

Joanna TOMASIK

Instytut Informatyki Teoretycznej i Stosowanej, PAN

## WPLYW MODULOWANIA ŹRÓDŁA NA FUNKCJĘ AUTOKOWARIANCJI GENEROWANEGO STRUMIENIA DANYCH

**Streszczenie.** W artykule pokazano, jakie cechy statystyczne mają strumienie danych przesyłane w sieciach szerokopasmowych. Analizowane strumienie pochodziły z transmisji ruchomych obrazów przesyłanych w standardzie MPEG. Następnie przedstawiono, jaki charakter ma funkcja autokowariancji strumieni generowanych przez pewną klasę źródeł modelowanych za pomocą łańcuchów Markowa i jakie strumienie rzeczywiste mogą być w taki sposób modelowane. Przedstawiono model przełącznika sieciowego, sformułowanego za pomocą formalizmu sieci automatów stochastycznych. Strumienie pakietów napływające na wejście przełącznika są superpozycją strumieni o różnym charakterze.

## AN IMPACT OF SOURCE MODULATION ON AN AUTOCORRELATION FUNCTION OF A GENERATED DATA STREAM

**Summary.** In the paper statistical features of data streams in broadband networks are discussed. The analyzed streams resulted from video transmissions, compressed and coded according to the MPEG standard. Further, it is discussed how an autocovariation function of a data stream generated by a certain group of sources modelled by Markov chains looks like and what kind of real streams may be modelled this way. As an example, a model of a multiservice network switch is presented. It is constructed with the Stochastic Automata Network formalism. The incoming packet stream is a superposition of streams of different types.

---

<sup>0</sup>Część niniejszego opracowania została zrealizowana w ramach grantu KBN 8 T11C 038 16.

## 1. Wstęp

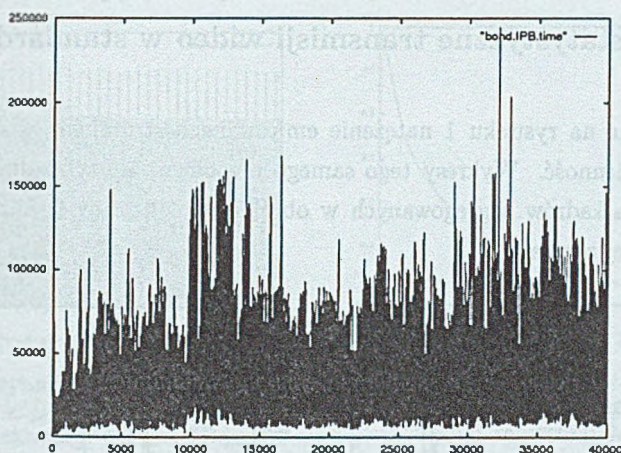
Najintensywniej rozwijającym się obecnie standardem sposobu transmisji w sieciach rozległych jest transmisja ATM (*Asynchronous Transfer Mode*). Zapewniające ją procedury znajdują się na drugiej warstwie według standardowego ISO/OSI. Zdefiniowanym na tym poziomie ramkom odpowiadają komórki ATM (*ATM cells*) o stałej długości (53 bajty). Łącze fizyczne, które realizuje transmisje ATM, charakteryzuje się szerokim dostępnym pasmem i dlatego w danym momencie może być wykorzystane przez różne strumienie przesyłanych danych. Przydział pasma odbywa się na zasadzie multipleksacji statystycznej [5]. Zachodzące w sieci transmisje mają różne wymagania wobec łącza. Przede wszystkim przy zestawianiu połączenia określany jest typ żądanej usługi (CBR — rezerwacja pasma o stałej szerokości, VBR — przydział pasma w zależności od aktualnej intensywności transmisji, ABR — wykorzystywanie aktualnie wolnego pasma w deklarowanym zakresie przy użyciu informacji przesyłanych w sprzężeniu zwrotnym, UBR — wykorzystanie aktualnie wolnego pasma bez zapewniania jakichkolwiek parametrów co do jego minimalnej szerokości (*best-effort service*) [11]). Dla usługi wykorzystującej zmienny przydział pasma podawane są też średnia intensywność ruchu (*Mean Bit Rate, MBR*) i szczytowa intensywność ruchu (*Peak Bit Rate, PBR*). Algorytm CAC (*Call Acceptance Control*) podejmuje decyzję, czy nowe połączenie może być zrealizowane, z zapewnieniem mu żądanej jakości. Jednocześnie z rozwojem samej technologii sieciowej postępują prace dotyczące modelowania strumieni ramek w sieciach, gdyż ruch ten jest na tyle specyficzny, że klasyczne metody modelowania stają się niewystarczające.

## 2. Transmisje obrazów ruchomych w sieciach komputerowych

Istotnym typem danych, przesyłanych w sieciach rozległych, są dane multimedialne, a w szczególności przesył obrazów ruchomych. Przesył takich danych następuje w formie skompresowanej i zakodowanej, aby minimalizować zajętość pasma przez daną transmisję. Rozpowszechnionym standardem kompresji obrazów ruchomych jest MPEG (*Moving Picture Coding Expert Group*) [4]. Poszczególne obrazy (klatki, kadry) mają ustaloną rozdzielczość (np.  $350 \times 250$ ,  $640 \times 480$ ). W jednostce czasu transmitowana jest ustalona liczba obrazów (typowo 24 obrazy na sekundę, 30 obrazów na sekundę). Na opis piksela przeznaczona jest 24 bity. Kompresja w tym standardzie odbywa się zarówno przez usuwanie redundancji w ramach pojedynczego obrazu, jak i w ramach całej ich sekwencji. Efektem kompresji i kodowania są obrazy trzech typów:

- I (*Intra*) — obrazy, do których stosowana jest wyłącznie kompresja przestrzenna,
- P (*Predicted*) — obrazy, do których stosowana jest kompresja przestrzenna oraz kompensacja ruchu w stosunku do poprzedniego kadru I- bądź P-,
- B (*Bi-directional*) — obrazy, do których stosowana jest kompresja przestrzenna oraz kompensacja ruchu w stosunku do poprzedniego kadru I- bądź P- i następnego kadru I- bądź P-, albo też kompensacja ruchu przebiega w stosunku do średniej z obrazu poprzedniego i następnego (interpolacja).

Zakodowane kadry są zbierane w grupy o stałej długości i powtarzającym się schemacie. Przykładowa grupa kadrów (*Group of Pictures, GOP*) ma postać: IBBPBBPB-BPBB. Przykładowa intensywność strumienia takich danych pochodzących z transmisji filmu przedstawiona jest na rysunku 1. Kadry nie są kodowane w kolejności wyświetlania, ponieważ kadr B zależy od przyszłości, może być otrzymany dopiero po zdygitalizowaniu następnego obrazu.



Rys. 1. Transmisja dla standardu MPEG; film „Goldfinger”  
Fig. 1. MPEG Transmission; film "Goldfinger"

Do tego rodzaju transmisji stosuje się transmisję VBR, która jest w stanie zapewnić wymagania czasowe i pozwala na utrzymanie jednakowej jakości poszczególnych obrazów. O jakości przesyłu wideo decydują następujące wielkości ([9]):

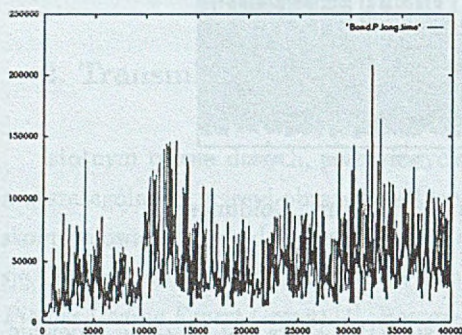
- strata komórek (*cell loss*) — kompresja obrazu cyfrowego usuwa występowanie redundancji w poszczególnych obrazach, co powoduje wzrost wpływu strat komórek na

jakość obrazu. Trudno określić jednoznacznie, jakie prawdopodobieństwo strat można uznać za tolerowalne. Jego wartość zależy od szeregu czynników, takich jak wrażliwość ludzkiego oka na różne typy pogorszenia jakości obrazu.

