

Tadeusz CZACHÓRSKI, Joanna DOMAŃSKA  
Instytut Informatyki Teoretycznej i Stosowanej PAN

## PROCESY MARKOWA W MODELOWANIU NATEŻENIA RUCHU W SIECIACH KOMPUTEROWYCH

**Streszczenie.** W artykule opisano sposób wykorzystania łańcuchów Markowa do generowania ruchu wykazującego cechę samopodobieństwa w skończonej skali czasu. Pomimo problemów z dopasowaniem tego typu źródeł ruchu do danych rzeczywistych dają one możliwość zastosowania dobrze poznanych teorii kolejkowych do modelowania pracy sieci komputerowych.

## MARKOV PROCESSES FOR MODELLING COMPUTER NETWORK TRAFFIC

**Summary.** The article presents the use of Markov Chains to generate a self-similar network traffic if a finite number of time scales is considered. In spite of problems to fit the model to the real data, this approach is recommended because it allows us to apply well known queueing models and modelling techniques for performance evaluation of existing communication and computer networks.

### 1. Wstęp

Przeprowadzane w ostatnich latach badania wykazały, że konwencjonalne metody modelowania nie oddają w pełni charakterystyki rzeczywistego ruchu w sieciach komputerowych. Źródła ruchu charakteryzowano zwykle zależnościami krótkoterminowymi, natomiast pomiary ruchu sieciowego wykazały, że zależności te są długoterminowe, co wiąże się z tzw. cechą samopodobieństwa procesów losowych [4]. Cechy tej nie uwzględniły konwencjonalne metody modelowania. W związku z tym zaczęły powstawać nowe

metody modelowania takich źródeł [5], [7], [10], [16]. Ich zaletą jest niewątpliwie dobry opis ruchu sieciowego przy niezbyt dużej liczbie parametrów. Jednakże nie dają one możliwości wykorzystania dobrze poznanych technik teorii kolejek do szacowania wydajności sieci komputerowych. Dlatego wskazane są próby dostosowania klasycznych metod modelowania źródeł do generowania ruchu samopodobnego [1], [12], [13].

Niniejszy artykuł pokazuje, jak za pomocą łańcuchów Markowa uzyskać źródło generujące ruch samopodobny (w skończonej skali czasu).

W części 2 opisano problemy modelowania sieci komputerowych związane z zauważoną w rzeczywistych strumieniach ruchu sieciowego cechą samopodobieństwa. Część 3 zawiera krótką charakterystykę procesów stochastycznych typu SSMP (ang. Special Semi-Markov Process) oraz opis markowskiego źródła ruchu samopodobnego. Część 4 opisuje problemy dopasowania takiego modelu do danych rzeczywistych. W części 5 podano wnioski wpływające z zastosowania modeli Markowa do modelowania samopodobnego ruchu w sieciach komputerowych.

## 2. Modelowanie sieci komputerowych

Modelowanie sieci komputerowych pomaga ich twórcom przewidzieć zachowanie projektowanej sieci, porównać różne topologie, zlokalizować miejsca zatłoczeń, scharakteryzować obciążenie sieci, określić liczbę i rodzaj koniecznych komponentów, a także przewidzieć pracę sieci przy rosnącym obciążeniu. Końcowym celem modelowania sieci jest jej optymalizacja na poziomie architektury w taki sposób, aby sieć spełniała (w najbardziej ekonomiczny sposób) wymagany standard jakości usług. Modele kolejkowe w postaci sieci stanowisk obsługi są uznanym i stosowanym od dawna narzędziem analizy i oceny efektywności pracy systemów komputerowych. Ich znaczenie rośnie wraz ze wzrostem złożoności badanych systemów (coraz trudniej zdać się na intuicję) oraz kosztem tych systemów (błędne decyzje podjęte na początku projektu powodują większe straty). Szczególne znaczenie ma modelowanie przy projektowaniu współczesnych sieci o zintegrowanych usługach, które muszą pogodzić wielu użytkowników generujących zróżnicowany ruch i oczekujących od sieci różnej jakości usług.

Jedną z tradycyjnych metod modelowania jest wykorzystanie łańcuchów Markowa: stanom badanego systemu odpowiadają stany łańcucha Markowa, a rozwiązując równania wiążące prawdopodobieństwa stanów łańcucha (są to równania algebraiczne w przypadku prawdopodobieństw stanów ustalonych i równania różniczkowe w przypadku

stanów nieustalonych) uzyskujemy prawdopodobieństwa stanów badanego systemu. Metoda ta pozwala uwzględnić dowolne zależności logiczne między stanami, wynikające np. ze stosowanych w modelowanej sieci mechanizmów regulacji natężenia ruchu, prowadzi jednak w przypadku realistycznych modeli do łańcuchów Markowa o bardzo dużej liczbie stanów. Wymagane jest w tym przypadku stosowanie specjalnych metod numerycznych do rozwiązywania bardzo dużych układów równań algebraicznych i różniczkowych.

W tradycyjnych modelach kolejkowych zakłada się, że strumień wejściowy klientów (pakietów, ramek, komórek – w przypadku modeli sieci komputerowych) jest charakteryzowany odstępami czasu pomiędzy nadejściami klientów, które są niezależnymi zmiennymi losowymi. Jednakże pomiary rzeczywistego ruchu sieciowego i ich analiza [16] wykazały, że ruch ten jest procesem losowym wykazującym silną autokorelację i mającym własność tzw. samopodobieństwa. Te cechy ruchu sieciowego zauważono między innymi wtedy, gdy projektowane za pomocą teorii kolejek rozmiary buforów okazały się zbyt małe [16]. Jest tak, ponieważ stosowanie uproszczonych, tradycyjnych modeli kolejkowych, pomijających autokorelację i samopodobieństwo strumienia klientów, prowadzi do bardzo znacznego zaniżenia długości przewidywanej kolejki. Samopodobieństwo procesu opisuje jego długoterminowe zależności i powoduje bardzo podobne zachowanie się procesu dla różnych skal czasu [8].

