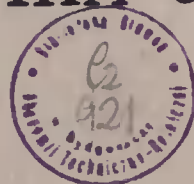


AKADEMIA TECHNICZNO-ROLNICZA
IM. JANA I JĘDRZEJA ŚNIADECKICH
W BYDGOSZCZY

ZESZYTY NAUKOWE 106

TELEKOMUNIKACJA –
ELEKTRONIKA 5



BYDGOSZCZ – 1983

AKADEMIA TECHNICZNO-ROLNICZA
IM. JANA I JĘDRZEJA ŚNIADECKICH
W BYDGOSZCZY

ZESZYTY NAUKOWE 106

TELEKOMUNIKACJA –
ELEKTRONIKA 5



BYDGOSZCZ – 1983

PRZEWODNICZĄCY KOMITETU REDAKCYJNEGO
doc. dr hab. Juliusz Skonieczny

REDAKTOR NAUKOWY
doc. dr hab. Lech J. Weiss

OPRACOWANIE REDAKCYJNE I TECHNICZNE
mgr Halina Koziolkiewicz, Alfons Grzenkiewicz

Wydano za zgodą Rektora
Akademii Techniczno-Rolniczej
w Bydgoszczy

ISSN-0209-0589

WYDAWNICTWA UCZELNIANE AKADEMII TECHNICZNO-ROLNICZEJ
w BYDGOSZCZY

Wyd. I. Nakład 100 + 50. Ark. wyd. 6,95, Ark. druk. 7,75, Papier kl. V drukowy B1
Oddano do druku 6.12.83. Druk ukończono w grudniu 83. Zam. nr 597/83
Cena 84 zł MNSzWiTK-10
UZMP – BYDGOSZCZ

Nr 100 P-201
507-70 A

S p i s t r e ś c i

	str.
1. Ewa Szlachta, Mirosław Konieczny - Wstęp do sieci Petri	5
2. Ewa Szlachta, Mirosław Konieczny - Zastosowanie sieci Petri do modelowania układu realizującego regułę dostępu w prostym systemie telekomunikacyjnym	16
3. Antoni Zabłudowski, Andrzej Sobólski - Ocena efektywności rekurencyjnego algorytmu wyznaczania miar niezawodnościowego grafu probabilistycznego	25
4. Antoni Zabłudowski, Bożydar Dubalski - Pseudolosowy generator prostego strumienia Poissona	34
5. Marian Molski, Bożydar Dubalski - Koncepcja generatora cyfrowego wektorów losowych o składowych niezależnych i dowolnych rozkładach	44
6. Jerzy Piotr Pawlak - Funkcjonalna struktura procesu awaryjnego zamykania ruchu wewnętrznego w koncentratorze CSADz centrali elektronicznej E-10	53
7. Józef Pieprzyk - Autentyfikacja informacji w systemach i sieciach komputerowych	61
8. Andrzej Błaszczak, Ireneusz Żmidziński - Ocena komutacji zintegrowanej w modelu sieci telekomunikacyjnej	70
9. Tadeusz Uhl - Wyznaczanie okresu uaktualniania tablic doboru tras w sieciach z komutacją pakietów	80
10. Krzysztof Chmara - Charakterystyki węzła sieci teleinformatycznej o ograniczonej pamięci buforowej	90
11. Tadeusz Wysocki jr - O pewnym sposobie modulacji przebiegu harmonicznego sygnałem cyfrowym	99
12. Lech J. Weiss, Tadeusz Wysocki - Sposób cyfrowego pomiaru maksymalnej różnicy chwilowych wartości przebiegów odkształconych ...	111
13. Andrzej Pawlak, Jan Piątkowski, Roman Wiatr - Metodyka doboru silnika stemplownicy statyczno-dynamicznej	117



Ewa Szlachta
 Mirosław Konieczny

WSTĘP DO SIECI PETRI

W pracy przedstawiono podstawowe pojęcia opisujące sieci Petri. Pokazano, że stanowią doskonałą metodę opisu systemów działających równocześnie i asynchronicznie. Przedstawiono również dwa sposoby prowadzenia analizy działania badanego modelu.

1. WSTĘP

Sieci Petri powstały w wyniku poszukiwań prostej i jednocześnie skutecznej metody służącej do opisu systemów działających asynchronicznie i równocześnie. Jako model matematyczny systemu sieć Petri reprezentuje jego działanie w postaci układu równań i nierówności, co umożliwia opracowanie komputerowych metod analizy i syntezy systemu. Sieci Petri dobrze opisują zarówno systemy, w których procesy zachodzą sekwencyjnie, jak i równocześnie. Pozwalają również na modelowanie systemów na różnych poziomach abstrakcji, tak że modele na niższym poziomie opisu mogą być budowane na podstawie modeli uzyskanych na wyższym poziomie. Wielopoziomowe modelowanie systemu przy zastosowaniu tego samego narzędzia ułatwia systematyczną analizę i projektowanie, gdyż opracowane metody badania i weryfikacji modelu mogą być z powodzeniem stosowane na różnych poziomach abstrakcji.

2. SIECI PETRI

Definicja 1. Siecią Petri nazywamy uporządkowaną trójką $S = \langle P, T, R \rangle$, gdzie:

- P jest skończonym, niepustym zbiorem miejsc,
- T jest skończonym, niepustym zbiorem przejść, $P \cap T = \emptyset$
- R jest relacją binarną na $P \cup T$, taką że

$$R \subseteq P \times T \cup T \times P$$

oraz

$$\bigwedge_{t \in T} \bigvee_{p_1, p_2 \in P} /p_1, t/ \in R \wedge /t, p_2/ \in R$$

Miejsce p jest miejscem wejściowym /lub wyjściowym/ przejścia t , jeśli $/p, t/ \in R$ lub $/t, p/ \in R$. Zbiór wszystkich miejsc wejściowych /wyjściowych/

danego przejścia t oznaczamy jako $\text{Inp}/t/$ / $\text{Out}/t//$, a więc

$$\begin{aligned} \text{Inp}/t/ &= \left\{ p \in P \mid /p, t/ \in R \right\} \\ \text{Out}/t/ &= \left\{ p \in P \mid /t, p/ \in R \right\} \end{aligned}$$

Podobnie:

$$\begin{aligned} \text{Inp}/p/ &= \left\{ t \in T \mid /t, p/ \in R \right\} \\ \text{Out}/p/ &= \left\{ t \in T \mid /p, t/ \in R \right\} \end{aligned}$$

Miejsce p jest podzielne, jeśli jest ono miejscem wejściowym więcej niż jednego przejścia. Miejsce p jest miejscem z wolnym wyborem, jeśli jest podzielne i jedyne w zbiorze miejsc wejściowych/ dla wszystkich przejść dzielących je.

Definicja 2. Znakowaną siecią Petri nazywamy uporządkowaną parę $M = /S, m_0/$, gdzie:

- S jest siecią Petri,
- m_0 jest pewnym znakowaniem początkowym, tj. $m_0 : P \rightarrow \mathbb{N}$, gdzie \mathbb{N} jest zbiorem liczb naturalnych.

Niech dowolna funkcja $m : P \rightarrow \mathbb{N}$ będzie znakowaniem sieci M . Przejście t jest przygotowane przez znakowanie m , jeśli każde miejsce wejściowe tego przejścia zawiera co najmniej jeden znacznik. Zbiór wszystkich przygotowanych przejść przez dane znakowanie m oznaczamy jako $T/m/$. Każde przygotowane przejście w danym znakowaniu m może zapalić się. Kiedy przejście pali się, znacznik usuwany jest z każdego miejsca wejściowego i umieszczony w każdym z miejsc wyjściowych. Wyznacza to nowe znakowanie sieci.

Definicja 3. Znakowanie m_j jest osiągalne bezpośrednio ze znakowaniem m_i , co oznaczamy $m_i \rightarrow m_j$, jeśli istnieje przejście t przygotowane w znakowaniu m_i takie, że:

$$\bigwedge_{p \in P} m_j/p/ = \begin{cases} m_i/p/ - 1, & \text{jeśli } p \in \text{Inp}/t/ \text{ i } p \notin \text{Out}/t/ \\ m_i/p/ + 1, & \text{jeśli } p \in \text{Out}/t/ \text{ i } p \notin \text{Inp}/t/ \\ m_i/p/ & , \text{ jeśli } p \in \text{Inp}/t/ \text{ i } p \in \text{Out}/t/ \\ & p \notin \text{Inp}/t/ \text{ i } p \notin \text{Out}/t/ \end{cases}$$

Znakowanie m_j jest osiągalne ze znakowania m_i w sieci M , co oznaczamy $m_i \xrightarrow{K} m_j$, wtedy i tylko wtedy, gdy istnieje ciąg znakowań $/m_1, m_2, \dots, m_k/$ taki, że $m_1 = m_i$, $m_k = m_j$ i każde m_l jest bezpośrednio osiągalne z m_{l-1} , dla $l = 2, \dots, k$.

Zbiór $Z/M/$ osiągalnych znakowań znakowanej sieci Petri M jest zbiorem wszystkich znakowań osiągalnych ze znakowania początkowego m_0 , tj.:

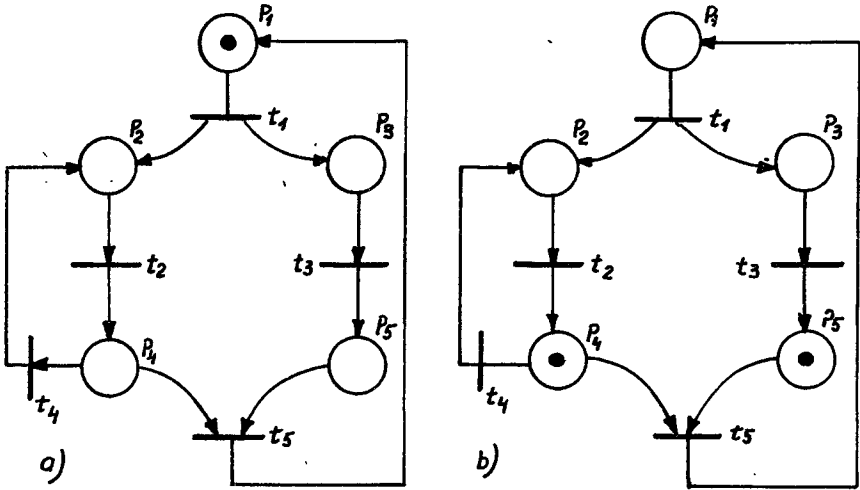
$$Z/M/ = \left\{ m \mid m_0 \xrightarrow{K} m \right\}$$

Przez dowolne znakowanie m rozumiemy wektor, którego i -ta składowa reprezentuje liczbę znaczników w miejscu p_i sieci M .