- **opóźnienie komórek** (*cell delay*) — przesył komórek w standardzie MPEG zawsze jest obciążony pewnym opóźnieniem, wynikającym z kodowania/dekodowania obrazu, które dodatkowo może być powiększone przez opóźnienia wynikające z dostępu do łącza, z buforowania komórek oraz opóźnienia transmisji. Opóźnienie to powinno być minimalizowane w przypadku transmisji interaktywnych (np. wideokonferencje), zatem dla transmisji tego typu nie zaleca się stosowania zmian charakteru strumienia (*shaping*), które spowodowałyby dodatkowe opóźnienie wynikające z buforowania komórek. Wartość opóźnienia dla transmisji interaktywnych tolerowana przez odbiorcę nie jest jednoznacznie określona. W przypadku transmisji telewizyjnych czy pokazów filmowych opóźnienie to nie ma zasadniczego wpływu na jakość przekazu postrzeganą przez odbiorcę.

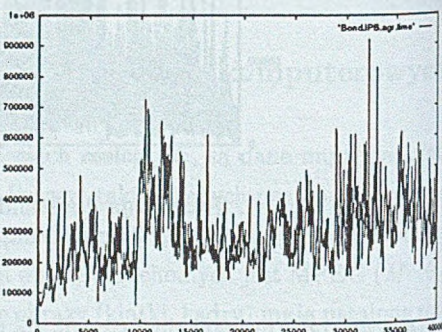
### 3. Cechy statystyczne transmisji wideo w standardzie MPEG

Przedstawione na rysunku 1 natężenie emitowanego strumienia w czasie wykazuje bardzo dużą zmienność. Wykresy tego samego przebiegu, ale tylko dla kadrów P (rysunek 2) czy dla kadrów zagregowanych w obrębie jednej grupy (rysunek 3) stają się wizualnie bardziej regularne.



Rys. 2. Przebieg ramek MPEG, tylko ramki P dla sekwencji 40000 ramek (Bond — Goldfinger)

Fig. 2. MPEG frames, P frames only for the 40000 frames sequention (Bond — Goldfinger)



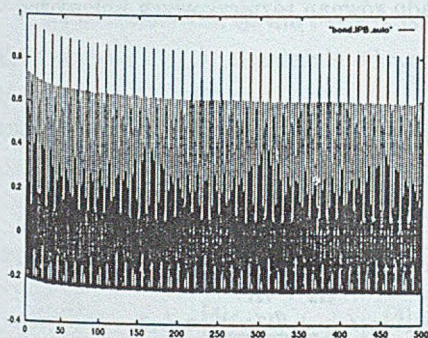
Rys. 3. Przebieg ramek MPEG, zagregowanych w ramach jednej grupy, dla sekwencji 40000 ramek

Fig. 3. MPEG frames, with aggregation made for each GOP, for the 40000 frames sequention

Różnice w kształtach poszczególnych krzywych można wytłumaczyć różnym zachowaniem się funkcji autokowariancji. Funkcja ta jest estymowana na podstawie wzoru [2]:

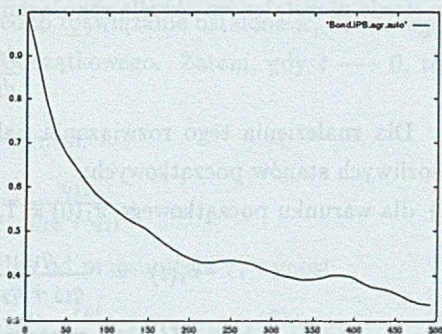
$$c(k) = \frac{1}{N} \sum_{i=1}^{N-k} (x(i) - \bar{x})(x(i+k) - \bar{x}),$$

gdzie  $x(i)$  jest realizacją procesu w chwili  $i$ , a  $\bar{x}$  jest średnią wartością procesu. Wartości  $k$  zmieniają się co 1 od 0 do  $K$ , przy czym  $K$  nie powinno być większe od  $\frac{N}{4}$ , gdzie  $N$  jest liczbą danych wartości procesu. Współczynnik autokowariancji jest estymowany jako:  $r(k) = c(k)/c(0)$ ,  $k = 0, 1, 2, \dots, K$ . Funkcja współczynnika autokowariancji policzona dla pełnego strumienia danych MPEG charakteryzuje się tym, że tendencja spadkowa jest słaba, a moduły wartości liczbowych współczynnika autokowariancji są relatywnie duże (rysunek 4). Tymczasem dla tej samej sekwencji MPEG, ale zagregowanej w ramach poszczególnych grup kadrów, funkcja przypomina funkcję wykładniczą i charakteryzuje się względnie szybkim spadkiem do wartości 0 (rysunek 5). Można więc zakładać, że na poziomie grup kadrów strumień danych może być modelowany za pomocą procesu o eksponencjalnie zanikającej funkcji autokorelacji.



Rys. 4. Współczynnik autokowariancji w funkcji przesunięcia dla oryginalnej sekwencji MPEG (Bond — Goldfinger)

Fig. 4. Autocovariance coefficient as a function of an offset for entire MPEG sequence (Bond — Goldfinger)



Rys. 5. Współczynnik autokowariancji w funkcji przesunięcia dla zagregowanej sekwencji MPEG (Bond — Goldfinger)

Fig. 5. Autocovariance coefficient as a function of an offset for aggregated MPEG sequence (Bond — Goldfinger)

#### 4. Modelowanie strumieni danych o zmiennym natężeniu

Próbie modelowania strumieni danych o zmiennym natężeniu, na przykład takich, jakie opisano powyżej, można podjąć w konwencji łańcuchów Markowa. W tym celu trak-

tuje się wejściowy strumień danych jako proces Poissona modulowany łańcuchem Markowa (*Markov Modulated Poisson Process, MPPP*). Koncepcja takiego procesu polega na tym, że przez pewien okres trwający średnio  $\frac{1}{\pi_i}$  zgodnie z rozkładem wykładniczym (czas trwania  $i$ -tej fazy) źródło generuje porcje danych zgodnie z rozkładem Poissona z intensywnością  $\lambda_i$ . Po upływie czasu przeznaczanego na  $i$ -tą fazę źródło przechodzi do kolejnej, np.  $j$ -tej fazy. Czasowy przebieg natężenia strumienia danych może zostać skwantowany na pewną liczbę poziomów odpowiadającą liczbie faz źródła modulowanego oraz można na jego podstawie oszacować długość trwania poszczególnych faz.

Najprostszy przypadek takiego procesu występuje dla źródła dwustanowego. W pierwszej fazie intensywność pracy źródła wynosi  $\lambda_1$ , a w fazie drugiej wynosi  $\lambda_2$ . W szczególności, źródłem dwustanowym jest źródło ON/OFF — pracuje ono tylko w fazie ON, a w fazie OFF — milczy.