W związku z zauważonymi cechami ruchu sieciowego zaczęły powstawać nowe metody modelowania źródeł rzeczywistych oparte na samopodobnych procesach stochastycznych (ułamkowe ruchy Browna [10] i procesy ARIMA [16]), na deterministycznych nieliniowych mapach chaosu [5] i miarach multifraktalnych [7]. Ich zaletą jest niewątpliwie dobry opis ruchu sieciowego przy niezbyt dużej liczbie parametrów. Nie są jednak wolne od wad, np. za pomocą ułamkowych ruchów Browna nie można uzyskać różnego parametru Hursta dla różnych skali czasu. Trudno jednak wykorzystać te metody w analitycznych modelach kolejkowych. Zespół w Bellcore, analizujący własności statystyczne ruchu ramek w sieci Ethernet, ogłosił nawet koniec modeli markowskich w zastosowaniu do opisu rzeczywistego ruchu sieciowego [16].

Odrzucenie modeli Markowa jako nieprzydatnych do modelowania ruchu samopodobnego wynikało już z samej definicji procesów o zależnościach krótko- i długoterminowych [3]. Z definicji tych wynika, że do pierwszej kategorii procesów należą wszystkie procesy z powrotami skończonego rzędu, wszystkie skończone łańcuchy Markowa, natomiast do drugiej kategorii należą ułamkowe ruchy Browna, procesy ARIMA oraz mapy chaosu. Jednakże istnieje sposób, aby procesy o zależnościach krótkoterminowych znalazły zastosowanie w modelowaniu sieci komputerowych. Według matematycznej definicji sa-

mopodobieństwa, jeżeli proces wykazuje tę cechę dla kilku różnych skali czasu, to jest to proces o zależnościach krótkoterminowych, gdyż nie spełnia warunków nieskończoności. Jednakże do modelowania sieci taki proces jest wystarczający ze względu na to, że w praktyce wykorzystujemy zwykle do badań ograniczony zbiór danych. Taki proces wykazujący cechę samopodobieństwa dla kilku skal czasowych został nazwany procesem o zależnościach pseudodługoterminowych (ang. pseudo long-range dependent processes) [12].

Pojawiły się więc w literaturze próby opisu zmiennego w czasie, skorelowanego i samopodobnego procesu za pomocą tzw. procesów SMP (ang. Semi-Markov Processes). W części 3 zawarto bliższy opis tego typu procesów stochastycznych oraz pokazano, że za pomocą tego typu procesów można wygenerować ruch samopodobny.

### 3. Markowskie źródło ruchu samopodobnego

Proces Semi-Markowa (SMP - ang. Semi-Markov Process) został wprowadzony niezależnie przez Lévy'ego i Smitha [15] jako nowy rodzaj procesów stochastycznych. Procesy te są uogólnieniem ciągłych i dyskretnych procesów Markowa z przeliczalną przestrzenią stanów. Rozważając przypadek dyskretny, łańcuch Markowa jest modulowany i może zmienić swój stan w zdefiniowanym czasie:  $t_i, t_{i+1}, \dots$ . Pomiędzy chwilami czasowymi  $t_i$  i  $t_{i+1}$  łańcuch pozostaje w tym samym stanie. W każdym stanie przybysze są generowani według rozkładu  $\zeta_i$ , który jest uzależniony od stanu aktualnego i następnego. Proces SMPP (ang. Special Semi-Markov Process) jest to proces SMP, w którym rozkład  $\zeta_i$  dla danego stanu zależy wyłącznie od stanu aktualnego. Jeżeli  $\zeta_i$  jest rozkładem Poissona, wtedy proces SSMP nazywamy procesem MMPP (ang. Markov Modulated Poisson Process [1]). Niech  $X(t)$  będzie procesem stochastycznym reprezentującym liczbę komórek przybyłych w przedziale czasowym  $[t-1, t)$ . Niech  $Y_t = i$  oznacza  $i$ -ty stan modulatora w czasie  $t$  ( $i \in 1, 2, 3, \dots, n$ ). Przybycie komórek jest modulowane przez  $n$ -stanowy dyskretny łańcuch Markowa o prawdopodobieństwach tranzyjacji  $a_{ij}(t_1, t_2) = \text{prob}(Y_{t_2} = j | Y_{t_1} = i)$ . Niech  $\phi_{ij}$  oznacza prawdopodobieństwo przybycia  $j$  komórek w jednym przedziale czasowym przy stanie modulatora  $i$ . Inaczej  $\phi_{ij} = \text{prob}(X_t = j | Y_t = i)$ . Prawdopodobieństwa stanu modulowanego łańcucha Markowa są oznaczane przez  $\pi_i = \text{prob}(Y_t = i)$ , gdzie  $i$  jest stanem modulatora.

Poniższe rozważania zostaną zawężone do procesów z dyskretnym czasem. 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. jako lepiej oddające charakter zdarzeń zachodzących w systemach cyfrowych, gdzie szczelina czasowa jest podstawową jednostką czasu narzucającą rytm pracy systemu.

