовалось специальное программное обеспечение Wireshark [11], которое устанавливалось на всех хостах, и которое захватывало входящий и исходящий трафик с сетевой карты и записывало данные в режиме реального времени. Результатом его работы является сохраненные дампы с данными на каждом хосте. Захват данных производился несколько раз в течение рабочего сеанса пользователей на протяжении 20 дней. Время захвата сетевого трафика на хостах составляло 3600 секунд, что вполне достаточно, чтобы проанализировать сетевую активность данной части сети в конкретный период времени. Wireshark работает на большинстве современных ОС (Microsoft Windows, Mac OS X, UNIX) и является продуктом с открытым исходным кодом, распространяемым на основании лицензии GPL. Используя Wireshark, можно достаточно эффективно находить и диагностировать разнообразные проблемы, возникающие в сети.

В процессе захвата трафика в сети отслеживалось значение множества переменных для каждого хоста, поэтому полученные отчеты с данными фильтровались по следующим критериям: IP адрес хоста и TCP протокол передачи данных. Для дальнейшего анализа временных рядов [12] необходимо было провести процедуру агрегирования для преобразования исходных рядов $\psi = \{\psi(t_1), \psi(t_2), \dots, \psi(t_n)\}$ в эквидистантные, которые обладают постоянным шагом Δt по оси времени. Параметр Δt можно обозначить как степень агрегации. Для этого был сформирован новый ряд, полученный при помощи операции суммирования каждого исходного значения информации (TCP трафик) соответствующим заданным временным интервалом Δt . Следовательно, агрегированные значения преобразованного ряда можно представить в таком виде:

$$X_N = \sum_{i=1+(N-1)\Delta t}^{N\Delta t} \psi(t_i). \quad (1)$$

В результате работы алгоритма получается агрегированная эквидистантная реализация трафика $X = \{X(\Delta t), X(2\Delta t), \dots, X(N\Delta t)\}$, которая содержит N элементов. В процессе агрегации временных рядов выбирались разные временные интервалы $\Delta t = 0.1, 0.5, 1$. Ниже приведены графики (рис. 2, а, б и рис. 3, а, б) агрегированного трафика для одинакового уровня агрегации $\Delta t = 0.1$ для различных хостов модельной сети.

Как можно видеть на рис. 2, а, б и рис. 3, а, б наблюдается периодичность работы TCP протокола и, как и ожидалось, сетевой трафик носит взрывной характер на разных временных участках, а при разном шаге агрегирования временной ряд сохраняет свойство инвариантности.

Для оценивания параметра N самоподобных временных рядов использовался метод детрендированного флуктуационного анализа (ДФА) [13, 14].