Mówimy, że znakowanie m w sieci M jest znakowaniem z wolnym wyborem, jeśli każde miejsce przygotowujące więcej niż jedno przejście jest miejscem z wolnym wyborem, tj:

$$\bigwedge_{p \in P} \bigwedge_{t_1, t_j \in T/m/ \cap \text{Out}/p/} t_1 \neq t_j \Rightarrow \text{Inp}/t_1/ = \text{Inp}/t_j/ = \{p\}$$

Rozważmy następującą sieć Petri /rys.1/.



Rys.1. Przykład sieci Petri:

a/ zdarzenia równoczesne

b/ zdarzenia wykluczające się

Na rys.1a przy znakowaniu $m = /1,0,0,0,0/$ jedynym przejściem, które może się zapalić jest przejście t_1 . Po zapaleniu się przejścia t_1 otrzymujemy znakowanie $m' = /0,1,1,0,0/$, w którym przygotowane są przejścia t_2 i t_3 . Przejścia te mogą zapalić się równocześnie. Sieć jest wtedy znakowana przez $m'' = /0,0,0,1,1/$. Rys.1b przedstawia sytuację konfliktową, która powstała w wyniku zapalenia się przejść t_2 i t_3 . Przejścia t_4 i t_5 są przygotowane, lecz zapalenie się jednego z nich uniemożliwia zapalenie drugiego /znaczniki są niepodzielne/.

Definicja 4. Miejsce p w znakowanej sieci Petri nazywa się k -ograniczonym, jeśli istnieje takie $k \in \mathbb{N}$, że:

$$m/p/ \leq k \text{ dla wszystkich } m \in Z/M/$$

Jeśli warunek ten jest spełniony dla wszystkich miejsc sieci M , sieć taką nazywamy k -ograniczoną.

Definicja 5. Sieć Petri M nazywamy bezpieczną, jeśli jest ona 1-ograniczona.

Definicja 6. Przejście t w znakowanej sieci Petri M jest żywe, jeśli dla każdego znakowania $m \in Z/M/$ istnieje znakowanie m' osiągalne z m , w którym przejście t jest przygotowane do zapalenia, tj:

$$\bigwedge_{m \in Z/M/} \bigvee_{m'} m \xrightarrow{t} m' \quad \text{i} \quad t \in T/m'$$

Znakowana sieć Petri $M = /S, m_0/$ nazywa się żywą, jeśli każde jej przejście jest żywe.

Definicja 7. Znakowana sieć Petri M jest zachowawcza, jeśli suma znaczników dla dowolnego, osiągalnego znakowania jest stała, tj:

$$\text{dla każdego } m \in Z/M/ \quad \sum_{p \in P} m/p/ = \text{const.}$$

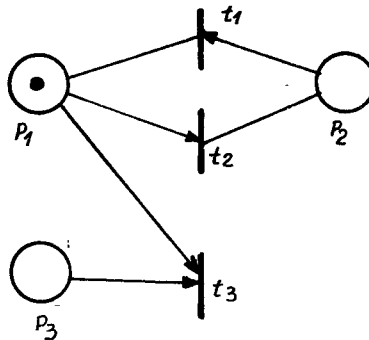
Z definicji 7 wynika, że każde przejście w sieci zachowawczej jest zachowawcze w tym sensie, że liczba miejsc wejściowych jest równa liczbie miejsc wyjściowych.

Definicja 8. Zbiór $I \subseteq P$ jest niezmienniczy, jeśli suma:

$$\sum_{p \in I} m/p/$$

jest stała dla każdego osiągalnego znakowania m i zbiór I nie ma podzbiorów właściwych, które są niezmiennicze.

Niech dana będzie następująca znakowana sieć Petri /rys.2/.



Rys.2. Znakowana sieć Petri

Przy tak zadanym znakowaniu początkowym, $m_0 = /1, 0, 0/$, tylko przejście t_2 jest przygotowane. Przejścia t_1 i t_2 są zachowawcze. Przejście t_3 nie jest zachowawcze w tym sensie, że jeśli kiedykolwiek pali się, powoduje ubytek znaczników o 2. Jednak dla dowolnego znakowania początkowego m_0 , w którym liczba znaczników w miejscu p_3 jest równa zero, przejście to nie zapali się i z tego powodu liczba znaczników w sieci jest zachowana.

Przejście t_3 i miejsce p_3 mogą być usunięte z sieci bez zmiany jej działania.

3. METODY ANALIZY DZIAŁANIA SIECI PETRI

Wykorzystując podany niżej przykład dotyczący modelu sieci Petri rozwiązującego problem wzajemnej wyłączoneści, omówimy dwie metody analizy jej działania. Metodami tymi będą:

- a/ metoda podzbiorów niezmienniczych,
- b/ metoda drzewa osiągalności.

Rozważmy system dwóch działań:

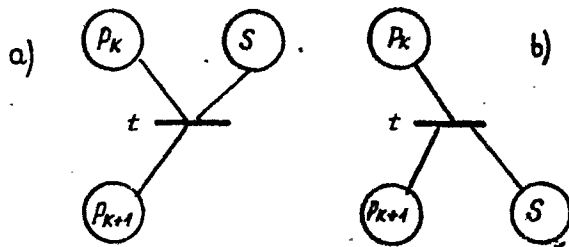
- /1/ - producent /proces wejściowy lub urządzenie wejściowe/, który uzyskane dane umieszcza w ograniczonym buforze B
- /2/ - konsument /proces obliczeniowy lub urządzenia/, który pobiera dane z buforu i działa nad nimi.

Procesy /1/ i /2/ są asynchroniczne. Należy przy tym uniemożliwić równoczesny ich dostęp do buforu, jak również dopuścić do jego nadmiaru lub niedomiaru. Zdefiniujemy operacje P i V nad specjalną zmienną S nazwaną semaforem:

P/S/ jeśli $S > 0$ wtedy $S := S - 1$, w przeciwnym przypadku czekaj

V/S/ $S := S + 1$

Modelowanie z wykorzystaniem semafora S ilustruje rys.3.



Rys.3. Przykłady modelowania z wykorzystaniem semafora;

a/ operacja P/S/

b/ operacja V/S/

Operacje P i V wykorzystujemy do rozwiązania problemu wzajemnej wyłączoneści /rys.4/. Przy tak zadany znakowaniu początkowym przygotowanymi są przejścia t_1 i t_2 . Semafor $S = 1$. Tylko jedno z tych przejść może się zapalić. Załóżmy, że pali się przejście t_1 . Wykonujemy operację P nad semaforem S, po której $S = 0$. Przejście t_2 może zapalić się wówczas, gdy wykonamy operację V nad semaforem S /zapala się przejście t_2 /.

3.1. Metoda podzbiorów niezmienniczych

Metoda ta polega na znalezieniu podzbiorów niezmienniczych sieci. Zdefiniujemy macierz incydencji dla znakowanej sieci Petri $M = /S, m_0/$.

Definicja 9. Macierzą incydencji $[C] = [C/t, p/]$ $m \times n$ znakowanej sieci Petri M nazywamy macierz określoną następująco:

$$C/t, p/ = \begin{cases} -1 & \text{jeśli } /p, t/ \in R \text{ i } /t, p/ \notin R \\ +1 & \text{jeśli } /t, p/ \in R \text{ i } /p, t/ \notin R \\ 0 & \text{jeśli zachodzą pozostałe przypadki} \end{cases}$$

gdzie:

$t \in T$, $p \in P$, m jest mocą zbioru T , n jest mocą zbioru P .

Rozważmy układ równań $C \cdot y = 0$, którego rozwiązaniem jest wektor binarny y . Zbiór miejsc odpowiadający niezerowym składowym wektora y stanowi podzbiór niezmienniczy. Dla sieci z rys.5 rozważany układ przyjmie następującą postać:

$$\begin{array}{c} t_1 \\ t_2 \\ t_3 \\ t_4 \end{array} \begin{array}{c} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \\ P_7 \end{array} \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Stąd otrzymujemy:

$$\begin{aligned} -y_1 + y_2 &= 0 \\ y_1 - y_2 - y_5 + y_7 &= 0 \\ -y_3 + y_4 + y_5 - y_7 &= 0 \\ y_3 - y_4 &= 0 \end{aligned}$$

Rozwiązaniem powyższego układu są następujące liniowo niezależne wektory:

$$y_a = /1100000/, \quad y_b = /0000101/, \quad y_c = /0011000/, \quad y_d = /0000010/$$

Wobec tego niezmienniczymi podzbiórami są:

$$\{P_1, P_2\}, \quad \{P_3, P_5\}, \quad \{P_4, P_7\}, \quad \{P_6\}$$

Dane podzbiory niezmiennicze wskazują na pewne właściwości dynamicznego zachowania się sieci:

- sieć jest ograniczona, bowiem każde miejsce należy do pewnego podzbioru niezmienniczego, a działanie swoje rozpoczyna od znakowania z ograniczoną liczbą znaczników,
- sieć jest zachowawcza /oczywiste/,
- sieć jest żywa; zgodnie z definicją sieci żywej należy pokazać, że dla każdego przejścia t i dowolnego znakowania $m \in Z/M/$ istnieje znakowanie

m' osiągalne z m , w którym przejście t jest przygotowane.

Rozważmy np. przejście t_2 i dowolne znakowanie m , w którym $m/p_2/ = 0$ lub $m/p_5/ = 0$. Jeśli $m/p_5/ = 0$, wtedy z niezmienniczości zbioru $\{p_5, p_7\}$ wynika, że $m/p_7/ = B$, jeśli $m/p_3/ = 0$, wtedy $m/p_4/ = 1$ / z niezmienniczości zbioru $\{p_3, p_4\}$ / i zapalić się może t_4 ; w przeciwnym wypadku $m/p_3/ = 1$ i pali się t_3 . Jeśli $m/p_2/ = 0$, wtedy $m/p_1/ = 1$ i pali się t_1 . Otrzymujemy znakowanie, w którym $m/p_2/ = 1$ i $m/p_5/ = 1$ i przejście t_2 jest przygotowane. Podobne rozważanie można przeprowadzić dla dowolnego przejścia t i znakowania m .

d/ wzajemna wyłączność; jeśli dwa przejścia są przejściami danego zbioru niezmienniczego i znakowanie początkowe jest takie, że suma znaczków miejsc niezmienniczych równa się 1, wtedy przejścia te się wzajemnie wykluczają i nie mogą palić się równocześnie. W myśl tego znakowanie początkowe i niezmienniczy zbiór p_6 gwarantują, że t_2 i t_3 są wzajemnie wyłączne.

e/ nadmiar i niedomiar buforu; zauważmy, że niedomiar buforu jest niemożliwy ponieważ t_3 nie pali się, jeśli bufor jest pusty $m/p_7/ = 0$. Nie wystąpi również nadmiar buforu, bowiem z niezmienniczości zbioru $\{p_5, p_7\}$ i znakowania $m/p_7/ = B$ wynika, że $m/p_5/ = 0$ i przejście t_2 nie pali się.