Niech intensywność przejścia między pierwszą a drugą fazą źródła wynosi  $q_{12}$ , a między drugą a pierwszą —  $q_{21}$ . Wówczas macierz generatora łańcucha modulującego jest macierzą  $Q$ , w której  $q_{11} = -q_{12}$  i  $q_{22} = -q_{21}$ . Prawdopodobieństwa stanów źródła w stanie nieustalonym określa rozwiązanie układu równań różniczkowych liniowych:

$$\frac{d\pi}{dt} = \pi Q.$$

Dla znalezienia tego rozwiązania należy rozważyć dwa przypadki w zależności od możliwych stanów początkowych:

A) dla warunku początkowego  $\pi_1(0) = 1, \pi_2(0) = 0$ :

$$\begin{aligned}\pi_1(t) &= \frac{q_{12}}{q_{12} + q_{21}} e^{-(q_{12} + q_{21})t} + \frac{q_{21}}{q_{12} + q_{21}} \\ \pi_2(t) &= -\frac{q_{12}}{q_{12} + q_{21}} e^{-(q_{12} + q_{21})t} + \frac{q_{12}}{q_{12} + q_{21}}\end{aligned}$$

B) dla warunku początkowego  $\pi_1(0) = 0, \pi_2(0) = 1$ :

$$\begin{aligned}\pi_1(t) &= -\frac{q_{21}}{q_{12} + q_{21}} e^{-(q_{12} + q_{21})t} + \frac{q_{21}}{q_{12} + q_{21}} \\ \pi_2(t) &= \frac{q_{21}}{q_{12} + q_{21}} e^{-(q_{12} + q_{21})t} + \frac{q_{12}}{q_{12} + q_{21}}\end{aligned}$$

Przy oznaczeniach  $\gamma = q_{12} + q_{21}$ ,  $\gamma_{12} = \frac{q_{12}}{\gamma}$ ,  $\gamma_{21} = \frac{q_{21}}{\gamma}$

$$\pi_1^{(A)} = \gamma_{12} e^{-\gamma t} + \gamma_{21}; \quad \pi_2^{(A)} = -\gamma_{12} e^{-\gamma t} + \gamma_{21}$$

$$\pi_1^{(B)} = -\gamma_{21} e^{-\gamma t} + \gamma_{21}; \quad \pi_2^{(B)} = \gamma_{21} e^{-\gamma t} + \gamma_{12}$$

Na tej podstawie można obliczyć autokorelację procesu modulowanego [6]:

$$\begin{aligned} R(t, t + \tau) &= E(\Lambda(t)\Lambda(t + \tau)) = \\ &= \pi_1(t)\lambda_1\pi_1(t + \tau|\pi_1(t) = 1)\lambda_1 + \pi_1(t)\lambda_1\pi_2(t + \tau|\pi_1(t) = 1)\lambda_2 + \\ &\quad + \pi_2(t)\lambda_2\pi_1(t + \tau|\pi_2(t) = 1)\lambda_1 + \pi_2(t)\lambda_2\pi_2(t + \tau|\pi_2(t) = 1)\lambda_2 \end{aligned}$$

Jeżeli proces startuje z warunku A, to:

$$\begin{aligned} R^{(A)}(t, t + \tau) &= \lambda_1^2\pi_1^{(A)}(t)\pi_1^{(A)}(\tau) + \lambda_1\lambda_2\pi_1^{(A)}(t)\pi_2^{(A)}(\tau) + \\ &\quad + \lambda_1\lambda_2\pi_2^{(A)}(t)\pi_1^{(B)}(\tau) + \lambda_2^2\pi_2^{(A)}(t)\pi_2^{(B)}(\tau) \end{aligned}$$

Wówczas średnia intensywność strumienia wynosi  $E^{(A)}(\Lambda(t)) = \lambda_1\pi_1^{(A)}(t) + \lambda_2\pi_2^{(A)}(t)$ ,  $E^{(A)}(\Lambda(t + \tau)) = \lambda_1\pi_1^{(A)}(t + \tau) + \lambda_2\pi_2^{(A)}(t + \tau)$ , a jego wariancja  $\text{Var}^{(A)}(\Lambda(t)) = R^{(A)}(t, t) - (E^{(A)}(\Lambda(t)))^2 = \lambda_1^2\pi_1^{(A)}(t) + \lambda_2^2\pi_2^{(A)}(t)$ .

Autokowariancja procesu modulowanego zostanie policzona na podstawie wzoru [6]:

$$C^{(A)}(t, t + \tau) = R^{(A)}(t, t + \tau) - E^{(A)}(\Lambda(t))E^{(A)}(\Lambda(t + \tau)).$$

W stanie ustalonym, tj. dla  $t \rightarrow 0$ , istnieje jedno rozwiązanie ustalone  $\pi$ , do którego dąży rozwiązanie  $\pi(t)$ , niezależnie od warunku początkowego. Zatem, gdy  $t \rightarrow 0$ , to  $C(t, t + \tau) \rightarrow C(\tau)$  [3].

Rozwiązanie w stanie ustalonym  $\pi = [\pi_1, \pi_2]$  wynosi:

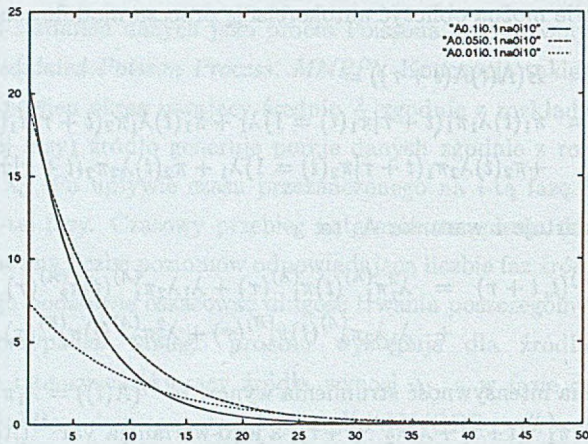
$$\pi_1 = \frac{q_{21}}{q_{12} + q_{21}}, \quad \pi_2 = \frac{q_{12}}{q_{12} + q_{21}},$$

zatem autokorelacja, której wartość zależy już tylko od przesunięcia  $\tau$ , wynosi:

$$R(\tau) = \pi_1\lambda_1\pi_1^{(A)}(\tau)\lambda_1 + \pi_1\lambda_1\pi_2^{(A)}(\tau)\lambda_2 + \pi_2\lambda_2\pi_1^{(B)}(\tau)\lambda_1 + \pi_2\lambda_2\pi_2^{(B)}(\tau)\lambda_2$$

i jest kombinacją liniową funkcji wykładniczych. Powoduje to jej relatywnie szybki spadek, co oznacza, że w przypadku takiego procesu można mówić tylko o samopodobieństwie krótkodystansowym (*short-range self-similarity*) [11]. Przykładowe przebiegi funkcji autokowariancji dla źródła dwustanowego typu ON/OFF przedstawione są na rysunku 6. Dla krzywych przedstawionych na tym wykresie parametry źródła są następujące: dla każdego przypadku intensywność w fazie ON — 0, intensywność w fazie OFF — 10.0 oraz czasy trwania faz odpowiednio a) po 10.0 każda faza, b) faza ON — 20.0, faza OFF — 10.0, c) faza ON — 100.0, faza OFF — 10.0.

W zależności od parametrów źródła, funkcja autokowariancji utrzymuje się przez różny czas powyżej zera, aczkolwiek jej spadek jest relatywnie szybki (tj. zgodnie z eksponentą).