$k$ -ty moment dyskretnego procesu stochastycznego  $X_t$  jest równy:

$$E[X_t^k] = \sum_{i \geq 0} x_i^k Pr(X_t = x_i)$$

$X_t$  jest zmienną losową reprezentującą liczbę komórek przybyłych w przedziale czasowym  $[t-1, t)$ .  $X_t$  może przyjmować wartości dodatnie lub być równy zero. W przypadku procesu SSMP  $k$ -ty moment jest równy:

$$E[X_t^k] = \pi_t \Lambda_t^{(k)} e, \quad (1)$$

gdzie  $e$  jest wektorem jednostkowym, a  $\pi_t = (Pr(Y_t = 1), \dots, Pr(Y_t = n))$  jest wektorem prawdopodobieństw stanów modulatora,  $(Y_t = i)$  jest  $i$ -tym stanem modulatora ( $i \in 1, \dots, n$ ) w czasie  $t$ . Dla procesu SSMP macierz  $\Lambda_t^{(k)}$  jest równa:

$$\Lambda_t^{(k)} = \text{diag}(E(X_t^k | Y_t = 0), \dots, E(X_t^k | Y_t = n))$$

Jeżeli proces stochastyczny  $X_t$  jest stacjonarny w szerszym sensie (WSS - ang. Wide Sense Stationary), to:  $E[X^k] = \pi \Lambda^{(k)} e$  ( $e$  jest wektorem jednostkowym).

Ogólny wzór na funkcję autokowariancji dyskretnego procesu stochastycznego  $X_t$  jest następujący, np. [11]:

$$C(X_{t+\tau}^s, X_t^k) = E[X_{t+\tau}^s, X_t^k] - E[X_{t+\tau}^s]E[X_t^k]$$

Dla procesu SSMP funkcja autokowariancji przyjmuje następującą postać:

$$C(X_{t+\tau}^s, X_t^k) = \pi_t \Lambda_t^{(k)} (A(t+\tau, t) - e \pi_{t+\tau}) \Lambda_{t+\tau}^{(s)} e,$$

gdzie

$$(A(t+\tau, t))_{kj} = a_{kj}(t+\tau, t) = Pr(Y_{t+\tau} = j | Y_t = k), \quad 1 \leq k, j \leq n$$

jest elementem macierzy tranzycji łańcucha Markowa. Jeżeli dyskretny proces stochastyczny  $X_t$  jest stacjonarny w szerszym sensie i dodatkowo  $k = s = 1$ , to wtedy:

$$C(X_{t+\tau}^s, X_t^k) = C(X_\tau^s, X_0^k) = C(X_\tau, X_0) = C(\tau) = \pi \Lambda (A^{|\tau|} - e \pi) \Lambda e, \quad |\tau| \geq 1$$

Jeżeli macierz  $A^{|\tau|}$  zostanie zdekomponowana, wtedy funkcja autokowariancji  $C(\tau)$  przyjmuje następującą postać:

$$C(\tau) = \sum_{j=2}^n \lambda_j^{|\tau|} \pi \Lambda G^{-1} P_j G \Lambda e, \quad (2)$$

gdzie  $\lambda_j$  są wartościami własnymi macierzy  $A$ , a macierze  $G$  i  $P$  spełniają następujące zależności:

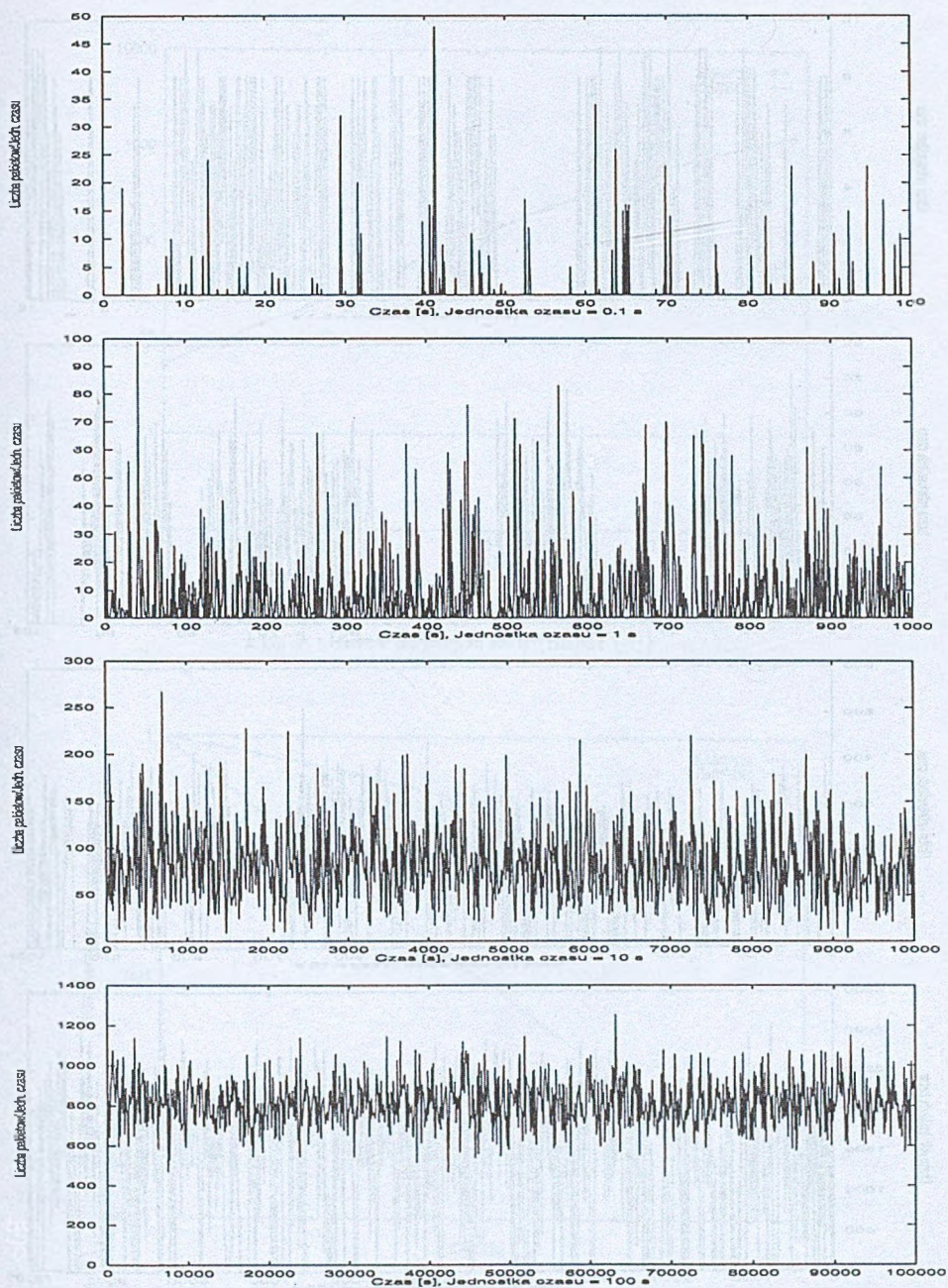
$$G^{-1}AG = \lambda_1 P_1 + \lambda_2 P_2 + \dots + \lambda_n P_n,$$