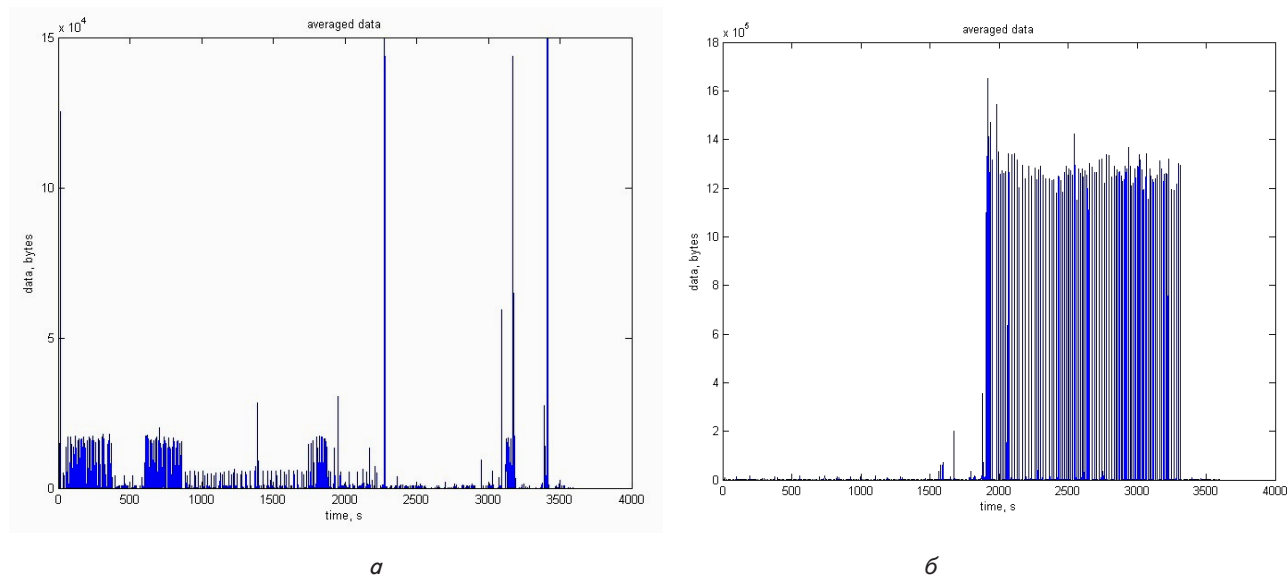


Рис. 2. Агрегированная реализация сетевого трафика при $\Delta t=0,1$: а – для хоста 1 ($H=0,651$); б – для хоста 3 ($H=0,653$)

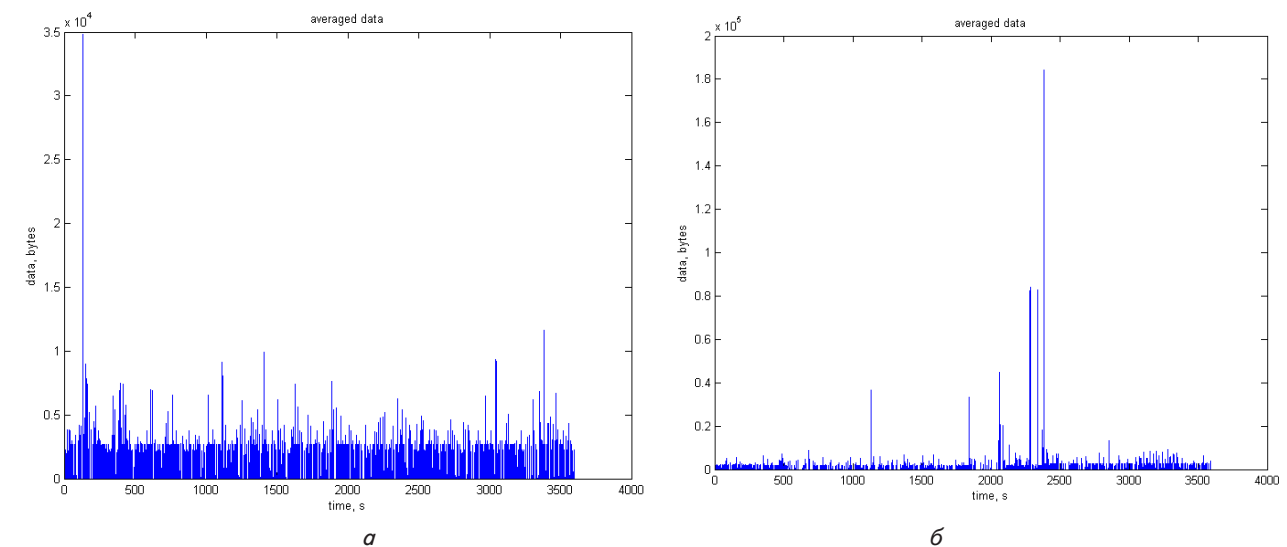


Рис. 3. Агрегированная реализация сетевого трафика при $\Delta t=0,1$: а – для хоста 4 ($H=0,512$); б – для хоста 5 ($H=0,654$)

Авторами было проведено исследование мультифрактальных характеристик агрегированных временных рядов. Значение вычисленных показателей Херста для реализаций агрегированного трафика для хостов рассматриваемой модельной сети при одинаковом уровне агрегации $\Delta t = 0,1$.