Zbiory niezmiennicze służą do weryfikacji zachowania się sieci Petri. Nie zawsze jednak w celu ich odszukania należy rozwiązać układ równań $C \cdot y = 0$. Czasem zbiory te są systematycznie uzyskiwane w wyniku pewnych prawideł wykorzystywanych podczas konstrukcji sieci.

3.2. Metoda drzewa osiągalności

Podstawowym podejściem do analizy sieci Petri jest użycie drzewa osiągalności. Powiemy, że $m \ll m'$, jeśli $\bigwedge_{p \in P} m/p/ \leq m'/p/$. Oznaczamy przez $d/m, t/ = m'$ znakowanie, jakie uzyskamy ze znakowania m po zapaleniu się przejścia t .

Zdefiniujemy wielkość w jako:

$$w + a = w, \quad w - a = w, \quad a < w \text{ dla wszystkich } a \in N$$

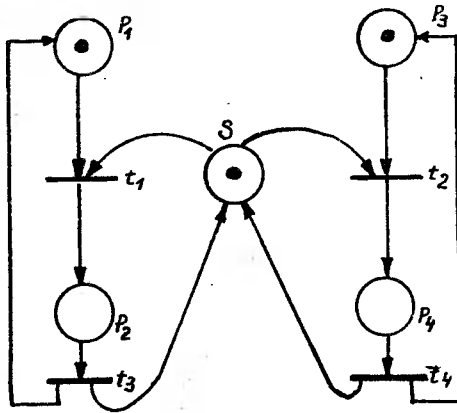
Wielkość w w stosunku do a może być symbolem nieskończoności.

Drzewo osiągalności składa się z węzłów i strzałek. Węzły oznaczane są elementami zbioru $Z/M/$, natomiast strzałki elementami zbioru T . Niech $M = /S, m_0/$ będzie znakowaną siecią Petri. Drzewo osiągalności dla tej sieci konstruujemy w następujący sposób:

a/ przyjmujemy znakowanie początkowe m_0 jako węzeł początkowy - korzeń drzewa,

b/ dla każdego węzła x w drzewie osiągalności znajdujemy zbiór $T/m/$.

Dla każdego, przygotowanego przejścia t_j przez znakowanie m w węźle x tworzymy nowy węzeł x' ze znakowaniem $d/m, t_j/$, a strzałką skierowaną, łączącą te węzły oznaczamy przez t_j . Proces ten powtarzamy jest dla wszyst-



Rys.4. Sieć Petri rozwiązująca problem wzajemnej wyłącności

Niech E reprezentuje liczbę pustych miejsc w buforze, natomiast F liczbę zajętych miejsc w buforze. Wówczas programy dla procesów /1/ i /2/ są następujące:

Semafor E : Początek /E/ = B

Semafor F : Początek /F/ = 0

Semafor M : Początek /M/ = 1

PRODUCENT:

PROD: produkowanie

P/E/

P/M/

"umieszczanie w buforze"

V/M/

V/F/

"skocz do PROD"

KONSUMENT:

KONS:

P/F/

P/M/

"wybieranie z buforu"

V/M/

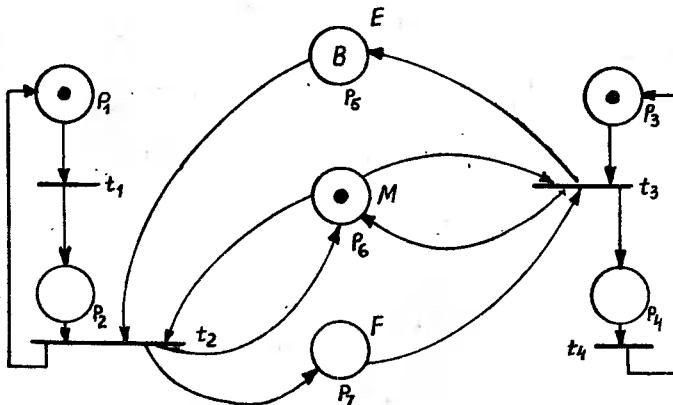
V/E/

"konsumpcja"

"skocz do KONS"

Semafor M reprezentuje wzajemną wyłącność zdarzeń.

Znakowana sieć Petri pokazana na rys.5 ilustruje pracę rozważanego systemu.



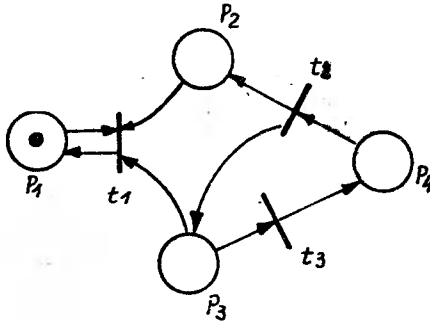
Rys.5. Znakowana sieć Petri modułująca system producent - konsument

kich nowych węzłów. Kontynuując go tworzymy całkowitą przestrzeń stanów. Ścieżka od znakowania początkowego do dowolnego węzła w drzewie odpowiada pewnej sekwencji działań.

Z tak podanej konstrukcji wynika, że może ona prowadzić do nieskończonej przestrzeni stanów. Aby określić skończone drzewo osiągalności należy ją zmodyfikować w następujący sposób:

- 1/ jeśli nowe znakowanie równe jest pewnemu już istniejącemu znakowaniu w drzewie, znakowanie to staje się węzłem końcowym,
- 2/ jeśli nowe znakowanie x jest większe od znakowania y znajdującego się w ścieżce od korzenia do węzła x , wtedy te składowe znakowania x , które są ostro większe od odpowiednich składowych znakowania y , zastępowane są symbolem w .

Jako przykład tej konstrukcji rozważmy sieć znakowaną z rys.6.



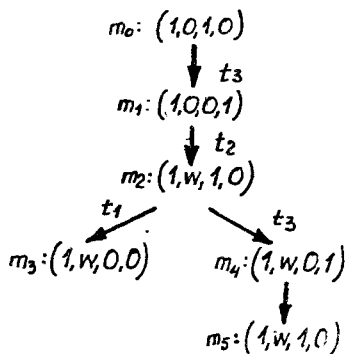
Rys.6. Znakowana sieć Petri z nieskończoną przestrzenią stanów

Znakowanie $m_0 = /1,0,1,0/$ stanowi korzeń drzewa. W tym znakowaniu przygotowane jest tylko przejście t_3 . Otrzymujemy nowy węzeł $m_1 = /1,0,0,1/$ i strzałkę od $/1,0,1,0/$ do $/1,0,0,1/$ oznaczoną przez t_3 . W znakowaniu tym zapala się przejście t_2 , w rezultacie czego otrzymujemy znakowanie $m_2 = /1,1,1,0/$. Teraz, ponieważ $/1,1,1,0/ \gg /1,0,1,0/$, drugi składnik zastępujemy symbolem w . Wynika to z faktu, że ciąg t_3, t_2 możemy zapalać dowolną ilość razy i w sposób nieograniczony zwiększać liczbę znaczników w miejscu P_2 . W znakowaniu $/1,w,1,0/$ przygotowane są dwa przejścia t_1 i t_3 . Możemy otrzymać dwa nowe węzły $m_3 = /1,w,0,0/$ i $m_4 = /1,w,0,1/$. Dla pierwszego z nich $T(/1,w,0,0/)$ jest zbiorem pustym i funkcja $d/m,t/$ jest nieokreślona. Z węzła $/1,w,0,1/$ po zapaleniu t_2 przechodzimy do węzła $m_5 = /1,w,1,0/$, który jest identyczny z węzłem m_2 .

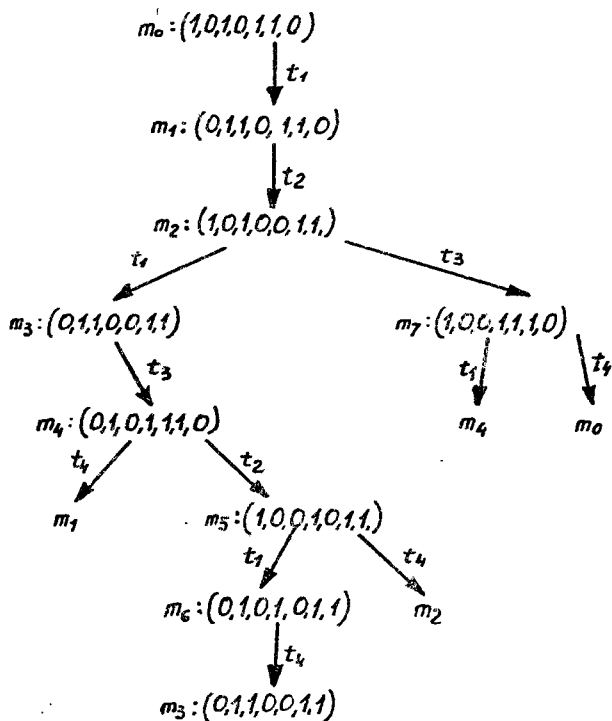
Rys.7 pokazuje kompletne drzewo osiągalności, natomiast rys.8 pokazuje drzewo osiągalności dla przykładu producent - konsument.

Zbiorem węzłów jest zbiór $Z/M/$ osiągalnych znakowań. Ponieważ w rozważanym drzewie nie ma węzłów zawierających symbol w , sieć jest ograniczona. Dla każdego znakowania zachodzi $\bigwedge_{p \in P} m/p/ \leq 1$, a więc sieć jest 1-ograniczona, czyli bezpieczna. Ponieważ suma znaczników w dowolnym znakowaniu jest równa 4, sieć jest zachowawcza. Z analizy drzewa wynika, że dowolne przejście może zapalić się po wcześniejszym zapaleniu się odpowied-

niego ciągu przejść. Wobec tego sieć jest żywa.