$$P_1 + P_2 + \dots + P_n = I,$$

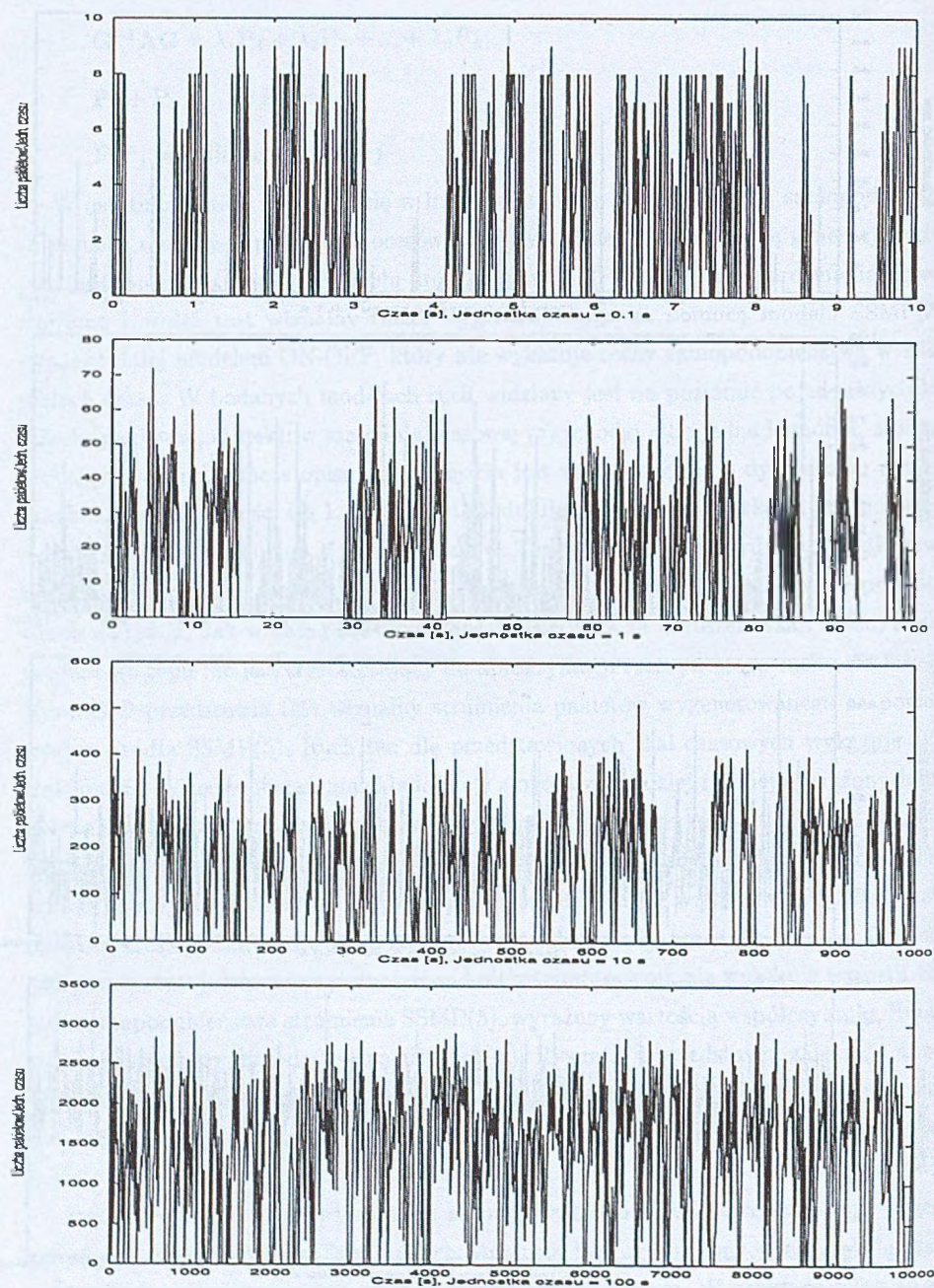
$$P_i P_j = 0 \text{ dla każdego } i \neq j$$