Максимальный показатель Ляпунова (Maximum Lyapunov Exponent) – величина, которая характеризует скорость разбегания близких траекторий, положительное значение которой обычно принимается как индикатор хаотического поведения системы. Вычисление максимального показателя Ляпунова производилось с помощью пакета утилит TISEAN [15], который предназначен для анализа временных рядов и основан на теории нелинейных детерминированных динамических систем или теории хаоса [16, 17]. Результатом работы TISEAN является набор данных, представляющих собой зависимость логарифма

коэффициента разбегания траекторий от времени – $S(\epsilon, m, \Delta n)$, который вычисляется следующим образом

$$S(\epsilon, m, \Delta n) = \frac{1}{N} \sum_{n_0=1}^N \ln \left(\frac{1}{|U(S_{n_0})|} \times \sum_{S_n \in U(S_{n_0})} |S_{n_0+\Delta n} - S_{n+\Delta n}| \right), \quad (2)$$

где ϵ – окрестность точки S_{n_0} ; m – размерность фазового пространства; Δn – время, а $U(S_{n_0})$ – окрестность точки S_{n_0} диаметра ϵ .

Результат вычислений максимального показателя Ляпунова для хостов компьютерной сети с помощью программы TISEAN приведен ниже на рис. 4, а, б и рис. 5, а, б для уровня агрегации $\Delta t = 0,1$ временных рядов.

Так как вычисленный максимальный показатель Ляпунова для исследуемых хостов положительный, то система находится в хаотическом состоянии, что негативно влияет на производительность модельной сети.

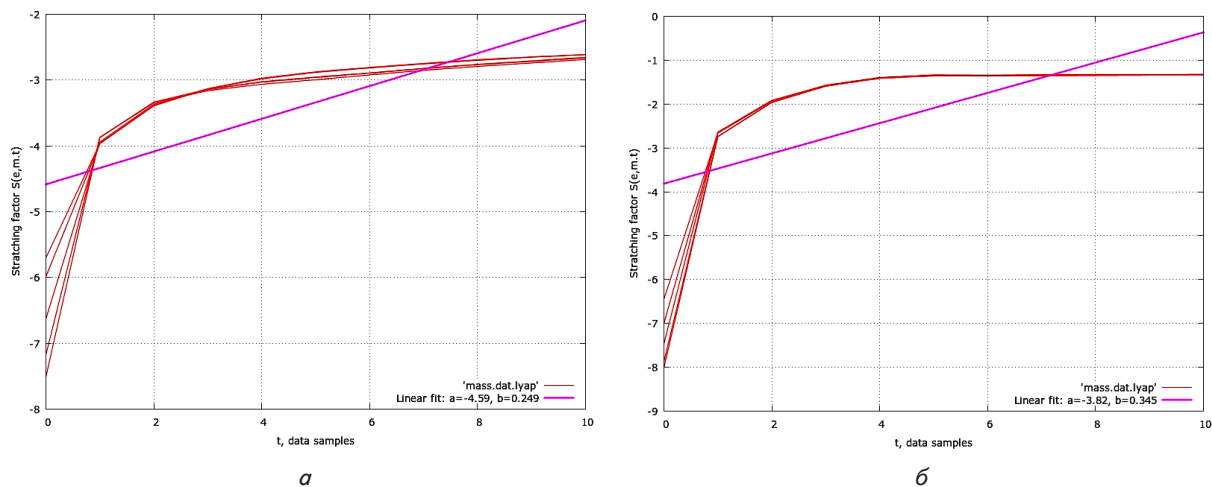


Рис. 4. Вычисление показателя Ляпунова при $\Delta t=0,1$: а – для хоста 1 ($\lambda \sim 0,249$); б – для хоста 3 ($\lambda \sim 0,345$)