Rys. 6. Funkcja autokowariancji dla źródła ON/OFF

Fig. 6. Autocovariation function for an ON/OFF source

## 5. Modele analityczne w formalizmie sieci automatów stochastycznych

Metodą służącą do ułatwienia konstruowania skomplikowanego modelu Markowa, a następnie obliczenia wynikającego z niego układu równań określających prawdopodobieństwa stanów łańcucha jest metoda sieci automatów stochastycznych, w skrócie SAN (ang. *Stochastic Automata Network*), która została wprowadzona w [7], a następnie rozwijana w [1, 8, 14]. W swoich założeniach jest ona metodą modułową, tj. sieć kolejek zostaje poddana dekompozycji na elementy, dla których buduje się opis ich działania w postaci grafu z etykietowanymi krawędziami, a następnie otrzymane dla tych grafów macierze połączeń z elementami będącymi wartościami etykiet krawędzi konsoliduje się za pomocą rachunku tensorowego macierzy, w celu uzyskania pełnego opisu całej sieci, tj. macierzy deskryptora. Ze względu na swoją specyfikę SAN jest przeznaczona do opisu systemów, które mogą być traktowane jako zbiór składowych, komunikujących się ze sobą wyłącznie przy synchronizacji czy przy zmianie swojego stanu w zależności od pozostałych elementów systemu.

Idea macierzowych operacji tensorowych została szczegółowo omówiona w [14]. Niech macierz  $C \in M(d)$ , gdzie  $M(d)$  oznacza zbiór wszystkich macierzy wymiaru  $d \times d$ ,

$$\bigwedge_{1 \leq i, j \leq d} c_{ij} \in \mathbb{R}.$$

Niech  $d$  nie jest liczbą pierwszą i  $d = d^{(1)}d^{(2)}$ ,  $d^{(i)} \neq 1$ ,  $i = 1, 2$ . Wówczas macierz  $C$



może być potraktowana jako macierz  $d^{(1)} \times d^{(1)}$  bloków będących macierzami  $d^{(2)} \times d^{(2)}$ . Przy takim podejściu do elementu macierzy  $C$  można się odwoływać, podając najpierw współrzędne bloku z przedziału  $1 \dots d^{(1)}$ , a następnie — współrzędne wewnątrz bloku z przedziału  $1 \dots d^{(2)}$ . Zatem:

$$C = (c_{ij})_{i,j \in L(d^{(1)}d^{(2)})},$$

gdzie  $L(d^{(1)}d^{(2)}) = \{\hat{k} : \hat{k} = (k^{(1)}, k^{(2)}) \wedge (1 \leq k^{(1)} \leq d^{(1)}) \wedge (1 \leq k^{(2)} \leq d^{(2)})\}$ , przy założeniu że zbiór ten jest uporządkowany leksykograficznie.

Blokowy zapis macierzy można uogólnić, gdy  $d = d^{(1)}d^{(2)} \dots d^{(N)}$ . Prowadzi to do zapisu:

$$C = (c_{ij})_{i,j \in L(d^{(1)}d^{(2)} \dots d^{(N)})}, \quad (1)$$

gdzie  $L(d^{(1)}d^{(2)} \dots d^{(N)}) = \{\hat{k} : \hat{k} = (k^{(1)}, k^{(2)}, \dots, k^{(N)}) \wedge \bigwedge_{1 \leq l \leq N} 1 \leq k^{(l)} \leq d^{(l)}\}$ , przy założeniu że zbiór ten jest uporządkowany leksykograficznie.

**Definicja 5.1.** *Iloczynem tensorowym macierzy  $A \in M(d^{(1)})$  i macierzy  $B \in M(d^{(2)})$  jest macierz  $C \in M(d^{(1)}d^{(2)})$ ,  $C = A \otimes B$  taka, że  $c_{ij} = a_{(i^{(1)}j^{(2)})} b_{(i^{(2)}j^{(2)})}$ , gdzie  $\hat{i} = (i^{(1)}, i^{(2)})$ ,  $\hat{j} = (j^{(1)}, j^{(2)})$ .*

Jeżeli  $A = [a_{ij}]_{i,j=1,2,\dots,d^{(1)}}$ ,  $B = [b_{ij}]_{i,j=1,2,\dots,d^{(2)}}$ , to

$$C = A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1,d^{(1)}}B \\ a_{21}B & a_{22}B & \cdots & a_{2,d^{(1)}}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{d^{(1)},1}B & a_{d^{(1)},2}B & \cdots & a_{d^{(1)},d^{(1)}}B \end{bmatrix}$$

**Definicja 5.2.** *Sumą tensorową macierzy  $A \in M(d^{(1)})$  i macierzy  $B \in M(d^{(2)})$  jest macierz  $D \in M(d^{(1)}d^{(2)})$ ,  $D = A \oplus B$ ,  $D = A \otimes Id_{d^{(2)}} + Id_{d^{(1)}} \otimes B$ , gdzie  $Id_{d^{(k)}}$  oznacza macierz jednostkową o wymiarze  $d^{(k)} \times d^{(k)}$ ,  $k = 1, 2$ .*

Jeżeli  $A = [a_{ij}]_{i,j=1,2,\dots,d^{(1)}}$ ,  $B = [b_{ij}]_{i,j=1,2,\dots,d^{(2)}}$ , to

$$D = A \oplus B = \begin{bmatrix} \text{diag}_{d^{(2)}}(a_{11}) + B & \text{diag}_{d^{(2)}}(a_{12}) & \cdots & \text{diag}_{d^{(2)}}(a_{1,d^{(1)}}) \\ \text{diag}_{d^{(2)}}(a_{21}) & \text{diag}_{d^{(2)}}(a_{22}) + B & \cdots & \text{diag}_{d^{(2)}}(a_{2,d^{(1)}}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{diag}_{d^{(2)}}(a_{d^{(1)},1}) & \text{diag}_{d^{(2)}}(a_{d^{(1)},2}) & \cdots & \text{diag}_{d^{(2)}}(a_{d^{(1)},d^{(1)}}) + B \end{bmatrix}$$

Niech  $d = d^{(1)}d^{(2)}$  i niech  $F(d^{(1)}d^{(2)})$  oznacza zbiór funkcji określonych na zbiorze par  $L(d^{(1)}d^{(2)})$  o wartościach w zbiorze  $\mathbb{R}$ . Dalej, niech  $M(d^{(1)}, F(d^{(1)}d^{(2)}))$  oznacza zbiór macierzy wymiaru  $d^{(1)} \times d^{(1)}$ , których współczynniki należą do  $F(d^{(1)}d^{(2)})$ , a zatem — gdy przyjąć, że  $Q \in M(d^{(1)}, F(d^{(1)}d^{(2)}))$  — są postaci  $q_{i,j}(\hat{i})$ ,  $\hat{i} \in L(d^{(1)}d^{(2)})$ . Jako współczynniki  $(i, j)$  macierzy  $Q$  należy więc rozumieć wartość funkcji  $q_{i,j}, q_{i,j} : L(d^{(1)}d^{(2)}) \rightarrow \mathbb{R}$ , dla argumentu  $\hat{i}$ .