Rys.7. Drzewo osiągalności dla sieci z rys.6



Rys.8. Drzewo osiągalności dla przykładu producent - konsument

L i t e r a t u r a

- [1] Agerwala T.: Putting Petri nets to work - IEEE grudzień 1979
- [2] Barzilai Z., Strasburger E., Yoeli M.: Concurrent system modelling by conditional Petri nets - Technical Report No 124, March 1978

- [3] Burkhard H.: Ordered firings in Petri nets - Journal of Information Processing and Cybernetics No 2-3, Akademie-Verlag Berlin 1981
- [4] Dadda L.: The synthesis of Petri nets for controlling purposes and the reduction of their complexity - 2nd Symp. on Microarchitecture Euromicro 1976, Amsterdam 1976, 251-259
- [5] Dawid R.: Modular design of asynchronous circuits by graphs - IEEE Trans. on Comp. C-26 1977
- [6] Grabowski J.: On the analysis of switching circuits by means of Petri nets - EIK 14, 1978
- [7] Iszkowski W., Maniecki M.: Programowanie współbieżne - WNT 1982, rozdział 11
- [8] Majewski M., Albicki A.: Algebraiczna teoria automatów - WNT 1980, rozdział 9
- [9] Merlin P.M., Farber D.J.: Recoverability of communication protocols: implications of a theoretical study - IEEE Trans. Comm. COM-24, 1976
- [10] Patil S.S.: Micro-control for parallel asynchronous computers. Euro-micro, North-Holland 1975
- [11] Peterson J.: Petri nets - Department of Computer Science - The University of Texas, 1977
- [12] Starke P.H.: Petri Netze - VEB Verlag der Wissenschaften, Berlin 1980
- [13] Valette R.: Analysis of Petri nets by stepwise refinements. J. Comp. Syst. Sci. 18, 1979
- [14] Winkowski J.: Formal theories of Petri nets and net simulation. Prace CO PAN nr 242

INTRODUCTION TO PETRI NETS

Summary

Basic notions describing Petri nets are given in the paper. It is shown that Petri nets are a comfortable way of describing simultaneous and asynchronous systems. The paper presents two methods of the operation analysis of the model under consideration.

ВВЕДЕНИЕ К СЕТИ ПЕТРИ

Резюме

В статье представлены основные сведения описывающие сети Петри. Показано, что они совершенно полезны для описания систем действующих одновременно и асинхронно. Приведены также два метода осуществления анализа действия исследуемой модели.

Ewa Szlachta
Miroslaw Konieczny

ZASTOSOWANIE SIECI PETRI DO MODELOWANIA
UKŁADU REALIZUJĄCEGO REGUŁĘ DOSTĘPU
W PROSTYM SYSTEMIE TELEKOMUNIKACYJNYM

W pracy przedstawiono możliwości zastosowania sieci Petri do modelowania systemów zdarzeń dyskretnych, asynchronicznych i równoczesnych. Przykład zastosowania sieci Petri dotyczy problemu zwiększenia efektywności wykorzystania K podkanałów łączących węzły A i B w prostym systemie telekomunikacyjnym.

1. WSTĘP

Z wykorzystaniem wspólnego kanału do przesyłania informacji pochodzących z wielu źródeł wiąże się dwa zagadnienia:

- rozbitcie kanału na podkanały o przepustowościach rzędu szybkości podawania informacji przez pojedyncze źródło,
- określenie zasady przydzielania danemu źródłu podkanału.

Pierwsze zagadnienie dotyczy różnych technik zwielokrotniania /np. czasowego lub częstotliwościowego/, drugie natomiast to wybór reguł przydzielania podkanałów. Reguły te nazywamy regułami dostępu.

Niniejsza praca jest przykładem zastosowania sieci Petri do modelowania układu realizującego regułę dostępu.

2. SIECI PETRI - PODSTAWOWE POJĘCIA

Definicja 1. Siecią Petri jest uporządkowana trójka $S = /P, T, R/$, gdzie:

- P - skończony, niepusty zbiór miejsc,
- T - skończony, niepusty zbiór przejść,
- R - relacja na $P \cup T$ taka, że:

$$R \subseteq P \times T \cup T \times P$$

Definicja 2. Znakowaną siecią Petri nazywamy uporządkowaną parę

$$= /S, m_0/$$
, gdzie:

S jest siecią Petri,

m_0 jest pewnym znakowaniem początkowym, tj:

$M =$

$m_0: P \rightarrow N$, gdzie N jest zbiorem liczb naturalnych.

W reprezentacji graficznej miejsca przedstawiane są jako kółka, przejścia jako kreski, krawędzie zaś jako strzałki. Znakowanie początkowe m_0 polega na umieszczeniu w każdym kółku reprezentującym miejsce p , m_0/p znaczników. Mówimy, że miejsce p jest miejscem wejściowym /lub wyjściowym/ przejścia t , jeśli $p, t \in R$ /lub $t, p \in R'$ /. Zbiór wszystkich miejsc wejściowych /wyjściowych/ danego przejścia t oznaczamy jako $\text{Imp}/t/$ / $\text{Out}/t/$ /.

Niech dowolna funkcja $m: P \rightarrow N$ będzie znakowaniem sieci M . Przejście t jest przygotowane przez znakowanie m , jeśli każde miejsce wejściowe tego przejścia zawiera co najmniej jeden znacznik. Każde przygotowane przejście w danym znakowaniu m może zapalić się. Wyznacza to nowe znakowanie.

Definicja 3. Znakowanie m_2 jest osiągalne bezpośrednio ze znakowania m_1 , co oznaczamy $m_1 \rightarrow m_2$, jeśli istnieje przejście t przygotowane w znakowaniu m_1 takie, że:

$$\bigwedge_{p \in P} m_2/p/ = \begin{cases} m_1/p/ - 1, & \text{jeśli } p \in \text{Imp}/t/ \text{ i } p \notin \text{Out}/t/ \\ m_1/p/ + 1, & \text{jeśli } p \in \text{Out}/t/ \text{ i } p \notin \text{Imp}/t/ \\ m_1/p/ & \text{w pozostałych przypadkach} \end{cases}$$

Oznaczmy przez $\alpha /p, t/$ liczbę strzałek od miejsca p do przejścia t , oraz przez $\beta /t, p/$ liczbę strzałek od przejścia t do miejsca p . Mówimy, że przejście t jest przygotowane przez znakowanie m , jeśli:

$$\bigwedge_{p \in \text{Imp}/t/} m/p/ - \alpha /p, t/ \geq 0$$

Zapalenie przejścia t polega na usunięciu $\alpha /p, t/$ znaczników z każdego miejsca p będącego wejściowym dla t i dodaniu $\beta /t, p/$ znaczników do każdego miejsca p będącego wyjściowym dla t . Wobec tego zapalenie się przejścia t zmienia znakowanie m na m' według następującej zasady:

$$\bigwedge_{p \in P} m'/p/ = m/p/ - \alpha /p, t/ + \beta /t, p/$$

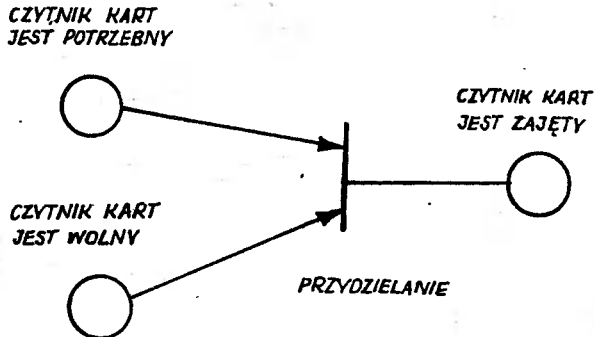
W szczególności, jeśli p nie jest miejscem wejściowym ani wyjściowym przejścia t wtedy $\alpha /p, t/ = \beta /t, p/ = 0$ i $m'/p/ = m/p/$.

Sieć Petri z powyższą modyfikacją nazywamy uogólnioną siecią Petri.

3. MODELOWANIE ZA POMOCĄ SIECI PETRI

W wielu przypadkach eksperymenty nad zjawiskami rzeczywistymi ze względu na koszt, a czasami niebezpieczeństwo ograniczamy do badania ich modelu. Reprezentowany jest on częstokroć za pomocą terminów matematycznych, które ujmują jego najważniejsze cechy. Sieci Petri modelują systemy, a w szczególności dwa ich aspekty: zdarzenie i warunek oraz relację pomiędzy nimi. W dowolnym czasie pracy systemu może pojawić się warunek, który

spowoduje zajście jednego ze zdarzeń. To zdarzenie może zmienić stan systemu, przerwać działanie warunku poprzedzającego i spowodować działanie następnego. Np. mamy dwa warunki: "potrzebny jest czytnik kart" oraz "czytnik kart jest dostępny". Warunki te mogą spowodować zdarzenie "przydzielanie czytnika kart". W wyniku zajścia tego zdarzenia przerwane zostają warunki "potrzebny jest czytnik kart" i "czytnik kart jest dostępny", a model przechodzi do stanu "czytnik kart jest niedostępny". Warunki i zdarzenie oraz ich zależności przedstawia rys.1, gdzie miejsca użyto do reprezentacji warunków, a przejścia do reprezentacji zdarzeń.



Rys.1. Prosty graf sieci Petri

Sieci Petri, podobnie jak modelowane systemy, są rozpatrywane jako ciąg dyskretnych zdarzeń, w którym jeden z możliwych porządków jest wyznaczony przez podstawową strukturę. Pozwala to na niezdeteterminowaną pracę sieci Petri. Jeśli w dowolnej chwili kilka przejść jest przygotowanych do zapalenia, wybór przejścia dokonywany jest w sposób przypadkowy. Niedeterminizm, chociaż korzystny z punktu widzenia modelowania, wprowadza znaczne utrudnienia podczas analizy sieci. Zapalenie się przejścia rozważane jest jako zdarzenie o czasie zerowym, chociaż w celu rozwiązania specjalnych problemów wprowadza się sieci, w których każdemu przejściu tęt przyporządkowuje się czas palenia różny od zera /tzw. sieci rzeczywiste Petri/.

4. PRZYKŁAD ZASTOSOWANIA SIECI PETRI DO MODELOWANIA UKŁADU REALIZUJĄCEGO REGUŁĘ DOSTĘPU

Jednym z zagadnień dotyczących zwiększania efektywności wykorzystania wspólnego kanału do przesyłania informacji pochodzących z wielu źródeł, jest wybór reguły przydzielania podkanałów. Najprostszą regułą jest reguła dostępu sztywnego. Polega ona na przydzielaniu danemu źródłu ustalonego podkanału na stałe. Przy takiej regule dostępu kanał wspólny wykorzystywany jest jako wiązka całkowicie niezależnych podkanałów. Większe wykorzystanie kanału uzyskujemy stosując regułę dostępu zmiennego polegającą na przydzielaniu danemu źródłu różnych podkanałów. Jedną z takich reguł jest

reguła ze zwalnianiem podkanałów, funkcjonująca w następujący sposób: jeśli źródło aktywizuje się, przydziela mu wolny podkanał, jeśli zaś źródło kończy swój cykl aktywności, zwraca podkanał do puli wolnych kanałów.