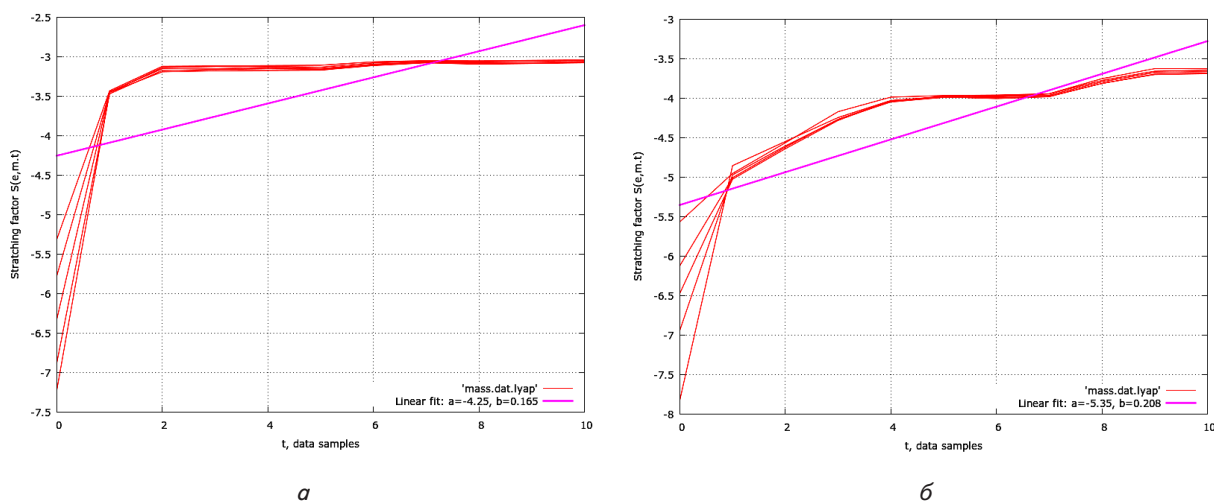


Рис. 5. Вычисление показателя Ляпунова при $\Delta t=0,1$: а – хоста 4 ($\lambda \sim 0,165$); б – для хоста 5 ($\lambda \sim 0,208$)

5. Выводы

В результате исследования были вычислены показатель Херста и максимальный показатель Ляпунова, что дает возможность оценку работы компьютерной сети по наблюдаемым временным рядам. На протяжении периода наблюдения за модельной сетью, показатель Херста H практически во всех случаях превышает значение $0,5$, что говорит о долгосрочной зависимости исследуемых рядов. Вычисленный максимальный показатель Ляпунова положительный и меняется в диапазоне $0,165 - 0,345$, что свидетельствует о наличии хаоса в рассматриваемой модельной сети. Как видно из представленных результатов, в наблюдаемой модельной сети присутствуют нежелательные хаотические яв-

ления, которые ухудшают производительность системы. Таким образом, данный подход дает возможность провести оценку влияния самоподобных процессов в современных компьютерных сетях, которые имеют долгосрочную зависимость, позволяя предсказать будущее поведение, обладая данными о недавнем прошлом.

В глобальном масштабе всей сети Internet решить проблему заторов и потерь пакетов, очевидно, не представляется возможным в связи с тем, что перестроить всю сеть нельзя в силу технических и экономических причин. Однако, в ограниченных по размерам сетях (даже довольно больших) возможно дать рекомендации по проектированию (и дальнейшей эксплуатации) таких сетей, которые позволят свести к минимуму отрицательные явления хаотизации.