**Definicja 5.3.** Niech rozważane będą macierze  $A \in M(d^{(1)}, F(d^{(1)}d^{(2)}))$  i  $B \in M(d^{(2)}, F(d^{(1)}d^{(2)}))$ . Wówczas uogólniony iloczyn tensorowy macierzy  $A \otimes B$  jest zdefiniowany jako  $C \in M(d^{(1)}d^{(2)})$ , gdzie:  $c_{ij} = a_{i^{(1)}j^{(1)}}(\hat{i})b_{i^{(2)}j^{(2)}}(\hat{i})$ ,  $\hat{i}, \hat{j} \in L(d^{(1)}d^{(2)})$ ,  $\hat{i} = (i^{(1)}, i^{(2)})$ ,  $\hat{j} = (j^{(1)}, j^{(2)})$ .

W sensie zmodyfikowanej definicji ogólnej przedstawionej w [14, 17] generator globalny dla modelu w czasie ciągłym przedstawiony jest wzorem:

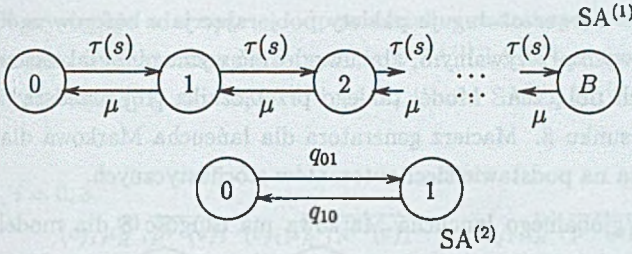
$$Q = \bigoplus_{i=1}^N Q^{(i)} + \sum_{j=1}^K \text{diag}_d(\tau_j(s)) \left( \bigotimes_{i=1}^N (S_j^{(i)})_p - \bigotimes_{i=1}^N (R_j^{(i)})_p \right), \quad (2)$$

gdzie macierze składowe mogą zawierać elementy funkcyjne.

Metoda sieci automatów stochastycznych pozwala na praktycznie natychmiastowe skonstruowanie odpowiedniego modelu dla procesu Poissona modulowanego łańcuchem Markowa oraz podaje algorytm wyznaczenia elementów odpowiadającego mu generatora (por. [15]). Sieć przedstawiona na rysunku 7 odpowiada źródłu poissonowskiemu modulowanemu łańcuchem dwustanowym. Źródło to emituje pakiety danych do bufora, w którym może się zmieścić  $B$  pakietów. Automat  $SA^{(2)}$  odpowiada łańcuchowi modulującemu, natomiast automat  $SA^{(1)}$  przedstawia stan bufora, tzn. liczbę znajdujących się w nim w danym momencie pakietów. Pakiety opuszczają bufor zgodnie z rozkładem wykładniczym z intensywnością  $\mu$ , natomiast napływają do niego z intensywnością  $\tau(s)$ ,  $s = (s^{(1)}, s^{(2)})$ , która zależy od stanu automatu  $SA^{(2)}$  i  $\tau(s^{(1)}, 0) = \lambda_1$ , a  $\tau(s^{(1)}, 1) = \lambda_2$ .

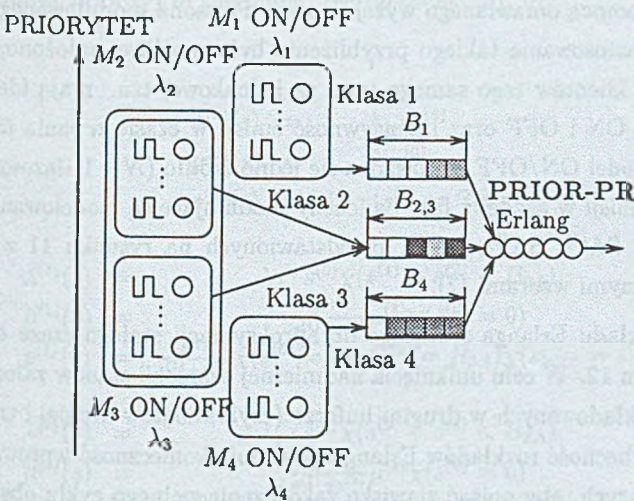
## 6. Model przełącznika sieciowego dla transmisji multimedialnych

Transmisje obrazów ruchomych występują w rzeczywistych sieciach jednocześnie ze strumieniami o innym charakterze. Strumień pakietów na wejściu przełącznika sieciowego jest wypadkową różnych strumieni. Przykładowa analiza zostanie przeprowadzona dla przełącznika sieci. Na strumień danych na jego wejściu składają się transmisje: pakietów



Rys. 7. Sieć automatów stochastycznych źródła dwustanowego  
 Fig. 7. Stochastic automata network for a two-states source

sygnalizacyjnych, pakietów wideo, pakietów głosu i pakietów transmisji plików [10, 13]. Każdy z tych typów transmisji ma inne wymagania wobec zestawianego połączenia i żąda dla siebie odpowiedniej jakości usług QoS (ang. Quality of Service). Ruch pakietów sygnalizacyjnych musi być ograniczany z uwagi na występowanie strat i opóźnień, ruch pakietów wideo i głosu wymaga minimalizowania opóźnień, a przesył plików jest najmniej uprzywilejowany.



Rys. 8. Model kolejkowy przełącznika sieciowego  
 Fig. 8. A queuing model of the multiservice switch

Pakiety należące do trzech typów ruchu składowane są w przełączniku w trzech odrębnych buforach. Serwer obsługuje pakiety pobierając je z buforów zgodnie z regulaminem priorytetowym przerywalnym, aby uwzględnić wymagania jakościowe uzgodnione dla poszczególnych połączeń. Model takiego przełącznika [16] przedstawiony jest jako sieć kolejek na rysunku 8. Macierz generatora dla łańcucha Markowa dla tego modelu jest skonstruowana na podstawie sieci automatów stochastycznych.

Wektor stanu globalnego łańcucha Markowa ma długość 8 dla modelu, w którym założony jest wykładniczy czas obsługi, natomiast dla aproksymacji stałego czasu obsługi przez rozkład Erlanga długość tego wektora wynosi 11.

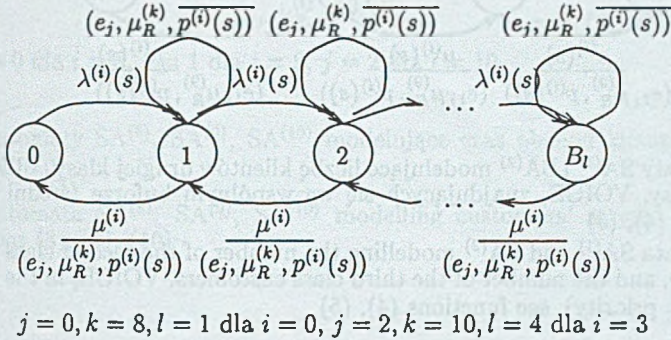
Automaty stochastyczne,  $SA^{(0)}$ ,  $SA^{(1)}$ ,  $SA^{(2)}$ ,  $SA^{(3)}$ , odpowiadające buforom, przedstawione są na rysunkach 9 i 10. Ich stan odpowiada liczbie klientów danej klasy znajdujących się w danym momencie w odpowiednim buforze (łącznie z klientem obsługiwanym). Ponieważ  $SA^{(1)}$  i  $SA^{(2)}$  reprezentują łącznie jeden bufor, to ich stany muszą spełniać warunek  $s^{(1)} + s^{(2)} < B_{2,3}$ . Regulamin priorytetowy wprowadzony jest przez transycje funkcyjne: klienci klasy drugiej i trzeciej mogą być jedynie obsługiwani wtedy, gdy nie ma w buforze klientów klasy pierwszej; klienci klasy czwartej mogą być obsługiwani jedynie wtedy, gdy w buforach nie oczekują klienci bardziej uprzywilejowani. Funkcje transycji opisują wzory (3), (4), (5), (6).

