

ОПРЕДЕЛЕНИЕ ХАРАКТЕРИСТИК СЕТЕВОГО ТРАФИКА КОМПЬЮТЕРА

М.К. Бойченко, И.П. Иванов, В.А. Лохтуров

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация
e-mail: noc@bmstu.ru; ivanov@bmstu.ru

Приведены результаты экспериментальных исследований самоподобия сетевого трафика исследуемого компьютера. Для определения параметра Херста найден масштабируемый диапазон последовательности случайных величин времени двойного оборота кадров, посылаемых от компьютера-источника к компьютеру-приемнику без промежуточных транзитных узлов сети между ними. Аналогичные исследования выполнены для потока кадров, посылаемых по адресу loopback, т.е. потока, поступающего на вход сетевого адаптера. Альтернативным для метода вычисления значения масштабируемого диапазона является метод агрегирования, используемый при определении дисперсии самоподобного процесса. На основе анализа полученных значений параметра Херста сделан вывод о самоподобии сетевого трафика компьютера.

Ключевые слова: компьютерная сеть, сетевой адаптер, самоподобие, сетевой трафик, параметр Херста, масштабируемый диапазон, агрегирование.

DEFINITION OF CHARACTERISTICS FOR COMPUTER'S NETWORK TRAFFIC

M.K. Boychenko, I.P. Ivanov, V.A. Lokhturov

Bauman Moscow State Technical University, Moscow, Russian Federation
e-mail: noc@bmstu.ru; ivanov@bmstu.ru

The results of experimental studies of self-similarity of the investigated computer network traffic, are presented. To determine the Hurst parameter, scalable sequence range value of random variables time double turnover is calculated for frames sent from the source computer to the destination computer without intermediate transit nodes between them. Similar investigations are carried out for the flow of frames sent to the address loopback, i.e. stream entering the network adapter input. An alternative way to calculate the scaled range value is an aggregation method used in determining the self-similar process dispersion. On the basis of analysis of the obtained Hurst exponent values calculated by two methods, it is concluded that the computer network traffic is self-similar.

Keywords: computer network, network interface card, self-similarity, network traffic, Hurst parameter, scaled range, aggregation.

Характер сетевого трафика во многом определяет характеристики буферной памяти транзитных и оконечных узлов компьютерных сетей. Толчком к исследованию характера трафика в различных сетях послужила публикация статьи Норроса [1], в которой автор привел результаты аналитического исследования бесконечного буфера с постоянным временем обслуживания, на вход которого поступает поток информации, смоделированный процессом дробного броуновского движения. Среди прочих результатов приведена зависимость размера буферной

памяти q от среднего коэффициента пропускной способности ρ (коэффициента нагрузки) буфера в виде

$$q = \left(\frac{\rho^{1/2}}{(1 - \rho)^H} \right)^{\frac{1}{1-H}}, \quad (1)$$

где H — параметр Херста (Hurst), введенный для анализа самоподобия исследуемых случайных процессов [2]. При $H = 0,5$ формула (1) дает классический результат теории массового обслуживания для систем с пуассоновским распределением интервалов между поступлением заявок и экспоненциальной длительностью обработки заявок [3].

Вместе с тем, анализ экспериментальных результатов многочисленных исследований, проводимых в различных локальных и глобальных сетях, показал значительное их отличие от теоретических результатов, полученных на основе классических допущений.

Подробный разбор результатов приведен, например, в работе [4]. Там же отмечено существенное отличие результатов анализа данных, приводимых исследователями лаборатории Bellcore, от полученных традиционными методами общепринятой теории очередей для Интернета.

Наиболее важным выводом во всех исследованиях являлось заключение о том, что чем больше нагрузка на Интернет, тем более близок параметр Херста к единице, следовательно, выше степень самоподобия трафика, а вопросы производительности приобретают наибольшую актуальность именно при высоких нагрузках. Среди наиболее вероятных причин самоподобия трафика в Интернете доминирует версия об асинхронности байтовых потоков в различных сегментах между транзитными узлами (асинхронность позиционирования кадров), вызванной мультиплексированием и демultipлексированием информации в сетях.

