STUDIA INFORMATICA 2002
Volume 23 Number 2A (48)

Tadeusz CZACHÓRSKI, Joanna DOMAŃSKA
Instytut Informatyki Teoretycznej i Stosowanej PAN

# WPŁYW SAMOPODOBNEJ NATURY RUCHU NA DŁUGOŚCI KOLEJEK W PRZEŁĄCZNIKACH SIECIOWYCH

**Streszczenie.** Artykuł opisuje wpływ doboru modelu źródła ruchu na szacowanie rozkładu długości kolejek w przełącznikach sieciowych. Porównano model bufora w przełączniku używający poissonowskiego źródła ruchu z modelem używającym źródła generującego ruch samopodobny.

# THE INFLUENCE OF SELF-SIMILAR NATURE OF NETWORK TRAFFIC ON QUEUE LENGTHS AT NETWORK SWITCHES

**Summary.** The article investigates the influence of stochastic features of the input traffic on the length of packet queues being formed in network switches. We compare queues having Poisson and self-similar input traffic. Both analytical and simulation models are used.

## 1. Wstęp

W konwencjonalnych modelach ruchu sieciowego używano zazwyczaj poissonowskich źródeł ruchu. Tego typu źródła ruchu były używane już wcześniej (przed powstaniem sieci komputerowych) do modelowania ruchu w telefonii. Ponieważ jednak natura ruchu sieciowego różni się znacznie od ruchu telefonicznego, dlatego też modele używające poissonowskich źródeł ruchu dawały błędne oszacowania parametrów ruchu sieciowego. Rzeczywisty ruch sieciowy wykazuje bowiem cechę samopodobieństwa [2] - można ją zaobserwować na różnych poziomach modelu warstwowego sieci komputerowych [5], [11], jak również przy różnych strukturach sieciowych [7], [17]. Niniejszy artykuł pokazuje

wpływ doboru modelu źródła ruchu na szacowanie rozkładu długości kolejek pakietów tworzących się w przełącznikach sieciowych. Porównano model bufora w przełączniku używający poissonowskiego źródła ruchu z modelem używającym źródła generującego ruch samopodobny.

Układ niniejszego artykułu jest następujący: w sekcji 2 zawarto krótki opis markowowskiego źródła ruchu samopodobnego wraz z dopasowaniem takiego źródła do danych rzeczywistych, sekcja 3 zawiera opis modelu bufora w przełączniku sieciowym, w sekcji 4 przedstawiono analizę rozkładu długości kolejki dla źródeł poissonowskich i samopodobnych, sekcja 5 zawiera wnioski wypływające z analizy.

## 2. Markowowskie źródło ruchu samopodobnego

W związku z zauważeniem cechy samopodobieństwa w rzeczywistym ruchu sieciowym powstały nowe metody modelowania takich źródeł [6], [9], [10], [17]. Podjęto również próby adaptacji konwencjonalnych metod modelowania do generowania ruchu samopodobnego [1], [12], [13], [3], [4].

W badaniach opisanych w niniejszym artykule wykorzystano model źródła samopodobnego, opisany szerzej w [3]. Model ten oparty jest na procesie typu SSMP(5) [14], [13], w którym nadejście klienta w przedziale czasowym o długości $t$ (tzw. *slocie czasowym*) uzależnione jest od stanu odrębnego łańcucha Markowa z dyskretnym czasem, tzw. *modulatora*, opisanego macierzą prawdopodobieństw przejść o wymiarach $5 \times 5$. Celem ułatwienia dopasowania takiego źródła do rzeczywistego ruchu sieciowego współczynniki macierzy prawdopodobieństw przejść uzależniono od dwóch parametrów:

$$\mathbf{A} = \left[\begin{array}{c|c|c|c|c} 1-1/a-1/a^2-1/a^3-1/a^4 & 1/a & 1/a^2 & 1/a^3 & 1/a^4 \\ \hline q/a & 1-q & 0 & 0 & 0 \\ \hline (q/a)^2 & 0 & 1-(q/a)^2 & 0 & 0 \\ \hline (q/a)^3 & 0 & 0 & 1-(q/a)^3 & 0 \\ \hline (q/a)^4 & 0 & 0 & 0 & 1-(q/a)^4 \end{array}\right]$$