W ostatnich latach pojawiły się w literaturze próby generacji ruchu samopodobnego za pomocą opisanych powyżej procesów Markowa. Dalej opisano próbę generacji ruchu samopodobnego za pomocą modelu źródła typu SSMP(5). W celu porównania przedstawiono również test wizualny ruchu wygenerowanego za pomocą modelu SSMP(2), zwanego dalej modelem ON-OFF, który nie wykazuje cechy samopodobieństwa w kilku skalach czasu. W badanych modelach ruch widziany jest na poziomie pojedynczych komórek (pakietów, ramek): w szczelinie czasowej przychodzi albo jedna komórka, albo nie przychodzi żadna. Proces opisujący nadejścia jest więc procesem o dyskretnym czasie i przybierającym wartości 0 i 1. Natomiast modulujący go proces Markowa jest procesem o ograniczonej do kilku (np. 5) liczbie stanów. Test wizualny potwierdzający brak samopodobieństwa w kilku skalach czasu ruchu wygenerowanego modelem ON-OFF przedstawiono na rys. 1. Jak widać, ruch ten ujednocila się wraz ze wzrostem skali czasu, a więc model tego typu nie jest wystarczający do aproksymacji rzeczywistego ruchu sieciowego. Rysunek 2 przedstawia test wizualny strumienia pakietów wygenerowanego za pomocą modelu źródła SSMP(5). Ruch ten dla przedstawionych skal czasowych wykazuje dużą zmienność — okresy obciążenia składają się z okresów bardziej i mniej obciążonych, podobnie jak to jest w przypadku rzeczywistego ruchu w sieciach komputerowych. Na rys. 3 przedstawiono zagregowany wskaźnik dyspersji IDC (ang. Index of Dispersion of Counts, por. [4]) strumienia SSMP(5). Jak widać, IDC rośnie wraz ze wzrostem skali czasu, natomiast wykreślony dla porównania zagregowany wskaźnik dyspersji dla procesu Poissona, który reprezentuje proces z zależnościami krótkoterminowymi, nie wykazuje wzrostu. Stopień samopodobieństwa strumienia SSMP(5), wyrażony wartością współczynnika Hursta, został wyznaczony metodą wariancji (por. [4]). Rysunek 4 przedstawia zależność współczynnika wariancji od skali czasu (w skali logarytmicznej). Nachylenie prostej (oszacowane metodą najmniejszych kwadratów) wynosi  $-0.472$ , co daje współczynnik Hursta równy  $0.764$ .

Powyższy przykład potwierdza, że za pomocą łańcuchów Markowa można wygenerować ruch samopodobny. Taki łańcuch Markowa jest ograniczony własną skalą czasu, a więc ruch można generować tylko w skończonej skali czasu. W praktyce bywa to jednak trudne ze względu na problemy związane z dopasowaniem parametrów modelu do rzeczywistych danych (patrz część 4).

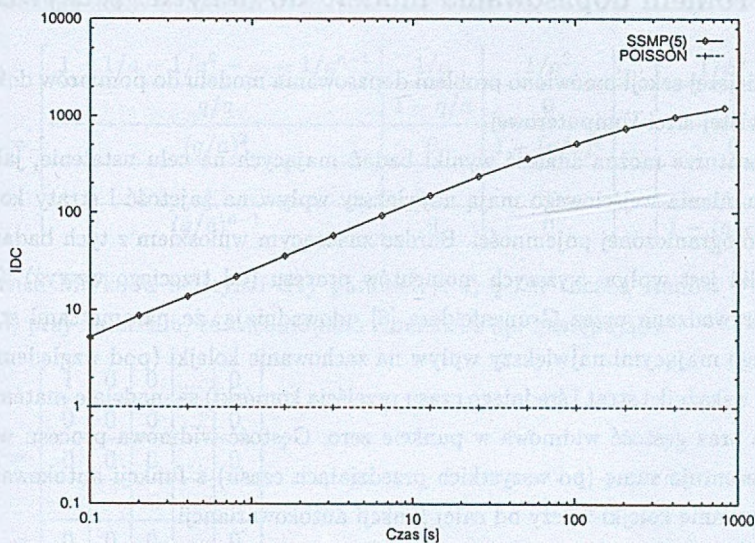


Rys. 1. Natężenie strumienia dla różnych jednostek czasu (model ON-OFF)  
Fig. 1. Packets per time unit over 4 time scales (ON-OFF model)



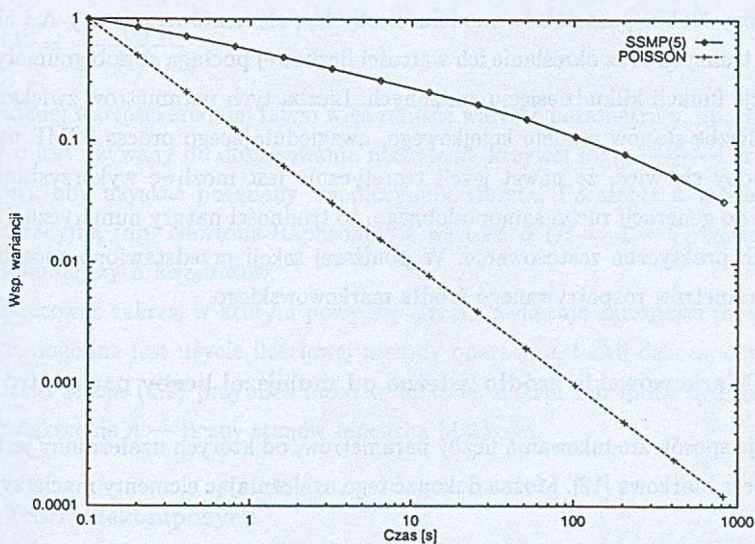
Rys. 2. Natężenie strumienia dla różnych jednostek czasu (model SSMP(5))  
 Fig. 2. Packets per time unit over 4 time scales (SSMP(5) model)





Rys. 3. Wskaźnik dyspersji(SSMP(5))

Fig. 3. Index of dispersion (SSMP(5))



Rys. 4. Zależność współczynnika wariancji od czasu – skala log-log (SSMP(5))

Fig. 4. "Variance-time" plot – log-log scale (SSMP(5))

## 4. Problem dopasowania modelu do danych rzeczywistych

W niniejszej sekcji omówiono problem dopasowania modelu do pomiarów dokonanych w rzeczywistej sieci komputerowej.