Liczba źródeł ON/OFF generujących klientów dla każdej z klas zmienia się i jest modelowana za pomocą omawianego wyżej procesu Poissona modulowanego łańcuchem Markowa. Aby zastosowanie takiego przybliżenia było możliwe, założono, że wszystkie źródła generujące klientów tego samego typu są jednakowe, tzn. mają identyczne rozkłady trwania faz ON i OFF oraz intensywność emisji w czasie trwania fazy ON. Tak więc zamiast  $N$  źródeł ON/OFF wprowadza się jedno źródło  $(N + 1)$ -fazowe o różnej intensywności transmisji w każdej z faz. Łańcuchy modulujące są modelowane za pomocą automatów  $SA^{(4)}$ ,  $SA^{(5)}$ ,  $SA^{(6)}$ ,  $SA^{(7)}$ , przedstawionych na rysunku 11 z transycjami funkcyjnymi opisanymi wzorami (7).

Automaty rozkładu Erlanga służącego do aproksymacji stałego czasu obsługi znajdują się na rysunku 12. W celu uniknięcia nadmiernej eksplozji stanów założono, że czas obsługi klientów składowanych w drugim buforze (czyli klientów drugiej i trzeciej klasy) jest identyczny. Obecność rozkładów Erlanga powoduje konieczność wprowadzenia zdarzeń synchronizujących, aby opisać zjawisko zakończenia pełnego cyklu obsługi Erlanga i opuszczenia stanowiska przez klienta. Automat  $SA^{(8)}$  odpowiada za obsługę klientów pierwszej klasy, a zdarzenie  $e_0$  inicjuje jednoczesną transycję w automacie  $SA^{(0)}$ , wzór (8),  $SA^{(9)}$  odpowiada obsłudze klientów drugiej i trzeciej klasy i występuje w nim zdarzenie  $e_1$

odpowiadające za jednoczesne tranzycje w automatach SA<sup>(1)</sup> albo SA<sup>(2)</sup>, wzór (9), a w automacie SA<sup>(10)</sup> występuje zdarzenie e<sub>2</sub> odpowiedzialne za tranzycję w SA<sup>(3)</sup>, wzór (10). Bardzo istotne jest, że jednoczesne wystąpienie w SA<sup>(1)</sup> i SA<sup>(2)</sup> zdarzenia e<sub>1</sub> jest zablokowane. Może ono wystąpić albo w SA<sup>(1)</sup> i SA<sup>(9)</sup>, albo w SA<sup>(2)</sup> i SA<sup>(9)</sup><sup>1</sup>.

SA<sup>(i)</sup>, i = 0, 3



Rys. 9. Automaty SA<sup>(0)</sup> i SA<sup>(3)</sup> modelujące liczbę klientów pierwszej klasy o najwyższym priorytecie, SIGNAL, i liczbę klientów o najniższym priorytecie, DATA, funkcje (3), (6)

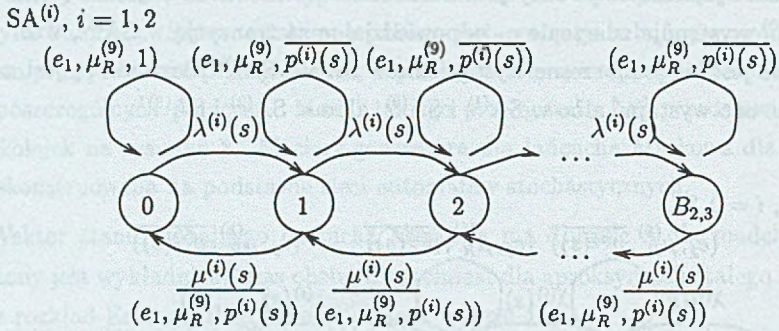
Fig. 9. Automata SA<sup>(0)</sup> and SA<sup>(3)</sup> which model the number of the first class customers, SIGNAL, (the highest priority) and the number of the last class customers, DATA, (the lowest priority) in the buffer, see functions (3), (6)

$$\begin{aligned}
 \lambda^{(0)}(s) &= \lambda_1 \cdot s^{(4)} \\
 \mu^{(0)}(s) &= \mu^{(0)} \\
 p^{(0)}(s) &= 1 \cdot \chi(s^{(8)} = R_1)
 \end{aligned}
 \tag{3}$$

$$\begin{aligned}
 \lambda^{(1)}(s) &= \lambda_2 \cdot s^{(5)} \chi(s^{(1)} + s^{(2)} < B_{2,3}) \\
 \mu^{(1)}(s) &= \frac{s^{(1)}}{s^{(1)} + s^{(2)}} \mu^{(1)} \chi(s^{(0)} = 0) \\
 p^{(1)}(s) &= \frac{s^{(1)}}{s^{(1)} + s^{(2)}} \chi(s^{(0)} = 0) \chi(s^{(9)} = R_{2,3}) \chi(s_{succ}^{(2)} = s^{(2)})
 \end{aligned}
 \tag{4}$$

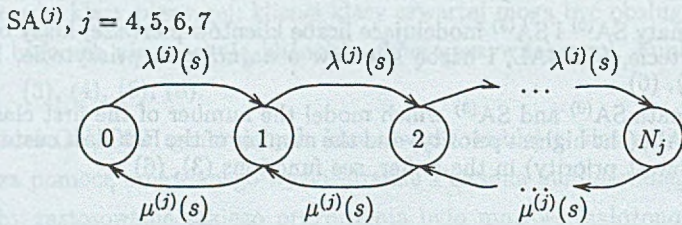
$$\begin{aligned}
 \lambda^{(2)}(s) &= \lambda_3 \cdot s^{(6)} \chi(s^{(1)} + s^{(2)} < B_{2,3}) \\
 \mu^{(2)}(s) &= \frac{s^{(2)}}{s^{(1)} + s^{(2)}} \mu^{(2)} \chi(s^{(0)} = 0) \\
 p^{(2)}(s) &= \frac{s^{(2)}}{s^{(1)} + s^{(2)}} \chi(s^{(0)} = 0) \chi(s^{(9)} = R_{2,3}) \chi(s_{succ}^{(1)} = s^{(1)})
 \end{aligned}
 \tag{5}$$

<sup>1</sup>W poniższych wzorach  $\overline{p^{(i)}(s)} = 1 - p^{(i)}(s)$ , i = 0, 1, 2, ..., 10;  $\chi(a) = 1$ , jeśli a jest prawdziwe i  $\chi(a) = 0$  w przeciwnym wypadku.