W przykładzie rozważać będziemy sytuację, w której dwa węzły A i B połączone są kanałem składającym się z K podkanałów. W przypadku, gdy liczba źródeł przyjętych do systemu nie przekracza ilości podkanałów, stosujemy regułę dostępu sztywnego. W momencie pojawienia się źródła nadmiarowego, w systemie stosujemy regułę ze zwalnianiem podkanałów. Zakładamy, że ciąg informacji podawanych przez źródło składa się z ciągu zaczynających się w sposób nierytmiczny bloków przedzielonych przerwami. Mówimy, że źródło w momencie podawania bloku informacji jest w stanie aktywnym. Poza tym różnicą będziemy następujące stany źródła: bierny i gotowy.

Rozważmy następującą sytuację: w systemie znajduje się K źródeł /stosujemy regułę dostępu sztywnego/ i w pewnym momencie przyjęte jest K+1 źródło /nadmiarowe/. System przechodzi w stan, w którym stosujemy regułę ze zwalnianiem podkanałów. W momencie tym zastać możemy każde z wcześniej przyjętych K źródeł w jednej z dwóch sytuacji: transmisji bloku lub przerwy. Zakładamy, że jeśli moment zmiany reguły dostępu zostanie źródło w stanie przerwy, następny blok rozważanego źródła przesłany będzie jeszcze za pomocą tego samego podkanału, w przeciwnym przypadku po zakończeniu podawania bloku źródło oddaje podkanał do puli wolnych kanałów.

W czasie pracy systemu z regułą ze zwalnianiem podkanałów, każde źródło traci informację przez czas jaki upływa od chwili rozpoczęcia przez nie podawania bloku do momentu, w którym jakiegokolwiek inne źródło zakończy swój cykl aktywności.

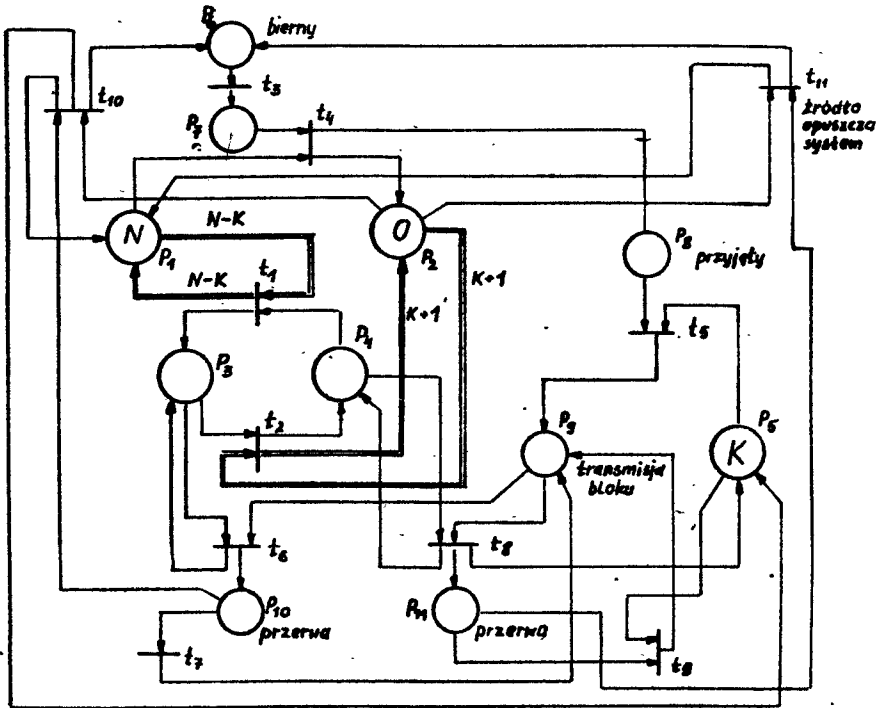
W przedstawionej poniżej sieci Petri modelującej rozważane zagadnienie wspólne dla wszystkich źródeł są miejsca p_1, p_2, p_5 . W miejscu p_1 w znakowaniu początkowym umieszczamy N znaczników /dopuszczalna ilość źródeł w systemie/, natomiast w miejscu p_5 , które jest pulą wolnych podkanałów, K znaczników.

4.1. Opis działania sieci

Opis działania sieci /oraz graf/ ograniczymy do warunków i zdarzeń jakie mogą mieć miejsce z punktu widzenia pojedynczego źródła. Dla pozostałych źródeł ciągi możliwych zdarzeń są identyczne.

Po zapaleniu się przejścia t_3 źródło ze stanu "bierny" przechodzi do stanu "gotowy". Zostaje przyjęte do systemu pod warunkiem, że liczba aktualnie przebywających w systemie źródeł jest mniejsza od N. Załóżmy, że warunek ten jest spełniony i pali się przejście t_4 . Źródło zostaje przyjęte do obsługi zwiększając jednocześnie o 1 w miejscu p_2 liczbę przyjętych użytkowników. Miejsca p_3 i p_4 sygnalizują stan sieci i są warunkiem wzajemnie wykluczających się zdarzeń. W znakowaniu początkowym $m/p_j/ = 1$ i sieć jest w stanie pracy z regułą dostępu sztywnego. Jeśli ilość źródeł przyjętych do systemu przekroczy wartość K, pali się przejście t_2 i sieć przechodzi

dzi do stanu pracy z regułą ze zwalnianiem podkanałów. Podobnie, jeśli w międzyczasie ilość źródeł w systemie zmniejszy się do wartości K , pali się przejście t_1 i sieć wraca do stanu pracy z regułą dostępu sztywnego.



Rys. 2. Sieć Petri modelująca układ realizujący regułę dostępu

Po zapaleniu się przejścia t_5 /pod warunkiem, że wolny jest chociaż jeden podkanał/, następuje transmisja bloku. Źródło w dowolnej chwili może zakończyć swój cykl aktywności. Jeśli sieć jest w stanie pracy z regułą dostępu sztywnego pali się przejście t_6 . W stanie przerwy źródło może się w każdej chwili zaktywizować /pali się przejście t_7 /, bądź opuścić system /pali się przejście t_{10} /, oddając jednocześnie podkanał do puli wolnych kanałów /miejsce p_7 / i zmniejszając liczbę źródeł w systemie. W przypadku, gdy źródło po swoim cyklu aktywności zastanie sieć w stanie pracy z regułą ze zwalnianiem podkanałów, pali się przejście t_8 . Ponowne przejście źródła do stanu transmisji bloku wiąże się z pobraniem podkanału z puli wolnych kanałów /pali się przejście t_9 /.

4.2. Analiza działania sieci

W celu analizy działania sieci Petri modelującej układ realizujący regułę dostępu /rys.2/ odszukamy jej zbiory niezmiennicze. Niech C będzie macierzą incydencji. Układ równań $C \cdot y = 0$ dla rozważanego przykładu przedstawia się:

$$\begin{array}{c}
 t_1 \\
 t_2 \\
 t_3 \\
 t_4 \\
 t_5 \\
 t_6 \\
 t_7 \\
 t_8 \\
 t_9 \\
 t_{10} \\
 t_{11}
 \end{array}
 \begin{bmatrix}
 P_1 & P_2 & P_3 & P_4 & P_5 & P_6 & P_7 & P_8 & P_9 & P_{10} & P_{11} \\
 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\
 -1 & 1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\
 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 1 \\
 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & -1 \\
 1 & -1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & -1 & 0 \\
 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1
 \end{bmatrix}
 \begin{bmatrix}
 y_1 \\
 y_2 \\
 y_3 \\
 y_4 \\
 y_5 \\
 y_6 \\
 y_7 \\
 y_8 \\
 y_9 \\
 y_{10} \\
 y_{11}
 \end{bmatrix}
 =
 \begin{bmatrix}
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0
 \end{bmatrix}$$

Rozwiązaniem powyższego układu równań są wektory:

$$y_1 = /11000000000/$$

$$y_2 = /00110000000/$$

$$y_3 = /00000111111/$$

$$y_4 = /00001000110/$$

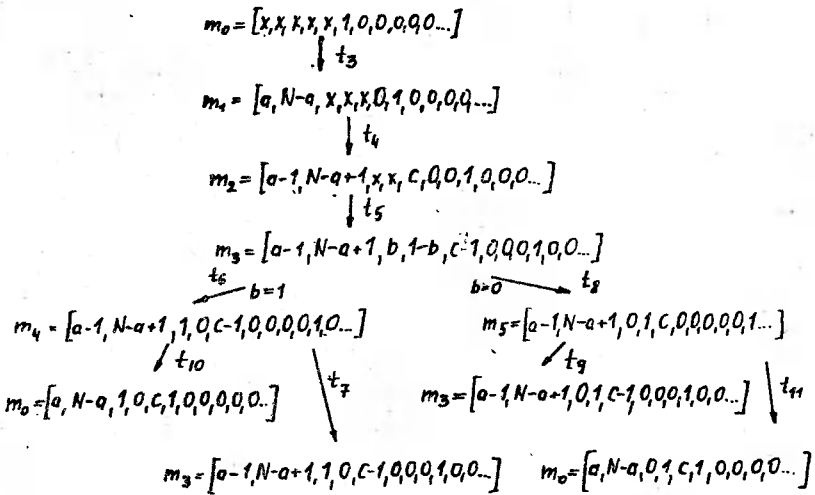
Wobec tego zbiorami niezmienniczymi są: $\{P_1, P_2\}$, $\{P_3, P_4\}$, $\{P_5, P_7, P_8, P_9, P_{10}, P_{11}\}$ oraz $\{P_5, P_9, P_{10}\}$.

Ponieważ każde miejsce należy do pewnego podzbioru niezmienniczego, a sieć działanie swoje rozpoczyna od znakowania z ograniczoną liczbą znaczników - sieć jest ograniczona. Dokładniej, miejsca p_1 i p_2 są N -ograniczone, miejsce p_5 K -ograniczone, pozostałe zaś 1-ograniczone. Cała sieć jest N -ograniczona. Sieć nie jest zachowawcza, ponieważ w sieci istnieją przejścia, które nie są zachowawcze w tym sensie, że zachodzi równość mocy zbiorów card Inp/t = card Out/t. Takimi przejściami w sieci są t_5, t_8, t_9 i t_{10} . Z niezmienniczości podzbiorów $\{P_1, P_2\}$ oraz $\{P_3, P_4\}$ wynika, że system może znaleźć się w jednym z dwóch stanów:

a/ pracy z regułą dostępu sztywnego,

b/ pracy z regułą zwalniania podkanałów.

Rozważmy drzewo osiągalności dla sieci z rys.2.