W literaturze można znaleźć wyniki badań mających na celu ustalenie, jakie parametry strumienia wejściowego mają największy wpływ na zajętość i straty komórek w buforach o ograniczonej pojemności. Bardzo znaczącym wnioskiem z tych badań jest to, że niewielki jest wpływ wyższych momentów procesu (od trzeciego wzwyż) [9]. Badania przeprowadzane przez Grunenfeldera [6] udowadniają, że parametrami strumienia wejściowego mającymi największy wpływ na zachowanie kolejki (pod względem średniej zajętości, wskaźnika strat i średniego czasu przejścia komórki) są: nadzieja matematyczna, wariancja oraz gęstość widmowa w punkcie zero. Gęstość widmowa procesu w punkcie zero reprezentuje sumę (po wszystkich przedziałach czasu) z funkcji autokowariancji, a więc zachowanie kolejki zależy od całej funkcji autokowariancji.

Korzystając z modeli Markowa opisanych w sekcji 3 można więc próbować aproksymować funkcję autokowariancji zmierzonych danych rzeczywistych (dla procesów samopodobnych ma ona przebieg hiperboliczny) funkcją autokowariancji, będącą sumą funkcji wykładniczych (wzór 2). Jest to jednak bardzo skomplikowane ze względu na dużą liczbę parametrów. Dobór parametrów modelu: liczby faz, elementów macierzy  $\Lambda$  i elementów macierzy tranzycji oraz określanie ich wartości liczbowej pociąga za sobą numeryczną minimalizację funkcji kilkudziesięciu zmiennych. Liczba tych parametrów zwiększa też iloczynowo liczbę stanów modelu kolejkowego, uwzględniającego proces SSMP na wejściu. Wydawałoby się więc, że nawet jeżeli teoretycznie jest możliwe wykorzystanie modeli Markowa do generacji ruchu samopodobnego, to trudności natury numerycznej uniemożliwiają ich praktyczne zastosowanie. W poniższej sekcji przedstawiono sposób redukcji liczby parametrów rozpatrywanego źródła markowskiego.

### 4.1. Markowskie źródło zależne od mniejszej liczby parametrów

Istnieje sposób zredukowania liczby parametrów, od których uzależniony jest modulowany proces Markowa [12]. Można dokonać tego uzależniając elementy macierzy tranzycji

A od niewielkiej liczby parametrów, np.:

$$A = \begin{bmatrix} 1 - 1/a - 1/a^2 - \dots - 1/a^{n-1} & 1/a & 1/a^2 & \dots & 1/a^{n-1} \\ q/a & 1 - q/a & 0 & \dots & 0 \\ (q/a)^2 & 0 & 1 - (q/a)^2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ (q/a)^{n-1} & 0 & 0 & \dots & 1 - (q/a)^{n-1} \end{bmatrix}$$

Taki łańcuch Markowa ma tylko trzy parametry:  $a$ ,  $q$  i  $n$  (liczba stanów w łańcuchu Markowa), przy założeniu, że zdefiniowano macierz  $A$  np. następująco:

$$A = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

W przypadku takiego modelu źródła wystarczy dopasować tylko dwa parametry: wartość średnią i wykładnik Hursta (plus liczbę stanów w łańcuchu Markowa).

Wartość oczekiwana powyższego modelu przyjmuje następującą postać:

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

Dla zadanej wartości średniej łatwo więc znaleźć wartość parametru  $q$ , np. iteracyjnie. Parametr  $a$  jest używany do dostosowania nachylenia krzywej  $\log_{10} \frac{\text{var}(X^{(m)})}{\sigma^2}$  w stosunku do  $\log_{10}(m)$ , aby uzyskać pożądany współczynnik Hursta. Parametr  $a$  można znaleźć metodą iteracyjną (np. Newtona-Raphsona), a wartość  $\beta$  ( $H = 1 - \frac{\beta}{2}$ ) oszacować np. metodą najmniejszych kwadratów.

Aby oszacować zakres, w którym powyższy proces wykazuje zależności pseudodługoterminowe, dogodnie jest użycie ilościowej metody opartej na teorii dekompozycji Courtois. Poniższa sekcja (4.2) przybliży nieco tę metodę. Zakres ten może być rozszerzany poprzez zwiększenie  $n$  — liczby stanów łańcucha Markowa.

#### 4.2. Teoria dekompozycji

Teoria dekompozycji została stworzona przez Courtois przeszło 20 lat temu. Analiza Courtois [2] opiera się na obserwacji, że duże systemy obliczeniowe można traktować jako systemy prawie zupełnie dekomponowalne. Systemy te są zbudowane w strukturze

hierarchicznej — z dużymi zależnościami pomiędzy składnikami tego samego poziomu i małymi zależnościami pomiędzy pozostałymi składnikami. Zjawisko prawie zupełnej dekomponowalności zostało zaobserwowane również w innych dziedzinach: w ekonomii, biologii, genetyce itp. Pionierami w tej dziedzinie byli Simon i Ando [14]. Przeprowadzali oni badania w dziedzinie fizyki i ekonomii i na ich podstawie stwierdzili, że agregacja zmiennych w prawie zupełnie dekomponowalnym systemie powinna rozdzielać analizę dynamiki krótko- i długoterminowej. Udowodnili oni dwa główne twierdzenia w tej dziedzinie. Pierwsze z nich mówi, że system prawie całkowicie dekomponowalny może być analizowany jako system całkowicie dekomponowalny pod warunkiem, że zależności pomiędzy grupami (poziomami w strukturze hierarchicznej) są bardzo słabe w porównaniu z zależnościami wewnątrz danej grupy. Drugie twierdzenie mówi, że rezultaty krótkoterminowe mogą być traktowane jako długoterminowe tak długo, jak długo panujemy nad względnym zachowaniem zmiennych tej samej grupy.