Вместе с тем, мало исследовались потоки кадров от самих источников для выявления и оценки их самоподобия [5, 6]. В настоящей статье приведены результаты определения характера сетевого трафика, поступающего со выхода сетевого адаптера персонального компьютера в транспортную систему Интернета, а также характера трафика, поступающего на вход этого же сетевого адаптера.

Для исключения влияния транзитных узлов сети на характер трафика два идентичных компьютера подключались друг к другу кроссовым сетевым кабелем. С помощью утилиты *nanoping* осуществлялась посылка серии эхо-запросов протокола ICMP (Internet Control Message Protocol) с компьютера-источника на компьютер-приемник, от которого возвращался эхо-ответ. На компьютере-источнике замерялось время прохождения кадров в обе стороны (RTT — Round Trip Time) [7].

Результаты измерения импортировались в Excel для дальнейшей обработки и определения параметра Херста для случайных значений RTT, полученных в каждой посылке эхо-запроса. При первой обработке для последовательности случайных величин x_i (значение RTT в i -й посылке эхо-запроса), используя критерий R/S [4], определяли математическое ожидание M и среднеквадратическое отклонение S по известным соотношениям:

$$\begin{aligned} M &= \frac{1}{N} \sum_{i=1}^N x_i; \\ X &= \frac{1}{N} \sum_{i=1}^N x_i^2; \\ D &= X - M^2; \\ S &= \sqrt{D}. \end{aligned} \quad (2)$$

При второй обработке последовательности x_i определяли следующие параметры:

$$\begin{aligned} L_{\max} &= \max_{1 \leq j \leq N} \sum_{i=1}^j (x_i - M); \\ L_{\min} &= \min_{1 \leq j \leq N} \sum_{i=1}^j (x_i - M); \\ R &= L_{\max} - L_{\min}. \end{aligned} \quad (3)$$

Отношение R/S , названное Херстом масштабированным диапазоном (scaled range) [2], для больших значений N хорошо описывается эмпирической формулой [4]

$$\frac{R}{S} \sim \left(\frac{N}{2}\right)^H \quad \text{при } H > 0,5. \quad (4)$$

Логарифмируя обе части выражения (4), получаем зависимость $\ln \frac{R}{S} \approx H \ln \frac{N}{2}$, откуда

$$H \approx \frac{\ln R/S}{\ln N/2}. \quad (5)$$

Варьируя значения числа посылок эхо-запросов утилиты *nanoping*, получаем значения R , S и H для различных N . Результаты расчетов приведены в табл. 1.

Результаты определения параметра Херста по отношению R/S

| | | | | | |
|------------|--------|--------|--------|--------|--------|
| N | 5000 | 10000 | 20000 | 40000 | 60000 |
| $\ln(N/2)$ | 7,824 | 8,517 | 9,210 | 9,903 | 10,309 |
| R/S | 1118,6 | 1527,8 | 2291,5 | 2561,3 | 3214,2 |
| $\ln(R/S)$ | 7,020 | 7,332 | 7,737 | 7,848 | 8,075 |
| H | 0,897 | 0,861 | 0,840 | 0,792 | 0,783 |

Значения H указывают на самоподобный характер трафика, что может быть следствием асинхронности тактовых частот функционирования сетевых адаптеров персональных компьютеров и их синхронизации преамбулой каждого передаваемого и возвращаемого кадра. Это предположение проверялось при исследовании характера трафика, поступающего на вход сетевого адаптера компьютера-источника, применением утилиты *nanoping* с сетевым адресом 127.0.0.1 (loopback), так как в этой ситуации эхо-запрос не поступает в аппаратно-программный тракт сетевого адаптера, а эхо-ответ формируется на сетевом уровне операционной системой компьютера-источника. Результаты определения H приведены в табл. 2.