Przyjęto założenie, że pakiet przychodzi w slocie czasowym $t$, jeżeli modulator jest w stanie pierwszym. Wartość średnia takiego procesu nadejść jest równa prawdopodobieństwu, że modulator jest w stanie pierwszym:

$$E[X] = \frac{1-(1/q)}{1-(1/q)^5} \tag{1}$$

Jeżeli taki model ma odzwierciedlać rzeczywisty ruch sieciowy, należy dopasować wartość średnią procesu oraz tzw. parametr Hursta, który charakteryzuje stopień samopodobieństwa procesu [2] do danych zebranych w sieci. Równanie (1) pozwala oszacować parametr $q$ procesu na podstawie zmierzonej wartości średniej. Parametr $a$ macierzy A jest używany do dopasowania wartości parametru Hursta.

Poniżej przedstawiono uzyskane wartości parametrów $a$ i $q$ dla:

- ruchu pakietów w sieci Ethernet (na podstawie pomiarów udostępnianych przez Laboratorium w Bellcore),
- ruchu ramek MPEG-2 (na podstawie pomiarów dokonanych w Laboratorium Komputerowych Systemów Multimedialnych w IITiS PAN).

Dopasowane parametry $q$ wynoszą odpowiednio $q = 0.684268$ dla ruchu Ethernet oraz $q = 1.075841$ dla przesyłu ramek MPEG-2. Natomiast znaleziony parametr $a$ dla ruchu Ethernet wynosi $a = 11.249898$, co odpowiada parametrowi Hursta równemu $H = 0.85$. Odpowiednio, znaleziony parametr $a$ dla przesyłu ramek MPEG-2 wynosi $a = 2.755029$, co odpowiada parametrowi Hursta równemu 0.673906.

W niniejszym artykule wykorzystano źródła ruchu samopodobnego o powyższych parametrach, co daje następujące macierze prawdopodobieństw przejść modulatora:

- dla parametru Hursta równego $H = 0.85$ (oznaczana poniżej jako *macierz1*):

$$A = \begin{bmatrix} 0.902444078 & 0.08888975 & 0.007901388 & 0.000702352 & 0.0000624319 \\ 0.060824412 & 0.939175588 & 0 & 0 & 0 \\ 0.003699609 & 0 & 0.99630039 & 0 & 0 \\ 0.000225027 & 0 & 0 & 0.999774973 & 0 \\ 0.0000136871 & 0 & 0 & 0 & 0.99998631289 \end{bmatrix}$$

- dla parametru Hursta równego $H = 0.673906$ (oznaczana poniżej jako *macierz2*):

$$A = \begin{bmatrix} 0.440099181 & 0.362972586 & 0.131749098 & 0.047821311 & 0.017357825 \\ 0.39050079 & 0.609499210 & 0 & 0 & 0 \\ 0.152490867 & 0 & 0.847509133 & 0 & 0 \\ 0.059547804 & 0 & 0 & 0.940452196 & 0 \\ 0.023253464 & 0 & 0 & 0 & 0.976746536 \end{bmatrix}$$

## 3. Model bufora w przełączniku sieciowym

Celem zbadania wpływu doboru modelu źródła ruchu na szacowanie długości kolejek w przełącznikach sieciowych utworzono analityczny model bufora w takim przełączniku. Aby potwierdzić prawidłowość uzyskanych wyników, utworzono również analogiczny model symulacyjny.