W przypadku problemu rozważanego w niniejszym artykule zagadnienie jest odwrotne: zakłada się, że ruch sieciowy składa się z różnych skali czasu. Łańcuch Markowa zaproponowany w poprzednim rozdziale jest faktycznie dekomponowalny na kilka poziomów. W pierwszym kroku należy wykonać dekompozycję dla tylko jednego poziomu. Różni się ona od tej zaproponowanej przez Simona, Ando i Courtoisa. Jak wykazano wcześniej, rozważany łańcuch Markowa można jednoznacznie scharakteryzować poprzez macierz tranzycji  $A$  (o wymiarze  $n \times n$ ) oraz wektor prawdopodobieństw stanów  $\pi$  ( $\pi_{t+1} = \pi_t A$ ). Macierz  $A$  jest prawie zupełnie dekomponowalna. Niech macierz  $A^*$  będzie macierzą całkowicie dekomponowalną (składa się z kwadratowych podmacierzy umieszczonych na przekątnej):

$$A^* = \begin{bmatrix} A_1^* & \dots & 0 & 0 \\ 0 & A_2^* & \dots & 0 \\ 0 & \dots & \dots & 0 \\ 0 & 0 & \dots & A_N^* \end{bmatrix} \quad (3)$$

Pozostałe elementy tej macierzy równe są 0. Macierz  $A_{IJ}^*$  jest podmacierzą macierzy  $A^*$  powstałą na przecięciu  $I$ -tego zbioru wierszy i  $J$ -tego zbioru kolumn. Element  $a_{i,jj}$  jest elementem na przecięciu  $i$ -tego wiersza i  $j$ -tej kolumny macierzy  $A_{IJ}^*$ . Z macierzą  $A^*$  związany jest nowy wektor prawdopodobieństw stanów  $\pi^*$  ( $\pi_{t+1}^* = \pi_t^* A^*$ ). Przyjmujemy następujące oznaczenia:  $\pi^*$  jest poziomym wektorem tego samego rozmiaru co  $A_{II}^*$ , gdzie  $I = 1, \dots, N$ . Dla ułatwienia macierz  $A_{II}^*$  oznaczmy przez  $A_i^*$ , gdzie  $i = 1, \dots, N$ . Jest to macierz kwadratowa o rozmiarze  $n(i) \times n(i)$ , gdzie  $\sum_{i=1}^N n(i) = n$ . Każda podmacierz

$A_i^*$  ma własny zbiór wartości własnych  $\lambda^*(i_j)$ . W celu ułatwienia zakładamy, że są one uporządkowane:  $\lambda^*(1_I) = 1 > \lambda^*(2_I) \geq \dots \geq \lambda^*(n(I)_I)$ , gdzie  $I = 1, \dots, N$ ;  $\lambda^*(1_I) = 1$ , ponieważ macierz jest stochastyczna. W przypadku macierzy  $A$  sytuacja jest inna, ponieważ tylko jedna wartość własna tej macierzy ( $\lambda(1_1)$ ) jest równa 1. Zakładamy, że macierz ta ma również uporządkowane wartości własne.

Wychodząc z założenia diagonalności macierzy  $A$  można napisać:

$$A = P^{-1}DP,$$

gdzie  $P$  jest macierzą przejść, a  $D$  można zapisać następująco:

$$D = \sum_i \lambda_i P_i = \sum_{i=1}^{n(I)} \sum_{l=1}^N \lambda(l_I) P_I(l),$$

$P_I$  jest rzutnikiem, tzn.  $p_{ij} = 0$  dla każdego  $i, j \neq l$ , a  $p_{ll} = 1$ . Tak więc:

$$A = P^{-1}P_1(1)P + \sum_{I=2}^N \lambda(1_I)P^{-1}P_I(1)P + \sum_{I=1}^N \sum_{i=2}^{n(I)} \lambda(i_I)P^{-1}P_I(i)P \quad (4)$$

Dla macierzy  $A^*$  można zapisać podobne wyrażenie:

$$A^* = P^{-1}P_1(1)P + \sum_{I=2}^N \lambda^*(1_I)P^{-1}P_I(1)P + \sum_{I=1}^N \sum_{i=2}^{n(I)} \lambda^*(i_I)P^{-1}P_I(i)P \quad (5)$$

Wyrażenie  $\lambda(i_I)P^{-1}P_I(i)P$  można zastąpić przez macierz  $Z(i_I)$  ( $z_{kl}(i_I)$  są elementami tej macierzy). Własności macierzy  $Z(i_I)$  można znaleźć w [2]. Zakładając, że macierz  $A$  ma postać  $A = A^* + \epsilon C$ , pierwsze twierdzenie Simona i Ando [14] brzmi następująco:

- Dla dowolnej dodatniej liczby  $\zeta$  istnieje liczba  $\epsilon_\zeta$  taka, że dla  $\epsilon \leq \epsilon_\zeta$   
 $\max_{k,l} |z_{kl}(i_I) - z_{kl}^*(i_I)| < \zeta$ ,  
 gdzie  $2 \leq i \leq n(I), 1 \leq I \leq N, 1 \leq k, l \leq n$ .