Результаты определения параметра Херста по отношению R/S для loopback

| | | | | | |
|------------|---------|---------|---------|----------|----------|
| N | 5000 | 10000 | 20000 | 40000 | 60000 |
| $\ln(N/2)$ | 7,824 | 8,517 | 9,210 | 9,903 | 10,309 |
| R/S | 155,240 | 247,195 | 289,027 | 1452,502 | 3208,180 |
| $\ln(R/S)$ | 5,045 | 5,510 | 5,667 | 7,262 | 8,073 |
| H | 0,645 | 0,647 | 0,615 | 0,733 | 0,783 |

Значительное отличие параметра Херста от 0,5 позволяет сделать вывод, что самоподобным является поток кадров, поступающих на вход сетевого адаптера компьютера-источника.

Альтернативный отношению R/S метод установления самоподобия потока кадров может базироваться на определениях в точности самоподобного (exactly self-similar) или асимптотически самоподобного (asymptotically self-similar) процесса [4], опирающегося на n -агрегирование путем суммирования исходной серии случайных величин в блоке размером n , не перекрывающих друг друга, с вычислением средних значений этих блоков, т.е. переходом к последовательности новых случайных величин:

$$x_k^n = \frac{1}{n} \sum_{i=kn-(n-1)}^{kn} x_i, \quad k = 1, 2, \dots, m. \quad (6)$$

Дисперсия самоподобного процесса подчиняется зависимости [4]:

$$D(x_k^n) = \frac{D(x_i)}{n^\beta}, \quad (7)$$

где β — параметр, связанный с введенным ранее параметром Херста,

$$\beta = 2(1 - H). \quad (8)$$

Следовательно, для установления параметра Херста достаточно получить исходную случайную последовательность x_i , вычислить дисперсию этой последовательности и найти дисперсию $D(x_k^n)$. После этого значение параметра β находят по формуле

$$\beta = \log_n \left(\frac{D(x_i)}{D(x_k^n)} \right), \quad (9)$$

а параметр Херста определяют по соотношению

$$H = 1 - \frac{\beta}{2}. \quad (10)$$

Для экспериментальных исследований был написан скрипт генерации m серий эхо-запросов программой *nanoping* по n посылок в каждой, листинг которого приведен на рисунке.

```
#!/bin/sh
w=`grep ping $1 |cut -d: -f2`
ping_arg=`grep ping $1 |cut -d: -f3`
ping_cmd=nanoping
while read s
do
  if echo $s |grep -q node
  then
    src=`echo $s |cut -d: -f1`
    dst=`echo $s |cut -d: -f2`
    count=`echo $s |cut -d: -f3`
    while [ $count -gt 0 ]
    do
      ssh $src "(sleep 2; $ping_cmd $ping_arg -w $w $w
        $dst 2>&1 >${1}.${dst}.${count}.res) &"&
      count=$((count - 1))
    done
  fi
done < $1
```

Листинг генерации m серий эхо-запросов по n посылок в каждой

В табл. 3 приведены значения параметра Херста для серии из 10000 посылок по n эхо-запросов ICMP пакетами размером 14,72 кБ (длина

полезной нагрузки Ethernet-кадра равна 15 кБ) и периодом 0,001 с. Значение β рассчитывались по формуле

$$\beta = \frac{\ln D(x_i) - \ln D(x_k^n)}{\ln n},$$

а H – по формуле (9).