### 3.1. Model markowowski

Celem modelowania działania bufora w przełączniku sieciowym utworzono model markowowski z czasem dyskretnym. Modele kolejkowe z dyskretnym czasem, mimo iż bardziej złożone w stosunku do modeli z ciągłym czasem, zyskały sobie miejsce w literaturze w początkach lat 90. (np. [18]) jako lepiej oddające charakter zdarzeń zachodzących w systemach cyfrowych, gdzie szczelina czasowa jest podstawową jednostką czasu narzucającą rytm pracy systemu.

Jako model bufora w przełączniku sieciowym użyto kolejek czasu dyskretnego typu Geo/D/1/N oraz SSMP(5)/D/1/N. Geo oznacza, że rozkład czasu pomiędzy nadejściami klientów jest geometryczny, tzn. klient przychodzi w jednostce czasu z prawdopodobieństwem $\alpha$, a nie przychodzi z prawdopodobieństwem $1-\alpha$. SSMP(5) oznacza, że rozkład czasu pomiędzy nadejściami klientów jest procesem typu SSMP, tzn. że nadejście klienta w jednostce czasu zależy od stanu modulatora (sekcja 2). N oznacza maksymalną liczbę klientów w systemie. D oznacza stały czas obsługi. Jako jednostkę czasu przyjęto czas obsługi.

W przypadku modelu typu Geo/D/1/N stanem systemu w danej jednostce czasu jest liczba klientów w systemie, natomiast w przypadku kolejki SSMP(5)/D/1/N jest to para liczb (liczba klientów w systemie, stan modulatora). Łatwo zauważyć, że jedynymi możliwymi stanami w przypadku kolejki Geo/D/1/N są stany 0 i 1, a w przypadku kolejki SSMP(5)/D/1/N są to stany: $(0,i)$, $(1,i)$, gdzie $i = 0, 1, ..., 4$. Oznacza to, że kolejka jest zawsze pusta. Dlatego też, aby zbadać zachowanie kolejki, należy w powyższych modelach uwzględnić co najmniej dwa źródła. Liczba stanów rośnie wtedy w przypadku kolejki Geo/D/1/N do wartości $N+1$, natomiast w przypadku modelu z dwoma źródłami SSMP(5) jest rzędu $25 * (N+1)$. Dalsze zwiększanie liczby źródeł w przypadku modelu Geo/D/1/N nie powoduje wzrostu liczby stanów, natomiast dla modelu SSMP(5)/D/1/N liczba stanów wzrasta iloczynowo – np. dla 5 źródeł i $N = 100$ liczba stanów modelu SSMP(5)/D/1/N przekracza 300000.

Komplikuje to numeryczne rozwiązanie takiego modelu, gdyż liczba stanów łańcucha Markowa determinuje liczbę równań w układzie równań, który trzeba rozwiązać, aby uzyskać prawdopodobieństwa poszczególnych stanów (na podstawie których można dopiero obliczyć poszukiwany rozkład długości kolejki). W układzie równań

$$p = pP$$

macierz $P$ jest macierzą prawdopodobieństw przejść dyskretnego łańcucha Markowa. Jest to macierz stochastyczna (suma elementów w każdym wierszu jest równa jeden), a jej elementy $p_{ij}$ stanowią prawdopodobieństwa przejścia ze stanu $i$ do stanu $j$. Wektor $p$ jest poszukiwanym wektorem prawdopodobieństw stanów.

Ponieważ macierz prawdopodobieństw przejść $P$ jest macierza rzadką, dlatego też konieczne jest zastosowanie specjalnej metody rozwiązywania dużych, rzadkich układów równań liniowych. Zastosowano metodę tzw. podprzestrzeni Kryłowa [16]. Jest to metoda projekcyjna, której podstawową ideą jest aproksymacja dokładnego rozwiązania sekwencją rozwiązań dla podprzestrzeni o rozmiarze dużo mniejszym od rozmiaru macierzy $P$. Analizę uzyskanych wyników zawarto w sekcji 4.