Литература

1. Sheluchin, O. I. Self-Similar Processes in Telecommunications [Text] / O. I. Sheluchin, S. M. Smolskiy, A. V. Osin. – New York: John Wiley & Sons, 2007. – 320 p.
2. Willinger, W. Self-similarity through high-variability: Statistical analysis of Ethernet LAN traffic at the source level [Text] / W. Willinger, M.S. Taqqu, R. Sherman, D.V. Wilson // IEEE/ACM Trans. Netw. – 2007. – Vol. 5, Issue 1. – P. 71–86. doi: 10.1109/90.554723

3. Leland, W. E. On the self-similar nature of ethernet traffic [Text] / W. E. Leland, M. S. Taqqu, W. Willinger, D. V. Wilson // IEEE/ACM Transactions of Networking. – 1994. – Vol. 2, Issue 1. – P. 10–15. doi: 10.1145/166237.166255
4. Park, K. Self-similar network traffic: An overview. In: Self-similar network traffic and performance evaluation. Chapter 1 [Text] / K. Park, W. Willinger // Eds. New York: Wiley, 2000. — 19 p. doi: 10.1002/047120644x.ch1
5. Guillemin, F. Burstiness in broadband integrated networks [Text] / F. Guillemin, J. Boyer, A. Dupuis // Performance Evaluation. – 1992. – Vol. 15, Issue 3. – P. 163–176. doi: 10.1016/0166-5316(92)90032-c
6. Hanaya, Y. S. High-performance implementation of in-network traffic pacing for small-buffer networks [Text] / Y. S. Hanaya, A. Dwaraki, K. Huc, T. Wolf // Computer Communications. – 2013. – Vol. 36, Issue 13. – P. 1450–1459. doi: 10.1016/j.comcom.2013.07.002
7. Larsson, C. Chapter 8 – Flow-Controlled Packet Networks. Design of Modern Communication Networks [Text] / C. Larsson. – Methods and Applications, 2014. – P. 237–271. doi: 10.1016/B978-0-12-407238-1.00008-7
8. Карпухин, А. В. Особенности реализации протокола TCP в современных компьютерных сетях [Текст] / А. В. Карпухин // Системы обработки информации. – 2009. – Изд. 6, Вып. 80. – С. 49–53.
9. Feng, W. The failure of TCP in High-Performance Computational Grids [Text] / W. Feng, P. Tinnakornsriruphap // In Proceedings of International Conference on Parallel Processing (ICPP'00), 2000. – Article №. 37.
10. Feng, W. The Adverse Impact of the TCP Congestion-Control Mechanism in Distributed Systems [Text] / W. Feng, P. Tinnakornsriruphap // In Proceedings ICPP'00 of International Conference on Parallel Processing, 2000. – P. 299–306. doi: 10.1109/icpp.2000.876145
11. Пакет программ Wireshark и сопутствующая документация [Электронный ресурс] / Режим доступа: <http://www.wireshark.org/>
12. Петров, В. В. Статистический анализ сетевого трафика [Текст] / В. В. Петров. – МЭИ, ИРЭ, Москва, 2003. – 47 с.
13. Kantelhardt, J. W. Multifractal detrended fluctuation analysis of non-stationary time series [Text] / J. W. Kantelhardt, S. A. Zschiegner, A. Bunde, S. Havlin, E. Koscielny-Bunde, H.E. Stanley // Physica A. – 2002 – Vol. 316. – P. 87–114. doi: 10.1016/S0378-4371(02)01383-3
14. Кириченко, Л. О. Исследование выборочных характеристик, полученных методом мультифрактального флуктуационного анализа [Текст] / Л. О. Кириченко // Вестник НТУУ «КПИ». Информатика, управление и вычислительная техника. – 2011. – Вып 54. – С. 101–110.
15. Пакет программ TISEAN и сопутствующая документация [Электронный ресурс] / Режим доступа: <http://www.mpiikpks-dresden.mpg.de/~tisean/>
16. Kantz, H. Nonlinear Time Series Analysis, 2nd edition [Text] / H. Kantz and T. Schreiber. – Cambridge University Press, Cambridge, 2003. – 388 p.
17. Packard, N. H. Geometry from a Time Series [Text] / N. H. Packard, J. P. Crutchfield, J. D. Farmer, R. S. Shaw // Physical Review Letters. – 1980. – Vol. 45, Issue 9. – P. 712–716. doi: 10.1103/PhysRevLett.45.712