Таблица 3

Значения параметра Херста, найденные n -агрегированием для кроссового соединения компьютеров

| n | $D(x_k^n)$ | $\ln D(x_k^n)$ | $\ln n$ | β | H |
|-----|------------|----------------|---------|---------|-------|
| 1 | 923236587 | 20,643 | 0 | – | – |
| 2 | 701958416 | 20,369 | 0,963 | 0,395 | 0,802 |
| 4 | 655228947 | 20,300 | 1,386 | 0,204 | 0,898 |
| 6 | 581868276 | 20,182 | 1,792 | 0,257 | 0,871 |
| 8 | 522206043 | 20,074 | 2,079 | 0,274 | 0,863 |
| 16 | 254648132 | 19,355 | 2,773 | 0,464 | 0,768 |
| 32 | 104673850 | 18,466 | 3,466 | 0,628 | 0,686 |
| 64 | 53577259 | 17,797 | 4,159 | 0,684 | 0,659 |
| 128 | 30384449 | 17,229 | 4,852 | 0,704 | 0,648 |

Аналогичные результаты для посылки эхо-запросов по IP-адресу 127.0.0.1 (loorback) приведены в табл. 4.

Таблица 4

Значения параметра Херста, найденные n -агрегированием для loorback

| n | $D(x_k^n)$ | $\ln D(x_k^n)$ | $\ln n$ | β | H |
|-----|------------|----------------|---------|---------|-------|
| 1 | 898128261 | 20,616 | 0 | – | – |
| 2 | 669248228 | 20,322 | 0,693 | 0,424 | 0,788 |
| 4 | 652273846 | 20,296 | 1,386 | 0,231 | 0,855 |
| 6 | 679703981 | 20,337 | 1,792 | 0,156 | 0,922 |
| 8 | 694557493 | 20,359 | 2,079 | 0,124 | 0,938 |
| 16 | 579036718 | 20,177 | 2,773 | 0,158 | 0,921 |
| 32 | 531515633 | 20,091 | 3,466 | 0,151 | 0,924 |
| 64 | 587572601 | 20,192 | 4,159 | 0,102 | 0,949 |
| 128 | 600593579 | 20,213 | 4,852 | 0,083 | 0,958 |

В табл. 3 и 4 дисперсии случайных величин приведены в $нс^2$. Отметим, что для стационарного эргодического процесса ($H = 0,5$ и $\beta = 1$) дисперсия случайных n -агрегированных величин стремится к нулю со

скоростью, обратно пропорциональной числу агрегируемых данных, а для строго периодических (регулярных) последовательностей $H = 1$, $\beta = 0$, $D(x_i) = 0$ и $D(x_k^n) = 0$. Близость параметра Херста к единице (см. табл. 4) свидетельствует о том, что период посылки эхо-запросов операционной системой компьютера-источника почти постоянен. Случайные отклонения можно объяснить всегда присутствующей асинхронностью функционирования процессора, общей шины, дискового накопителя, оперативной памяти и прочих узлов в современных компьютерах. “Добавка случайности” сетевым адаптером, заставляет констатировать тот факт, что сетевой трафик любого компьютера уже на уровне его генерации носит самоподобный характер, что существенно усложняет его анализ и моделирование.

ЛИТЕРАТУРА

1. *Norros I.* A storage model with self-similar input // *Queueing Systems*. 1994. Vol. 16. P. 387–396.
2. *City Hurst H., Black R. and Simaika Y.* Long-term Storage: An Experimental Study, London, Constable, 1965.
3. *Клейкрок Л.* Вычислительные системы с очередями. М.: Мир, 1979. 595 с.
4. *Столлингс В.* Современные компьютерные сети. СПб.: Питер, 2003. 782 с.
5. *Erramilli A., Narayan O. and Willinger W.* Experimental Queuing Analysis with Long-Range Dependent Packet Traffic // *IEEE/ACM Transactions on Networking*, April, 1996.
6. *Иванов И.П.* Математические модели, методы анализа и управления в корпоративных сетях. Автореферат дис. . . д-ра. техн. наук. М.: МГТУ им. Н.Э. Баумана. 2010. 34 с.
7. *Иванов И.П., Кондратьев А.Ю., Лохтуров В.А.* Модернизация процесса измерений интервалов времени в операционных системах современных компьютеров // *Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение*. 2012. № 4. С. 44–59.