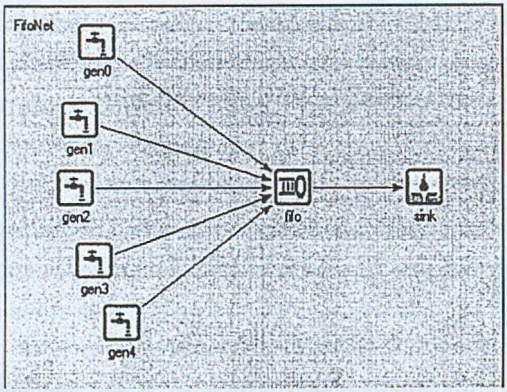
### 3.2. Model symulacyjny

Do stworzenia symulacyjnego modelu bufora w przełączniku sieciowym wykorzystano pakiet OMNeT++. Pakiet ten jest obiektowo zorientowanym symulatorem zdarzeń dyskretnych opartym na języku C++. Jest to pakiet niekomercyjny, dostępny w Internecie pod adresem: *http://www.hit.bme.hu/phd/vargaa/omnetpp.htm.*

Reprezentację utworzonego modelu w graficznym interfejsie pakietu OMNeT++ przedstawia rysunek 1. Utworzono dwa modele źródeł:

- geometryczne, jako dyskretny odpowiednik procesu Poissona,
- źródło samopodobne, oparte na procesie SSMP(5) opisanym w sekcji 2.

Symulacja takiego samopodobnego markowowskiego źródła wymagała ostrożności, ponieważ w macierzy A występowały duże różnice pomiędzy prawdopodobieństwami przejść. Zastosowano metodę kilkakrotnego losowania przy pojedynczej zmianie stanu modulatora – wprawdzie metoda ta spowalnia nieco generowanie klientów, ale za to jest precyzyjna i wiarygodna [13].




Rys. 1. Model bufora w przełączniku sieciowym
Fig. 1. Model of the considered network switches

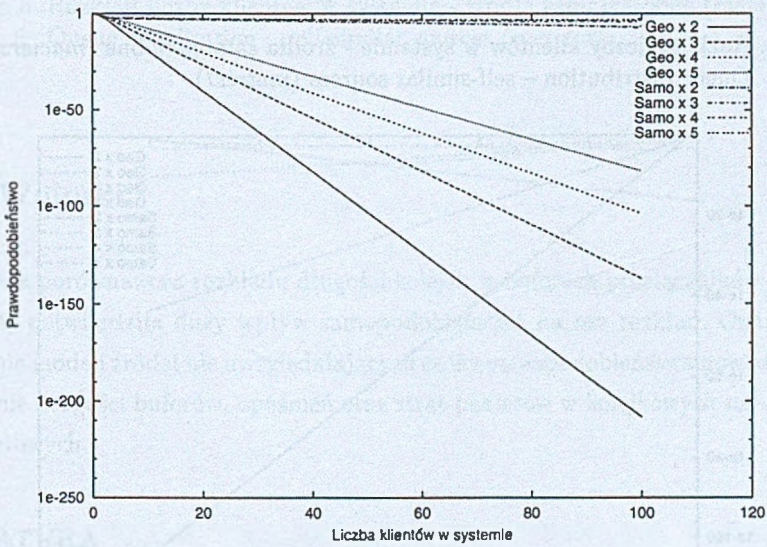
Stworzony model umożliwia dołączanie dowolnej liczby źródeł. Przyjęto stały czas obsługi równy slotowi czasowemu dyskretnego procesu nadejść. Celem uzyskania wiarygodnych wyników ustawiono zatrzymanie symulacji po wygenerowaniu 50 milionów klientów. Analizę uzyskanych wyników przedstawiono w sekcji 4.