Rys. 10. Automaty SA<sup>(1)</sup> i SA<sup>(2)</sup> modelujące liczbę klientów drugiej klasy, VIDEO, i trzeciej klasy, VOICE, znajdujących się we wspólnym buforze (średni priorytet), funkcje (4), (5)

Fig. 10. Automata SA<sup>(1)</sup> and SA<sup>(2)</sup> modelling the number of the second class customers, VIDEO, and the number of the third class customers, VOICE, in the buffer (the medium priority), see functions (4), (5)



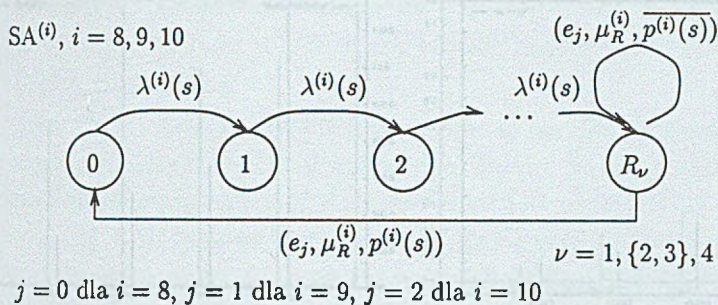
Rys. 11. Automaty SA<sup>(j)</sup>, j = 4, 5, 6, 7 modelujące liczbę aktywnych źródeł dla każdej z klas, aproksymacja MMPP, funkcje (7)

Fig. 11. Automata SA<sup>(j)</sup>, j = 4, 5, 6, 7 modelling active source number for each class respectively, MMPP approximation, see functions (7)

$$\begin{aligned} \lambda^{(3)}(s) &= \lambda_4 \cdot s^{(7)} \\ \mu^{(3)}(s) &= \mu^{(3)} \chi(s^{(0)} = 0) \chi(s^{(1)} = 0) \chi(s^{(2)} = 0) \\ p^{(3)}(s) &= 1 \cdot \chi(s^{(0)} = 0) \chi(s^{(1)} = 0) \chi(s^{(2)} = 0) \chi(s^{(10)} = R_4) \end{aligned} \tag{6}$$

$$\begin{aligned} \lambda^{(j)}(s) &= \alpha_{\text{OFF}}^{(j-4)} (N_j - s^{(j)}), \quad j = 4, 5, 6, 7 \\ \mu^{(j)}(s) &= \alpha_{\text{ON}}^{(j-4)} s^{(j)}, \quad j = 4, 5, 6, 7 \end{aligned} \tag{7}$$

$$\begin{aligned} \lambda^{(8)}(s) &= \mu^{(8)} \chi(s^{(0)} \neq 0) \\ p^{(8)} &= 1 \cdot \chi(s^{(0)} \neq 0) \end{aligned} \tag{8}$$



Rys. 12. Automaty  $SA^{(8)}$ ,  $SA^{(9)}$ ,  $SA^{(10)}$  modelujące czas obsługi klientów, funkcje (8), (9), (10)

Fig. 12. Automata  $SA^{(8)}$ ,  $SA^{(9)}$ ,  $SA^{(10)}$  modelling customers' service time, see functions (8), (9), (10)

$$\begin{aligned} \lambda^{(9)}(s) &= \mu_R^{(9)} \chi(s^{(0)} = 0) \chi(s^{(1)} \neq 0 \vee s^{(2)} \neq 0) \\ p^{(9)} &= 1 \cdot \chi(s^{(1)} \neq 0 \vee s^{(2)} \neq 0) \end{aligned} \quad (9)$$

$$\begin{aligned} \lambda^{(10)}(s) &= \mu_R^{(10)} \chi(s^{(0)} = 0) \chi(s^{(1)} = 0) \chi(s^{(2)} = 0) \chi(s^{(3)} \neq 0) \\ p^{(10)} &= 1 \cdot \chi(s^{(0)} = 0) \chi(s^{(1)} = 0) \chi(s^{(2)} = 0) \chi(s^{(3)} \neq 0) \end{aligned} \quad (10)$$

## 7. Wyniki numeryczne dla zaproponowanego modelu

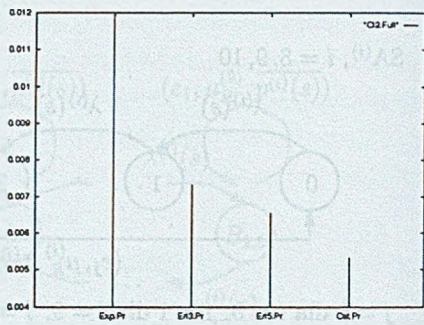
Model A składa się z 4 grup, z których każda zawiera 2 źródła ON/OFF. Długości buforów są następujące:  $B_1 = 1$ ,  $B_{2,3} = 3$ ,  $B_4 = 2$ , a pozostałe parametry modelu:  $\alpha_{1,OFF} = 1.204819$ ,  $\alpha_{1,ON} = 5.882353$ ,  $\alpha_{2,OFF} = 2.0$ ,  $\alpha_{2,ON} = 2.0$ ,  $\alpha_{3,OFF} = 1.333333$ ,  $\alpha_{3,ON} = 4.0$ ,  $\alpha_{4,OFF} = 1.1764706$ ,  $\alpha_{4,ON} = 6.666667$ ,  $\lambda_1 = 3.0$ ,  $\lambda_2 = 2.0$ ,  $\lambda_3 = 4.0$ ,  $\lambda_4 = 10.0$ , gdzie  $\alpha_{i,OFF}$  ( $\alpha_{i,ON}$ ) oznaczają intensywności zakończenia faz dla  $i$ -tej grupy źródeł.

Przy założonej jednakowej średniej czasu obsługi 0.1 zbadano trzy modele różniące się typem obsługi A.Exp (wykładniczy), A.Er13, A.Er15 (Erlang 3 rzędu, 5 rzędu), w których obciążenie na wejściu wynosiło około 80%. Liczba stanów dla tych modeli wynosiła 4860, 63'504 i 245'916, a ich wybrane wyniki, uzyskane z wygenerowanego na podstawie sieci automatów stochastycznych układu równań, rozwiązanego za pomocą metody Arnoldiego [12], tj. prawdopodobieństwa, że bufory są pełne są zestawione z wynikami



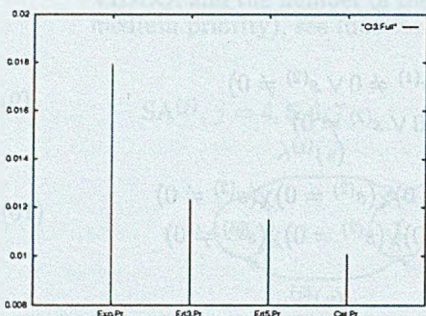
Rys. 13. Prawdopodobieństwo pełnego bufora dla klientów pierwszej klasy; model A

Fig. 13. Probability of full buffer for the first class customers; model A



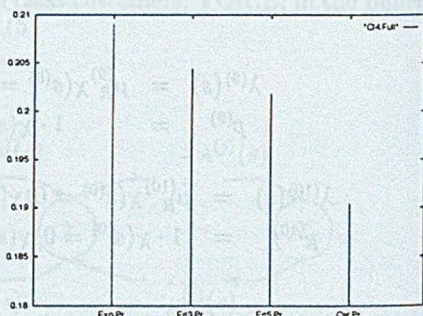
Rys. 14. Prawdopodobieństwo pełnego bufora dla klientów drugiej klasy; model A

Fig. 14. Probability of full buffer for the second class customers; model A



Rys. 15. Prawdopodobieństwo pełnego bufora dla klientów trzeciej klasy; model A

Fig. 15. Probability of full buffer for the third class customers; model A



Rys. 16. Prawdopodobieństwo pełnego bufora dla klientów czwartej klasy; model A

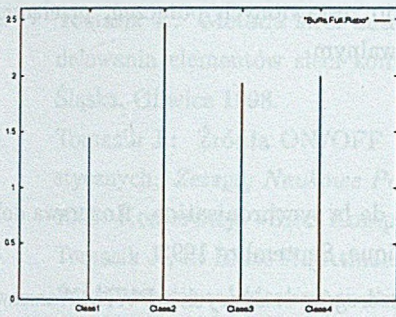
Fig. 16. Probability of full buffer for the fourth class customers; model A

modelu symulacyjnego, w którym czas obsługi jest stały, zostały przedstawione na rysunkach 13, 14, 15, 16. Zwiększanie liczby faz Erlanga powoduje, że otrzymane wyniki zbliżają się do tych otrzymanych dla stałego czasu obsługi (są bardziej optymistyczne niż uzyskane dla czasu obsługi zgodnego z rozkładem wykładniczym).