Rys.3. Drzewo osiągalności dla sieci z rys.2

Miejsca P_1, P_2, P_3, P_4 i P_5 są wspólne i znakowanie ich zależy również od zdarzeń jakie mogą mieć miejsce po przyjęciu do systemu innych źródeł. W celu zwiększenia czytelności drzewa miejsca te znakujemy przy pomocy symboli a, b oraz c , spełniających następujące warunki: $1 \ll a \ll N$, $b=1$ lub $b=0$, $0 < c \ll K$. Z rys.3 wynika, że sieć jest żywa, tzn. dla dowolnego znakowania m i dowolnego przejścia t istnieje znakowanie m' osiągalne z m , w którym przejście t jest przygotowane. W sieci nie istnieje znakowanie, w którym żadne z przejść nie jest przygotowane.

L i t e r a t u r a

- [1] Agerwala T.: Putting Petri nets to work - IEEE grudzień 1979
- [2] Baker H.G.: Petri nets and languages - MIT, Proj. MAC, Computation Structures Group Memo, 1972
- [3] Dadda L.: On the simulation of Petri nets a control tool - Euromicro Newsletter 2, 1976
- [4] Genrich H.J.: The Petri net representation of mathematical knowledge. GMD IB ISF-76-05, 1976
- [5] Grabowski J.: On the analysis of switching circuits by means of Petri nets - EIK 14, 1978
- [6] Godlewska J., Kalinowski J.: Zastosowanie sieci Petriego do modelowania złożonych systemów telekomunikacyjnych - Przegląd Telekomunikacyjny 7/80 str.226

- [7] Hack M.: Analysis of productions schemata by Petri nets - MIT, Proj. MAC, TR 94, 1972
- [8] Janicki R.: Synthesis of concurrent schemes - LNCS 64, 1978
- [9] Kasami T.: Vector addition systems and synchronization problems of concurrent processes - U. of Hawaii, Honolulu TR ARO-23, 1974
- [10] Keller R.M.: Vector replacement systems: a formalism for modelling asynchronous systems - Princeton U., CS Lab., TR 117, 1972
- [11] Kosaraju S.R.: Limitations of Dijkstra's semaphore primitives and Petri nets - Operating Systems Review 7, 1973
- [12] Kotov V.E.: An algebra for parallelism based on Petri nets - LNCS 64, 1978
- [13] Mazurkiewicz A.: In variants of concurrent programs - P. IFIP-INFOPOL Conf. on Information Processing, North Holland 1977
- [14] Miller R.E.: A comparison of some theoretical models of parallel computation - IEEE Trans Comp. C-22, 1973
- [15] Montanari U.: Networks of constraints: fundamental properties and applications to picture processing - Information Sci. 7, 1974
- [16] Ramchandani C.: Analysis of asynchronous concurrent systems by timed Petri nets - MIT, Proj. MAC, TR 120, 1974
- [17] Seidler J.: Analiza i synteza sieci łączności dla systemów teleinformatycznych - PWN Warszawa, 1979
- [18] Winkowski J.: Formal theories of Petri nets and net simulation - Polish Academy of Sci., Comp. Center, PAS Rep. 242, 1976
- [19] Weinstein C.: Fractional speech loss and talker activity model for TASI and for packed-switched speech - IEEE Trans. Commun. vol.26, 1978
- [20] Zuberek W.: Analiza efektywności jednostek centralnych maszyn cyfrowych oraz niektóre metody jej zwiększania - Praca doktorska, Warszawa 1976

THE APPLICATION OF THE PETRI NETS TO MODELLING
THE ACCESS RULE IN COMMUNICATION SYSTEMS

Summary

The paper gives an application possibility of Petri nets to modelling discrete, asynchronous and simultaneous system events. The example of the application of Petri nets is connected with an increase in the efficiency of utilization of K sublinks joining points A and B in communication systems.

ПРИМЕНЕНИЕ СЕТИ ПЕТРИ ДЛЯ МОДЕЛИРОВАНИЯ ПРАВИЛА ДОСТУПА В ПРОСТОЙ ТЕЛЕКОММУНИКАЦИОННОЙ СИСТЕМЕ

Резюме

В статье представлены возможности использования сети Петри для построения модели дискретных систем действующих асинхронно и одновременно. Пример применения сетей Петри касается проблемы увеличения эффективности использования. К подкавалов соединяющих узлы А и В в простой телекоммуникационной системе.

Antoni Zabłudowski
Andrzej Sobólski

OCENA EFEKTYWNOŚCI REKURENCYJNEGO ALGORYTMU WYZNACZANIA MIAR NIEZAWODNOŚCIOWYCH GRAFU PROBABILISTYCZNEGO

W pracy dokonano oceny, pod względem liczby obliczeń, dwóch algorytmów służących do znajdowania miar niezawodnościowych grafów probabilistycznych. Na podstawie obliczeń dla losowo wygenerowanych grafów określone zostały obszary efektywności obydwu rozpatrywanych algorytmów.

1. WSTĘP

Obliczanie parametrów probabilistycznych grafu odgrywa istotną rolę przy analizie niezawodnościowej wielkich systemów. Wyznaczenie parametrów probabilistycznych grafu służyć może zarówno do znajdowania funkcji niezawodności /dystrybuanty/ systemów mieszanych, w których każdy element opisany jest pewną funkcją niezawodności jak również do obliczania granicznych wartości procesu odnowy [3]. Wyznaczanie parametrów probabilistycznych grafów jest także podstawowym krokiem w algorytmach syntezy sieci telekomunikacyjnych, w których uwzględnia się wskaźniki niezawodnościowe określone dla tych sieci. Ze względu na to, że szybkość działania algorytmów służących do wyznaczania parametrów probabilistycznych często decyduje o możliwości zastosowania algorytmów w praktyce, stawia się im żądanie, aby ich czasy obliczeń były jak najkrótsze.

W teorii sieci telekomunikacyjnych problem wyznaczania parametrów probabilistycznych /miar niezawodnościowych/ odgrywa istotną rolę. Dla sieci tych definiuje się dwie podstawowe miary niezawodnościowe, a mianowicie:

- a/ prawdopodobieństwo istnienia połączenia między wybraną parą węzłów s oraz t , które przyjęto oznaczać przez P_{gt}
- b/ prawdopodobieństwo spójności sieci, które oznaczane jest przez P_c .

Obecnie znanych jest wiele metod służących do wyznaczania parametrów probabilistycznych dla określonego grafu. Dokładny przegląd znanych metod analizy grafu probabilistycznego /sieci/ podany został w pracy [5]. Biorąc pod uwagę kryterium jakim jest komputerowy czas obliczeń należy zwrócić uwagę na metody dekompozycji sieci oraz metody rekurencyjne jako te, dla których czasy te są najkrótsze. Jedną z takich metod, a mianowicie metoda

rekurencyjna, zaproponowana została w pracy [6], odmienna zaś - w pracy [1]. Celem niniejszej pracy jest oszacowanie efektywności działania rekurencyjnego algorytmu służącego do wyznaczania parametrów P_{st} i P_c , zaproponowanego w pracy [6]. Ze względu na trudności analitycznego wyznaczenia złożoności obliczeniowej tego algorytmu, za miernik efektywności przyjęto czas pracy komputera. Uzyskane wyniki porównano z oszacowaniem szybkości działania algorytmu przedstawionego w pracy Buzacotta [1].

2. PODSTAWOWE POJĘCIA ORAZ OPIS DZIAŁANIA ALGORYTMU

Zakładamy, że:

A1. Dany jest graf $G = \langle N, B \rangle$

gdzie:

$N = \{n_1, n_2, \dots, n_w\}$ jest zbiorem węzłów grafu

$B = \{b_1, b_2, \dots, b_k\}$ jest zbiorem gałęzi grafu

A2. Każdej gałęzi grafu przyporządkowana jest binarna zmienna losowa x_k , która określa stan gałęzi, przy czym:

$$P [x_k = 1] = p_k$$

$$P [x_k = 0] = q_k = 1 - p_k$$

dla $k = 1, 2, \dots, K$

A3. Zmienne losowe x_k są statystycznie niezależne.

Graf określony zgodnie z założeniami A1, A2, A3 będziemy nazywać dalej grafem probabilistycznym. Dla takiego grafu poszukiwać będziemy prawdopodobieństwa P_{st} istnienia połączenia między wybraną parą węzłów s oraz t , zgodnie z algorytmem podanym w [6].

Na początek podana zostanie szkieletowo zasada działania algorytmu. Jest ona następująca:

$$E1. V_0 = \{s\}, \quad P_{st} = 0, \quad k = 1$$

E2. Ze zbioru gałęzi grafu wybieramy dowolną gałąź, np. b_1 . Niech gałąź ta łączy węzły s, n_{w_1} . Węzły te umieszczamy w zbiorze węzłów sprawdzanych V_1 . Tworzymy podziały $\mathcal{T}_1^1(V_1)$ na zbiorze V_1 i obliczamy prawdopodobieństwa wystąpienia podziałów $\mathcal{T}_1^1(V_1)$.

E3. W k -tym kroku wybieramy gałąź b_k . Znajdujemy węzły n_i, n_j , umieszczamy je w zbiorze V_k i tworzymy podziały $\mathcal{T}_1^k(V_k)$. Prawdopodobieństwa $P[\mathcal{T}_1^k(V_k)]$ znajdujemy jako sumę iloczynów p_k lub q_k i prawdopodobieństw $P[\mathcal{T}_1^{k-1}(V_{k-1})]$.