REFERENCES

- [1] *Norros I.* A storage model with self-similar input. *Queueing Systems*, 1994, vol. 16, pp. 387–396.
- [2] *Hurst H., Black R., Simaika Y.* Long-term Storage: An Experimental Study, CityplaceLondon, Constable, 1965.
- [3] *Kleykrok L.* Russ. ed.: Vychislitel'nye sistemy s ocheredyami [Queueing Computer Systems]. Moscow, Mir Publ., 1979. 595 p.
- [4] *Stollings V.* Russ. ed.: Sovremennye komp'yuternye seti [Contemporary Computer Networks], St. Petersburg, Piter Publ., 2003. 782 p.
- [5] *Erramilli A., Narayan O., Willinger W.* Experimental Queuing Analysis with Long-Range Dependent Packet Traffic. *IEEE/ACM Transactions on Networking*, April 1996.
- [6] *Ivanov I.P.* Matematicheskie modeli, metody analiza i upravleniya v korporativnykh setyakh. Avtoreferat dis. dokt. tekhn. nauk [Mathematical Models, Methods of Analysis and Management in Corporate Networks. Dr. tehn. sci. diss. abstr.]. Moscow, MGTU im. N.E. Bauman, 2010. 34 p.

[7] Ivanov I.P., Kondrat'ev A.Yu., Lokhturov V.A. Modernization of process of measuring time intervals in operating systems of contemporary computers. *Vestn. Mosk. Gos. Tekh. Univ. im. N.E. Baumana, Priborost.* [Herald of the Bauman Moscow State Tech. Univ., Instrum. Eng.], 2012, pp. 44–59 (in Russ.).

Статья поступила в редакцию 9.11.2014

Бойченко Максим Константинович — начальник ИЦ структурного подразделения “Управление информатизации – Вычислительный центр” МГТУ им. Н.Э. Баумана. Специалист в области информационно-коммуникационных технологий. МГТУ им. Н.Э. Баумана, Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5.

Boychenko M.K. — leading programmer of the IT laboratory of the Administration on Informatization – Computing Center of the Bauman Moscow State Technical University. Specializes in the field of data-communication technologies. Bauman Moscow State Technical University, 2-ya Baumanskaya ul. 5, Moscow, 105005 Russian Federation.

Иванов Игорь Потапович — д-р техн. наук, доцент, заведующий кафедрой “Теоретическая информатика и компьютерные технологии”, проректор по информатизации и модернизации МГТУ им. Н.Э. Баумана. Автор более 40 научных работ в области информационно-коммуникационных технологий. МГТУ им. Н.Э. Баумана, Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5.

Ivanov I.P. — Dr. Sci. (Eng.), prorector in informatization and modernization of the Bauman Moscow State Technical University, head of "Theoretical Informatics and Computer Technologies" department of the Bauman Moscow State Technical University. Author of more than 40 publications in the field of data-communication technologies. Bauman Moscow State Technical University, 2-ya Baumanskaya ul. 5, Moscow, 105005 Russian Federation.

Лохтуров Вячеслав Александрович — ведущий электроник лаборатории АИС структурного подразделения “Управление информатизации – Вычислительный центр” МГТУ им. Н.Э. Баумана. Специалист в области информационно-коммуникационных технологий. МГТУ им. Н.Э. Баумана, Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5.

Lokhturov V.A. — leading electronics engineer of the AIS laboratory of the Administration on Informatization – Computing Center of the Bauman Moscow State Technical University. Specializes in the field of data-communication technologies. Bauman Moscow State Technical University, 2-ya Baumanskaya ul. 5, Moscow, 105005 Russian Federation.

Издательство МГТУ им. Н.Э. Баумана

Сдано в набор 1.03.2015
Формат 70 × 108/16
Заказ

Подписано в печать 7.04.2015
Усл.-печ. л. 13,31
Уч.-изд. л. 14,23

Отпечатано в типографии МГТУ им. Н.Э. Баумана