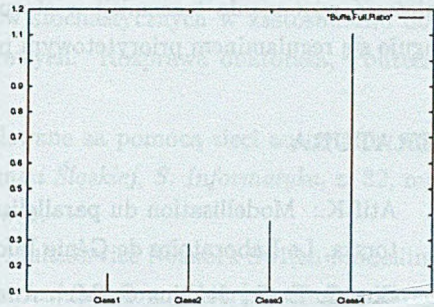
Dłożenie po jednym źródle do każdej grupy do modelu A.Er13 powoduje wygenerowanie łańcucha o 200'704 stanach (model B.Er13). Zwiększone w ten sposób obciążenie przełącznika można zaobserwować na rysunku 17 zestawiającym, ile razy prawdopodobieństwa pełnego bufora dla modelu B.Er13 są większe niż dla modelu A.Er13.

Model A.Er15 został rozszerzony przez dodanie po jednym miejscu w pierwszym





Rys. 17. Stosunek prawdopodobieństw modeli B.Erl3/A.Erl3, że bufory są pełne dla poszczególnych klas  
 Fig. 17. Proportion of full buffers probabilities for the all classes; Models: B.Erl3/A.Erl3



Rys. 18. Stosunek prawdopodobieństw modeli C.Erl5/A.Erl5, że bufory są pełne dla poszczególnych klas  
 Fig. 18. Proportion of full buffers probabilities for the all classes; Models: C.Erl5/A.Erl5

i w drugim buforze, co doprowadziło do skonstruowania łańcucha o 695'871 stanach (model C.Erl5). Porównanie stosunku prawdopodobieństw pełnych buforów dla obu tych modeli przedstawione jest na rysunku 18. Te prawdopodobieństwa są istotnie mniejsze w rozszerzonym modelu dla klientów pierwszej, drugiej i trzeciej klasy, natomiast rosną dla klientów klasy najmniej uprzywilejowanej.

## 8. Wnioski

Wprowadzenie źródeł ON/OFF do modelowania elementów sieci komputerowych pozwala na uwzględnienie w wejściowych strumieniach pakietów funkcji autokorelacji utrzymującej się przez pewien czas na poziomie różnym od zera. W modelach Markowa nie można otrzymać efektu długookresowego samopodobieństwa, gdyż funkcja współczynnika autokowariancji jest kombinacją liniową funkcji eksponencjalnych, ale można wprowadzić zależność krótkookresową. Tego typu strumienie są efektem procesu Poissona modulowanego łańcuchem Markowa (MMPP) i poddają się efektywnemu opisowi w formalizmie sieci automatów stochastycznych. Wówczas jeden z automatów tworzących sieć odpowiada modulującemu łańcuchowi Markowa.

Metoda Sieci Automatów Stochastycznych z tranzycjami funkcyjnymi umożliwia modelowanie złożonych regulaminów obsługi oraz omawianych strumieni pakietów, co zostało przedstawione na przykładzie przełącznika sieciowego obsługującego transmisje o różnym charakterze (transmisje kontrolne, obrazów i głosu oraz przesył plików). Aby

zróżnicować wymagania transmisji w stosunku do zestawionych połączeń, przełącznik posługuje się regulaminem priorytetowym przerywalnym.

## LITERATURA

1. Atif K.: Modélisation du parallélisme et de la synchronisation. Rozprawa doktorska, Le Laboratoire de Génie Informatique, Septembre 1992.
2. Box G. E. P., Jenkins G. M.: Analiza szeregów czasowych. PWN, Warszawa 1983.
3. Czachórski T.: Modele kolejkowe systemów komputerowych. Skrypt nr 1844. Wydawnictwo Politechniki Śląskiej, Gliwice 1994.
4. Lam S. S., Chow S., Yau D. K. Y.: An Algorithm for Lossless Smoothing of MPEG Video. In *ACM SIGCOMM*, London 1994.
5. Onvural R. O.: Asynchronous Transfer Mode Networks; Performance Issues. Artech House, Boston-London 1994.
6. Papoulis A.: Probability, Random Variables and Stochastic Processes. McGraw-Hill International Book Company, 10 wydanie, 1981.
7. Plateau B.: On the Stochastic Structure of Parallelism and Synchronization Models for Distributed Algorithms. *Proc. ACM Sigmetrics*, 13(2):147-154, 1985.
8. Quessette F.: De nouvelles méthodes de résolution pour l'analyse quantitative des systèmes parallèles et des protocoles. Rozprawa doktorska, Université de Paris-Sud, Centre d'Orsay, Novembre 1994.
9. Rose O., Frater M. R.: Impact of MPEG Video Traffic on an ATM Multiplexer. In R. Puigjaner, editor, *High Performance Networking*, volume 6, pages 157-168. Chapman & Hall, 1995.
10. Roussos J.K., Economou E.G., Philokyprou G.: Congestion control protocols for interconnected LANs supporting voice and data traffic. Technical report. Université d'Athènes, January 1992.
11. Stallings W.: High-Speed Networks; TCP/IP and ATM Design Principles. Prentice Hall, Upper Saddle River, New Jersey 1998.
12. Steward W. J.: Introduction to the Numerical Solution of Markov Chains. Princeton University Press, Princeton, New Jersey 1994.
13. Stuttgart H.J.: Network Evolution and Multimedia Communication. *IEEE MultiMedia*, 1995.

14. Tomasiak J.: Metoda sieci automatów stochastycznych w zastosowaniu do modelowania elementów sieci komputerowych. Rozprawa doktorska, Politechnika Śląska, Gliwice 1998.
15. Tomasiak J.: Źródła ON/OFF modelowane za pomocą sieci automatów stochastycznych. *Zeszyty Naukowe Politechniki Śląskiej, S. Informatyka*, z. 32, marzec 1998. Konferencja »Sieci Komputerowe«.
16. Tomasiak J., El Biaze H., Atmaca T.: Multiservice Network Switch Modelling. In *3rd International Meeting on the Numerical Solution of Markov Chains*, Zaragoza, Spain, September 1999. The paper submitted.
17. Tomasiak J., Quesette F.: Another approach to the stochastic automata network formalism. June 1997. PAPM97, Fifth International Workshop on Process Algebra and Performance Modelling.

Recenzent: Dr inż. Ewa Starzewska-Karwan

Wpłynęło do Redakcji 31 marca 1999 r.

#### Abstract

The paper discusses statistical features of data streams in broadband networks. The analyzed streams resulted from video transmissions, compressed and coded according to the MPEG standard. Autocovariance as a basic feature for estimating a degree of self-similarity was viewed. Further, it is discussed how the autocovariation function of a data stream generated by a certain group of sources modelled by Markov chains looks like and what kind of real streams may be modelled this way. A tool used for creating Markov chain models was the stochastic automata networks method (SAN). This method leads directly to a compact form of a matrix generator. As an example, a model of a multi-service network switch constructed with the Stochastic Automata Network formalism is presented. The incoming packet stream is a superposition of streams of different types (the signaling traffic, the real-time audio and video transmissions, the best-effort traffic).