E4. Sprawdzamy, czy istnieją podziały $\mathcal{T}_1^k(V_k)$ posiadające blok, w którym znajdują się jednocześnie węzeł s i t . Podziały takie \mathcal{T}_1^k usuwamy ze

zbioru podziałów oraz podstawiamy:

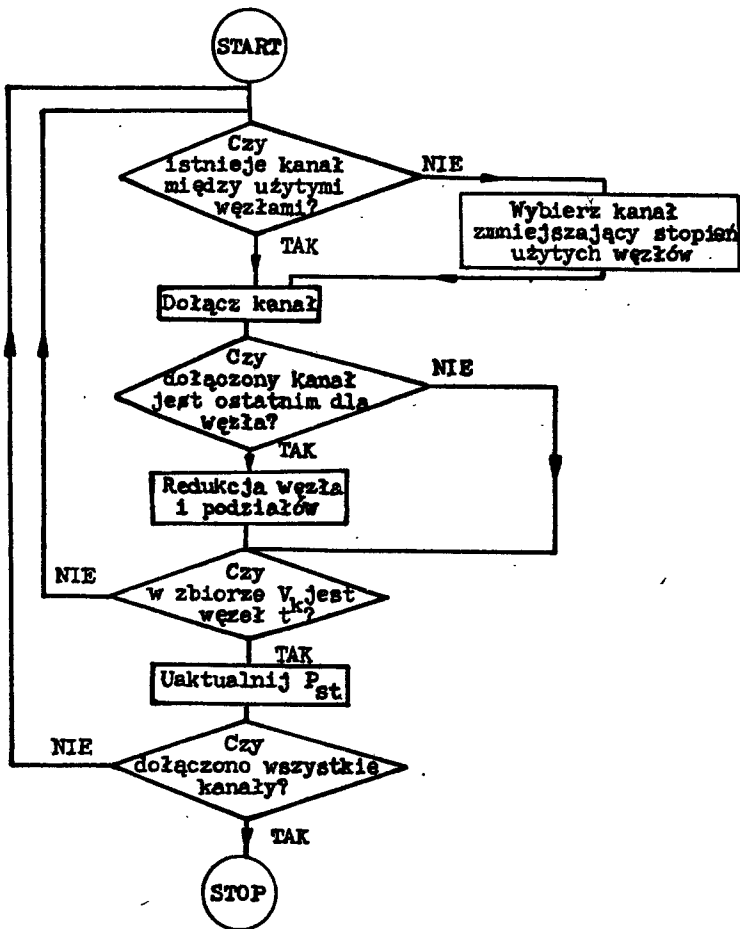
$$P_{st} = P_{st} + \sum_1 [P \pi_1^k]$$

E5. Sprawdzamy, czy istnieją gałęzie b_k

Tak: Wracamy do E3.

Nie: Koniec.

Schemat działania powyższego algorytmu przedstawiony został na rys.1.



Rys.1. Schemat blokowy zrealizowanego algorytmu rekurencyjnego

W trakcie pisania programu starano się maksymalnie ograniczyć wielkość tablic przechowywanych w pamięci komputera. Dlatego też w specjalny sposób dokonywano w każdym cyklu algorytmu wyboru gałęzi b_k , tak aby moc zbioru V_k była jak najmniejsza. I tak w pierwszej kolejności wybierane były te gałęzie, które łączą węzły znajdujące się w zbiorze węzłów sprawdzanych V_k /rozmiar tablic nie ulega zwiększeniu/. Ponadto możliwie szybko starano się eliminować te węzły w zbiorze V_k , dla których stopień wężła jest najmniejszy /w algorytmie stopień wężła określają tylko gałęzie nie dołączone/. Redukując takie węzły pozbywamy się również tych podziałów, w których węzeł źródłowy s jest izolowany od reszty grafu.

3. OCENA EFEKTYWNOŚCI ALGORYTMU

Podstawowymi kryteriami efektywności algorytmów są [4]:

- czas obliczeń algorytmu mierzony jako funkcja wielkości danych wejściowych
- obszar pamięci komputera wymagany przez algorytm, również zależny od wielkości danych wejściowych.

W przypadku rozpatrywanym w tej pracy jako dane wejściowe traktować będziemy następujące wielkości: średni stopień wężła, liczbę węzłów w grafie oraz strukturę topologiczną grafu. Zarówno średni stopień wężła oraz liczba węzłów są zdefiniowane /zadane każdorazowo do obliczeń/, natomiast topologia grafu przyjmuje charakter losowy /jest generowana losowo przez odpowiedni program/, dlatego też dla określenia czasu obliczeń rozpatrywać dalej będziemy czas uzyskany jako średni dla kilku wygenerowanych przypadków.

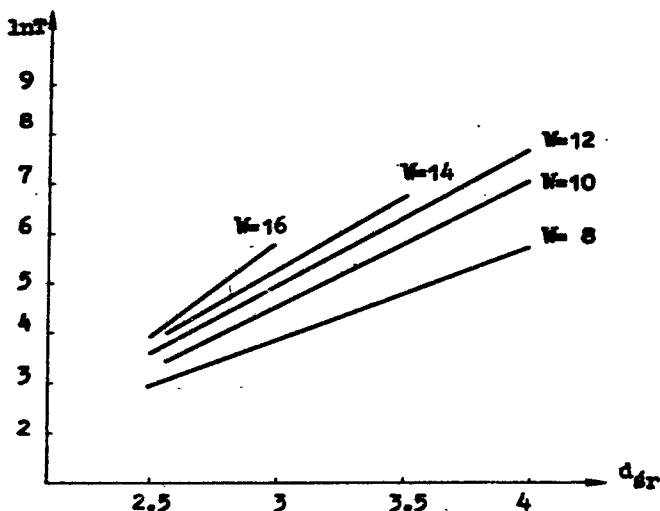
Stosując rekurencyjny algorytm służący do wyznaczania parametru P_{st} udało się ograniczyć do minimum zajmowany obszar pamięci komputera. Dlatego też w niektórych przypadkach uzyskano na maszynie „Odra 1204” rozwiązanie dla grafu o 20 węzłach /średni stopień wężła $d_{gr} = 2.5$ / w krótkim czasie /obliczenia trwały 69 sekund/. Dla wszystkich rozpatrywanych grafów o 12 węzłach rozwiązania znalezione zostały w czasie krótszym od 2700 sekund. Parametr P_{st} obliczono dla grafów o liczbie węzłów $W = 8, 12, 16$ i średnim stopniu wężła $d_{gr} = 2.5, 3, 3.5$ i 4 . Dla grafów zawierających 10 węzłów średni stopień wężła był równy $2.6, 3, 3.6$ i 4 , zaś dla grafów zawierających 14 węzłów średni stopień wężła był następujący: $2\frac{4}{7}, 3, 3\frac{4}{7}$ i 4 . We wszystkich przykładach mierzony był czas obliczeń. Pomierzone czasy zestawiono w tabeli podając jednocześnie średni czas obliczeń T dla określonej liczby węzłów i średniego stopnia wężła /tab.1/.

Logarytmiczną zależność średniego czasu obliczeń od średniego stopnia wężła dla grafów o zadanej liczbie węzłów przedstawiono na rys.2.

Tabela 1

d. W	2.5		2.6		2.7		3		3.5		3.6		3.7		4	
	Ar	T	Ar	T	Ar	T	Ar	T	Ar	T	Ar	T	Ar	T	Ar	T
8	19						61		28						436	
	20	19					56	55.75	125	86.75					309	283.25
	19	18					32		120						118	
10							74		74						270	
12							39	70							476	784.5
							83								912	
							109								816	
14							39								934	
16							108		484						2305	2161.5
							173	130	444	410					2615	
							94		106						1508	
14							145		606						2218	
16							179	214.25						989		
							31							x		
							56							x		
16							204							x		
														x		
														x		
16							222	488.5								
							391									
							1148									
16							193									

x - przekroczenie pamięci komputera



Rys.2. Logarytmiczna zależność średniego czasu obliczeń T od średniego stopnia wężła - d_{gr} / W - liczba wężłów w grafie/

Z wykresów przedstawionych na rys.2 wynika, że $\ln T$ jest liniową funkcją d_{gr} . Jako funkcję aproksymującą przyjęto

$$\ln T = a \cdot W \cdot d_{gr} + b \cdot W^c$$

Wyznaczając stałe a , b i c metodą najmniejszych sumy kwadratów otrzymano funkcję:

$$\ln T = 0.225 \cdot W \cdot d_{gr} - 0.043 \cdot W^{1.725} \quad (1)$$

Jak widać z otrzymanej zależności czas obliczeń w większym stopniu zależy od średniego stopnia wężła aniżeli od liczby wężłów w grafie. Wynika to po prostu z faktu, iż zwiększając liczbę gałęzi wychodzących z poszczególnych wężłów grafu zwiększa się wymiar tablic określających generowane na zbiorze V_k odpowiednie podziały. Jest to z kolei konsekwencją zwiększania się mocy zbioru sprawdzanych wężłów V_k w każdym kolejnym cyklu algorytmu.

4. WNIOSKI

Rekurencyjny algorytm, który poddany został ocenie, może być stosowany przy analizie grafów, w których stopień wężła jest znacznie mniejszy od liczby wężłów.

Porównując efektywność badanego algorytmu z najlepszym pod względem efektywności dla grafu zupełnego algorytmem zaproponowanym przez Buza-

cotta w pracy [1] znaleźć można zakres opłacalności wykorzystania rozważanego algorytmu. Czas pracy algorytmu Buzacotta dla grafu pełnego jest określony przez następującą liczbę mnożeń

$$M(W) = (W-1) \cdot (W+1) \cdot 3^{W-3} - 2^{W-1} + 1 \quad (2)$$

oraz następującą liczbę dodawań

$$A(W) = 3^{W-1} - 2^W + 1 \quad (3)$$

dla obliczania P_c , oraz

$$M(W) = 3^{W-2} + (W-1) \cdot (2^{W-2} - 2) - 1 \quad (4)$$

liczbę mnożeń dla obliczania P_{st} .

Zakładając, że znaleziona funkcja czasu obliczeń (1) określa również proporcjonalnie średnią liczbę operacji wykonywanych przez komputer mamy:

$$T = \bar{T} \cdot M \quad (5)$$

gdzie:

\bar{T} - średni czas działania operacji

M - liczba działań

Logarytmując stronami zależność (5) otrzymamy:

$$\ln T = \ln M + \ln \bar{T} \quad (6)$$

przy czym

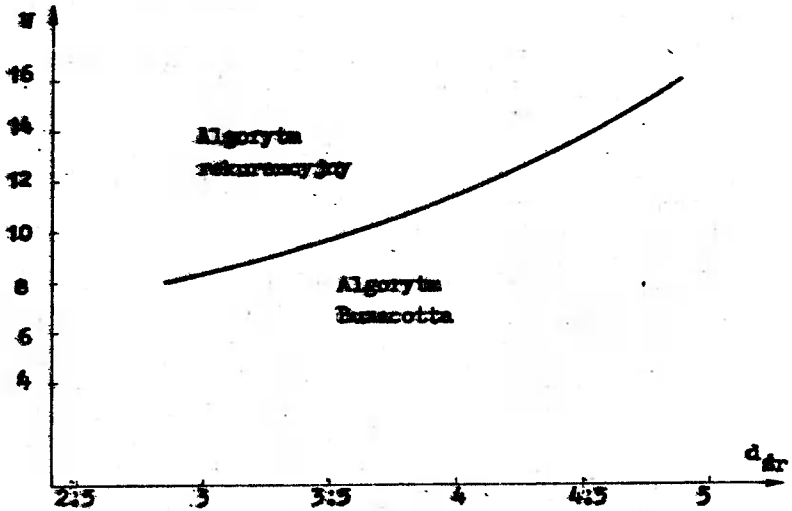
$$\ln \bar{T} = c = \text{const.}$$

Przyjmując, że niezależnie od topologii grafu algorytm Buzacotta wykonać powinien co najmniej

$$M = 3^{W-1} - 2^W + 1 \quad (7)$$

mnożeń [1], oraz ustalając wartość $c = -4$ /na podstawie [2]/ ocenić można zakres, w którym algorytm rekurencyjny posiada krótszy czas obliczeń od algorytmu Buzacotta. Wyniki takiej oceny przedstawiono na rys.3.