## 4. Analiza rozkładu długości kolejki dla źródeł poissonowskich i samopodobnych

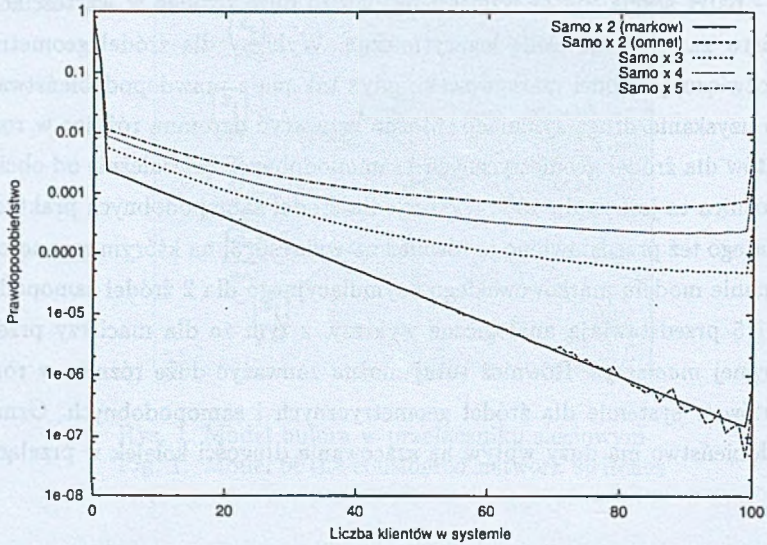
Aby porównać rozkład zajętości bufora w przełączniku sieciowym dla źródła geometrycznego i samopodobnego, należy odpowiednio dobrać parametr $\alpha$ źródła geometrycznego. Ponieważ $\alpha$ jest zarazem wartością średnią procesu geometrycznego, należy więc nadać jej wartość równą wartości średniej procesu nadejść SSMP(5). Dla *macierzy1* (patrz sekcja 2) wartość $E_1(X) = 0.081435101$, natomiast dla *macierzy2* wartość $E_2(X) = 0.230254427$. Przeanalizowano zajętość bufora w przełączniku przy różnych obciążeniach – zmieniano liczbę źródeł od dwóch do pięciu. W przypadku pięciu źródeł o wartości średniej $E_2(X)$ współczynnik wykorzystania systemu $\rho > 1$. Oznacza to, że stanowisko cały czas pracuje. Można to zauważyć w rozkładzie liczby klientów w systemie na wykresach 4 i 5.

Na wykresie 2 przedstawiono rozkład liczby klientów w systemie dla zmiennej

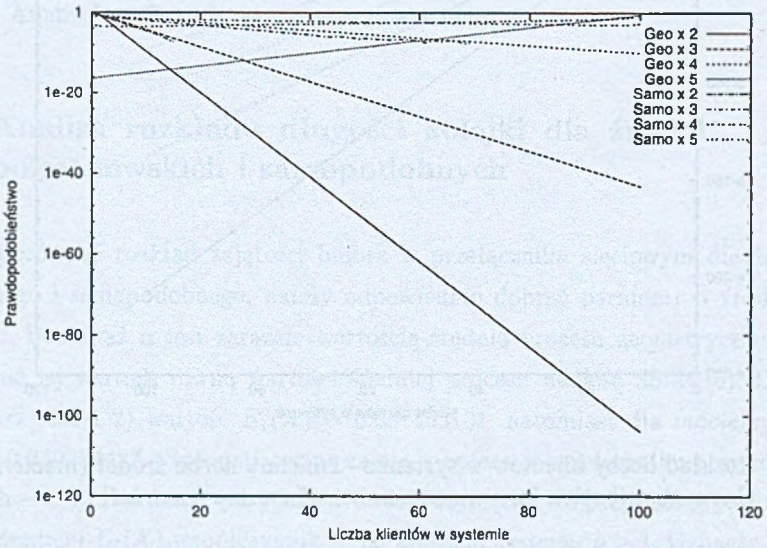
liczby źródeł geometrycznych i samopodobnych (macierz przejść modulatora równa *macierzy1* – patrz sekcja 2). Ze względu na bardzo duże różnice w wartościach prawdopodobieństw zastosowano skalę logarytmiczną. Wykresy dla źródeł geometrycznych uzyskano rozwiązując model markowowski, gdyż tak małe prawdopodobieństwa są niemożliwe do uzyskania drogą symulacji. Można zauważyć ogromną różnicę w rozkładzie liczby klientów dla źródeł geometrycznych i samopodobnych (niezależnie od obciążenia). Ponieważ różnica ta jest rzędu $10^{20}$, wykresy dla źródeł samopodobnych praktycznie się zlewają. Dlatego też przedstawiono je również na wykresie 3, na którym zaznaczono również porównanie modelu markowowskiego i symulacyjnego dla 2 źródeł samopodobnych. Rysunki 4 i 5 przedstawiają analogiczne wykresy, z tym że dla macierzy przejść modulatora równej *macierzy2*. Również tutaj można zauważyć duże różnice w rozkładzie liczby klientów w systemie dla źródeł geometrycznych i samopodobnych. Oznacza to, że samopodobieństwo ma duży wpływ na szacowanie długości kolejek w przełącznikach sieciowych.