Twierdzenie to jest bardzo ważne dla naszych rozważań. W centrum zainteresowań powinno się znaleźć zachowanie (w czasie) wektora  $\pi_i$  i porównanie z zachowaniem  $\pi_i^*$ . W związku z uporządkowaniem wartości własnych pierwsze wyrażenia ze wzoru (5) nie wprowadzają dużych zmian dla krótkiego okresu czasu ( $t < T_1$ ), ponieważ wartości własne  $\lambda(1_I), I = 1, \dots, N$  są bliskie jedności. Wynika stąd, że dla  $t < T_1$  dominującą zmienną częścią równania (4) jest część ostatnia, a więc  $\pi_i$  i  $\pi_i^*$  zachowują się podobnie. Dla  $T_1 < t < T_2$  zachowanie w czasie wektorów  $\pi_i$  i  $\pi_i^*$  jest definiowane ostatnimi częściami wzorów (odpowiednio) (4) i (5). Wewnątrz każdego z podsystemów  $A$  i  $A^*$

zostaje osiągnięty podobny stan równowagi. Dla  $T_2 < t < T_3$  najbardziej zmienną częścią (4) jest część druga. Dla  $t > T_3$  część pierwsza dominuje nad pozostałymi. Stan równowagi jest osiągalny dla całego systemu. Prawie zupełnie dekomponowalny system zmierza do punktu równowagi. System w stanie równowagi krótkoterminowej zmierza do punktu równowagi, ale wartości zmiennych wewnątrz każdego z podsystemów są zachowane. To dynamiczne zachowanie prawie zupełnie dekomponowalnej macierzy można podzielić na cztery (wspomniane już wcześniej) okresy [14]:

- 1) dynamika krótkoterminowa ( $t < T_1$ ),
- 2) stan równowagi krótkoterminowej ( $T_1 < t < T_2$ ),
- 3) dynamika długoterminowa ( $T_2 < t < T_3$ ),
- 4) stan równowagi długoterminowej ( $t > T_3$ ).

Ogólną postać macierzy prawie zupełnie dekomponowalnej  $A$  przedstawiono w sekcji 4.1. Ogólna postać macierzy  $A^*$  (zbudowanej w oparciu o teorię Courtois) jest następująca ( $q < a$ ):

$$A^* = \left[ \begin{array}{c|c|c|c|c} 1 - 1/a - 1/a^2 - \dots - 1/a^{n-2} & 1/a & 1/a^2 & \dots & 0 \\ \hline q/a & 1 - q/a & 0 & \dots & 0 \\ \hline (q/a)^2 & 0 & 1 - (q/a)^2 & \dots & 0 \\ \hline \dots & \dots & \dots & \dots & \dots \\ \hline 0 & 0 & 0 & \dots & 1 \end{array} \right] \quad (6)$$

Macierz  $A^*$  jest macierzą nieergodyczną. Porównanie zachowania wektorów  $\pi_t$  i  $\pi_t^*$  pozwala określić zakres, w którym proces opisany macierzą  $A$  wykazuje zależności pseudodługoterminowe.

## 5. Podsumowanie

W niniejszym artykule wykazano, że za pomocą modeli Markowa typu SSMP można wygenerować ruch samopodobny w skończonej skali czasu, a więc można wykorzystać modele Markowa do aproksymacji zachowania rzeczywistego ruchu sieciowego w pewnym przedziale czasowym. Nie jest zwykle konieczne rozważanie np. rocznych pomiarów ruchu w sieci (w takim przypadku należałoby rozważyć quasi-nieskończone modele Markowa),

a więc przedział czasowy oferowany przez źródła markowowskie należy uznać za wystarczający. Pomimo opisanych trudności związanych z dopasowaniem tego typu modelu do danych rzeczywistych daje on możliwość wykorzystania dobrze poznanych teorii kolejek do szacowania wydajności 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. Courtois P.J.: *Decomposability*, ACM Monograph Series, 1977.
3. Cox D.R.: *Long-range dependence: A review*. *Statistics: An Appraisal*, 1984.
4. Czachórski T., Domańska J., Sochan A.: Samopodobny charakter natężenia ruchu w sieciach komputerowych. VIII Konferencja SIECI KOMPUTEROWE, Krynica, 2001.
5. Erramilli A.: *Chaotic maps as models of packet traffic*. ITC 14, Elsevier Science Publishers B. V. (North-Holland), June 1994.
6. Grunenfelder R., Robert S.: Decisive arrival law parameters and a general finite capacity queueing problem. *Performance Evaluation*, vol. 23, 199-215, September 1995.
7. Jędrus S.: Modelowanie natężenia ruchu pakietów w sieciach komputerowych z wykorzystaniem miar multifrakalnych. Praca Doktorska, IITiS PAN, Gliwice 1999.
8. Leland W.E., Wilson D.V.: High Time Resolution Measurements and Analysis of LAN Traffic: Implications for LAN Interconnection. *IEEE Infocomm'91*, paper 11D.3.1.
9. Li S., Hwang C.L.: Queue Response to Input Correlation Functions: Continuous Spectral Analysis. *IEEE/ACM Transactions on Networking*, vol. 1, 678-692, December 1993.
10. Norros I.: On the use of fractional Brownian motion in the theory of connectionless networks. Technical contribution, TD94-33, September 1994.
11. Papoulis A.: *Probability, Random Variables and Stochastic Processes*. Mc Graw Hill, 1984.
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. Simon H., Ando A.: Aggregation of variables in dynamic systems. *Econometrica*, no. 29, 1961.
15. Smith W.L.: Regenerative stochastic processes. *Proc. Roy. Soc., Ser. A*, vol. 232, p. 6-31, London 1955.
16. 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.

Recenzent: Dr inż. Ewa Starzewska-Karwan

Wpłynęło do Redakcji 10 kwietnia 2001 r.

### Abstract

Recent studies of high quality traffic measurements display a complex nature of real traffic observed in computer and telecommunication networks. Classically, the traffic intensity, seen as a stochastic process, was characterized by short term dependencies; the characteristics of measured traffic, see e.g. [4] have shown that it is long term dependent (self-similar). New models of such sources have been developed [5], [7], [10], [16]. The advantage of these models is that they give a good description of the traffic behavior considering only few parameters. Their drawbacks consist in the fact that they do not allow the use of traditional and well known queueing models and modelling techniques.

In this paper, we present a Markov process which can behave as self-similar one in a limited time domain (it is not reasonable to consider years-long measurements for network dimensioning). Also some problems concerning the fitting procedures of such a source model to real data are discussed. The model makes possible the adaptation of traditional Markovian queueing models to evaluation of existing integrated services networks.