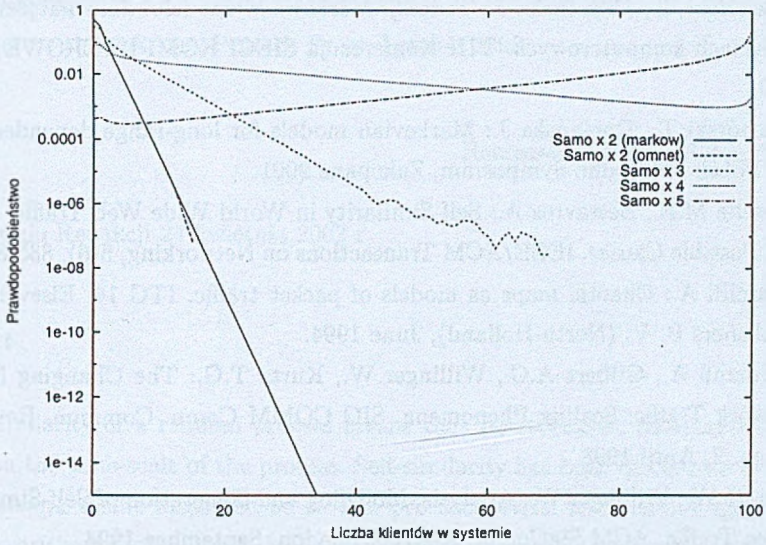

Rys. 2. Rozkład liczby klientów w systemie - zmienna liczba źródeł (*macierz1*)
Fig. 2. Queue distribution for various number of sources (*matrix1*)




Rys. 3. Rozkład liczby klientów w systemie - źródła samopodobne (*macierz1*)
Fig. 3. Queue distribution – self-similar sources (*matrix1*)




Rys. 4. Rozkład liczby klientów w systemie - zmienna liczba źródeł (*macierz2*)
Fig. 4. Queue distribution for various number of sources (*matrix2*)




Rys. 5. Rozkład liczby klientów w systemie - źródła samopodobne (*macierz2*)
Fig. 5. Queue distribution – self-similar sources (*matrix2*)

## 5. Wnioski

Analiza porównawcza rozkładu długości kolejek w buforach przełączników sieciowych (sekcja 4) potwierdziła duży wpływ samopodobieństwa na ten rozkład. Oznacza to, że stosowanie modeli źródeł nie uwzględniających cechy samopodobieństwa powoduje błędne szacowanie długości buforów, opóźnień oraz strat pakietów w kolejkowych modelach sieci komputerowych.

## LITERATURA

1. Baiocchi A., Melazzi N.B., Listanti M., Roveri A., Winkler R.: Loss Performance Analysis of an ATM Multiplexer Loaded with High-Speed ON-OFF Sources. IEEE-JSAC, vol. 9, no. 3, April 1991.
2. Czachórski T., Domańska J., Sochan A.: Samopodobny charakter natężenia ruchu w sieciach komputerowych. VIII Konferencja SIECI KOMPUTEROWE, Krynica 2001.

3. Czachórski T., Domańska J.: Procesy Markowa w modelowaniu natężenia ruchu w sieciach komputerowych. VIII Konferencja SIECI KOMPUTEROWE, Krynica 2001.
4. Czachórski T., Domańska J.: Markovian models for long-range dependent traffic. 8th Polish Teletraffic Symposium, Zakopane 2001.
5. Crovella M.E., Bestavros A.: Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes. IEEE/ACM Transactions on Networking, 5(6), 835-846, 1997.
6. Erramilli A.: Chaotic maps as models of packet traffic. ITC 14, Elsevier Science Publishers B. V. (North-Holland), June 1994.
7. Feldmann A., Gilbert A.C., Willinger W., Kurtz T.G.: The Changing Nature of Network Traffic: Scaling Phenomena. SIG COMM Comp. Commun. Review, vol. 28, no. 2, April 1998.
8. Garrett W., Willinger W.: Analysis, Modeling and Generation of Self-Similar VBR Video Traffic. ACM SigComm, 269-280, London, September 1994.
9. Jędruś S.: Modelowanie natężenia ruchu pakietów w sieciach komputerowych z wykorzystaniem miar multifraktalnych. Praca doktorska, IITiS PAN, Gliwice 1999.
10. Norros I.: On the use of fractional Brownian motion in the theory of connectionless networks. Technical contribution, TD94-33, September 1994.
11. Paxson V., Floyd S.: Wide Area Traffic: The failure of poisson modeling. IEEE/ACM Transactions on Networking, 3, 226-244, 1995.
12. Robert S., Le Boudec J.Y.: New Models for Pseudo Self-Similar Traffic. Performance Evaluation, vol. 30, (No 1-2): 57-68, July 1997.
13. Robert S.: Modélisation Markovienne du Trafic dans les Réseaux de Communication. PhD thesis, Ecole Polytechnique Fédérale de Lausanne, March 1996, No 1479, Switzerland.
14. Smith W.L.: Regenerative stochastic processes. Proc. Roy. Soc., Ser. A, vol. 232, p. 6-31, London 1955.
15. Stallings W.: High-Speed Networks: TCP/IP and ATM Design Principles, Prentice Hall, New Jersey 1998.
16. Stewart W.J.: An Introduction to the Numerical Solution of Markov Chains. Princeton Academic Press, 1994.
17. Willinger W., Leland W.E., Taqu M.S., Wilson D.: On the self-similar nature of Ethernet traffic (extended version). IEEE/ACM Transactions on Networking, February 1994.

18. Woodward M.E.: Communication and Computer Networks - Modelling with discrete-time queues. Pentech Press, London 1993.

Recenzent: Dr inż. Marcin Skowronek

Wpłynęło do Redakcji 24 kwietnia 2002 r.

## Abstract

Self-similarity of a random process means that its stochastic characteristics do not depend on the time-scale of the process. Self-similarity has been discovered in the intensity of network traffic measured on several protocol levels. It turns out that the traffic self-similarity has a great impact on a network performances. In particular, inregularity on large time scales and long-term autocorrelation related to it influence queue distributions observed in network switches, increases transmission delays and loss probabilities due to full buffers. Hence, it deteriorates quality of network services. The article investigates the influence of stochastic features of the input traffic on the length of packet queues being formed in network switches. We compare queues having Poisson and self-similar input traffic. Both analytical and simulation models are used.