Z otrzymanych wyników widać, że algorytm rekurencyjny wykazuje przewagę przy analizie grafów rzadkich. W praktyce najczęściej nie wykorzystuje się grafów zupełnych, stąd widać, że celowe jest również dla dowolnej topologii grafu stosowanie algorytmów nie będących najlepszymi dla grafu zupełnego.



Rys. 3. Obszary efektywności algorytmów

Literatura

- [1] Buzacott J.A.: A recursive algorithm for finding reliability measures related to the connection of nodes in a graph Networks vol.10 No 4, s.311-327
- [2] Jerzykiewicz K., Szpżepkiewicz J.: Algol 1204 - system programowania maszyny cyfrowej Odra 1204. PWN, Warszawa 1973
- [3] Kopociński B.: Zarys teorii odnowy i niezawodności. Warszawa 1973
- [4] Weide B.: A Survey of analysis techniques for discrete algorithms. Computing Surveys vol.9 No 4, December 1977, s.291-313
- [5] Zabłudowski A.: Algorytmy optymalizacji struktur topologicznych sieci. Rozprawy Nr 9, Bydgoszcz 1982
- [6] Zabłudowski A.: A recursive method for reliability measures evaluation. Prace IPI PAN Nr 483, Warszawa, czerwiec 1982

THE ESTIMATION OF THE EFFICIENCY OF THE NETWORK RELIABILITY
RECURSIVE ALGORITHM

Summary

The paper presents the estimation of the number of calculations for two reliability parameters evaluation algorithms. A comparison is made for several randomly generated probabilistic graphs. Using the obtained result there is defined the efficiency space for both the considered algorithms.

АНАЛИЗ ЭФФЕКТИВНОСТИ РЕКУРЕНТНОГО АЛГОРИТМА ДЛЯ НАХОЖДЕНИЯ НАДЕЖНОСТНЫХ ПОКАЗАТЕЛЕЙ СЕТИ

Резюме

В работе сделан анализ двух алгоритмов нахождения надёжных показателей вероятностных графов с точки зрения времени расчетов. На основе расчетов для нескольких случайных графов определены области в которых рассматриваемые алгоритмы эффективны.

Antoni Zabłudowski
Bożydar Dubalski

PSEUDOLOSOVY GENERATOR PROSTEGO STRUMIENIA POISSONA

W pracy przedstawiono koncepcję i schemat blokowy cyfrowego generatora losowego ciągu impulsów z wykładniczym rozkładem czasu pomiędzy kolejnymi impulsami. Generator taki nazywany jest generatorem prostego strumienia Poissona.

Otrzymane wyniki w pełni potwierdzają słuszność przyjętej koncepcji.

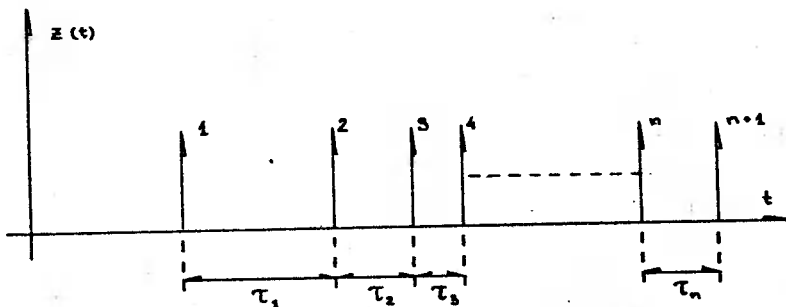
1. WSTĘP

Istnieje wiele fizycznych zjawisk, w których w losowych odstępach czasu pojawiają się krótkie impulsy o stałej bądź zmiennej amplitudzie. O zjawiskach takich mówimy, że tworzą one losowe źródło impulsów. Najczęściej zakłada się, że źródło to posiada charakter poissonowski. Przykładem poissonowskich źródeł impulsów mogą być zgłoszenia telefoniczne pojawiające się na centrali, zgłoszenia pojawiające się do obsługi w systemach kolejkowych, szumy impulsowe występujące w elementach elektronicznych lub systemach telekomunikacyjnych czy też uszkodzenia pojawiające się w złożonych systemach radioelektronicznych. Dla celów fizycznego symulowania systemów, w których występuje poissonowskie źródło impulsów należy zbudować urządzenie wiernie odtwarzające badane zjawisko. Urządzenie takie nosi nazwę generatora prostego strumienia Poissona. Z przedstawionych rozważań dotyczących charakteru źródła impulsów wynika, że generator strumienia Poissona musi być generatorem pewnego procesu stochastycznego, w którym na wyjściu generowane są krótkie impulsy pojawiające się w losowych chwilach czasowych. Dobrym modelem takiego generatora może być np. urządzenie, w którym jako źródło impulsów wykorzystuje się promieniotwórczy rozpad pierwiastków. Wadę generatorów wykorzystujących promieniotwórczy rozpad pierwiastków stanowi fakt, iż intensywność takiego strumienia jest stała. Stąd też zachodzi potrzeba budowy urządzeń, które mają co prawda charakter pseudolosowy, jednakże z wystarczająco wierną dokładnością aproksymują w pełni losowe modele źródeł oraz posiadają możliwość łatwej regulacji intensywności impulsów [2, 3].

Niniejsza praca poświęcona jest przedstawieniu takiego urządzenia.

2. WŁASNOŚCI POISSONOWSKIEGO STRUMIENIA IMPULSÓW ORAZ SPOSÓB JEGO GENERACJI

Poissonowski strumień impulsów może być traktowany jako uogólniona pochodna procesu Poissona [8]. Stąd wiele własności procesu Poissona przenosi się bezpośrednio na poissonowski strumień impulsów. Strumień taki może być przedstawiony tak, jak to pokazano na rys.1.



Rys.1. Przykładowa realizacja procesu w poissonowskim generatorze impulsów

Aby losowy strumień impulsów był prostym strumieniem Poissona potrzeba i wystarcza by:

- zmienna losowa τ_n określająca czas pomiędzy momentami pojawienia się n -tego oraz $n+1$ -tego impulsu posiadała rozkład wykładniczy jednakowy dla każdego $n=1, 2, \dots$,
- zmiennne τ_i, τ_j były statystycznie niezależne dla każdego $i \neq j$.

Jak wynika z warunku a/ dystrybuanta zmiennej τ_n dana jest w postaci:

$$P[\tau_n \leq t] = 1 - e^{-\lambda t} \quad (1)$$

gdzie:

λ - stała liczba.

Przyjmując ustalony przedział czasu o długości równej t_0 można określić rozkład liczby impulsów pojawiających się na wyjściu generatora strumienia Poissona w odcinkach czasu o długości t_0 . Rozkład ten opisany jest w postaci [8]:

$$P[Y = K] = e^{-\lambda t_0} \frac{(\lambda t_0)^K}{K!}, \quad K = 0, 1, 2, \dots \quad (2)$$

gdzie:

Y - zmienna losowa przedstawiająca liczbę impulsów w odcinku czasu równym t_0 .

Jak widać zmienna określająca liczbę impulsów w ustalonym czasie podlega rozkładowi Poissona. Prosty strumień Poissona może być w pełni okreś-

lany przez wyznaczenie rozkładu zmiennej Y .

Generowanie poissonowskiego strumienia impulsów odbywać się może w następujący sposób:

- a/ poprzez wykorzystanie generatora procesu zero-jedynkowego,
- b/ poprzez sumowanie wielu strumieni o dowolnym rozkładzie,
- c/ poprzez generowanie czasów o losowej długości, określonych zmienną o rozkładzie wykładniczym.

Pierwszy sposób generacji poissonowskiego strumienia impulsów wykorzystuje własności ciągu niezależnych zmiennych losowych / generowanych przez generator procesu zero-jedynkowego/ Z_1, Z_2, \dots, Z_n o rozkładach binarnych:

$$P[Z_i = 1] = p_i, \quad P[Z_i = 0] = q_i = 1 - p_i, \quad i = 1, 2, \dots, n \quad (3)$$

przy czym prawdopodobieństwo $p_i \ll 1$. Łatwo pokazać, że suma takich zmiennych losowych

$$Z = Z_1 + Z_2 + \dots + Z_n + \dots \quad (4)$$

jest zmienną losową o rozkładzie będącym przybliżeniem rozkładu Poissona z parametrem $a = p_1 + p_2 + \dots + p_n + \dots$. Dla $p_i = 0$ i $n \rightarrow \infty$ oraz $np_1 < \infty$ korzystając z twierdzenia Poissona otrzymujemy, że poszukiwany rozkład jest rozkładem Poissona. Zatem generator procesu binarnego staje się przybliżeniem generatora prostego strumienia Poissona. Metoda generacji strumienia tego typu wykorzystująca generator procesu binarnego jest jednak zbyt mało dokładna i dlatego może być stosowana tylko tam, gdzie nie jest wymagana duża dokładność. Generator prostego strumienia Poissona wykorzystujący podaną wyżej zasadę opisany został w pracy [1].

Sposób drugi pozwala na znacznie bliższe przybliżenie strumienia Poissona [4]. Jako źródła impulsów użyć można tutaj procesów zero-jedynkowych, dokładniej - generatorów krótkich impulsów pojawiających się w momentach przejścia procesu z $0 \rightarrow 1$. Prawdopodobieństwo p_i tego, że i -ty proces przyjmie wartość 1 powinno być bliskie zeru. Biorąc $N \gg 1$ takich procesów i sumując ich wyjścia otrzymamy strumień, który dobrze aproksymuje prosty strumień Poissona.

Trzecia z metod pozwalająca na uzyskanie prostego strumienia Poissona wymaga generacji czasów o rozkładzie wykładniczym. Jak wiadomo, na to aby dysponując generatorem rozkładu jednostajnego wygenerować zmienną losową daną dystrybuantą $F(x)$ o ciągłej funkcji odwrotnej F^{-1} trzeba postąpić w następujący sposób [9]:

- a/ wygenerować wartość R zmiennej losowej R o rozkładzie jednostajnym,
- b/ obliczyć dla każdej wartości R wielkość x szukanej zmiennej z zależności:

$$x = F^{-1}(R) \quad (5)$$

Dla prostego strumienia Poissona $F(x) = 1 - e^{-\lambda x}$, a stąd po wygenerowaniu wielkości R znajdziemy:

$$x = -\frac{1}{\lambda} \ln(1 - R) \quad (6)$$

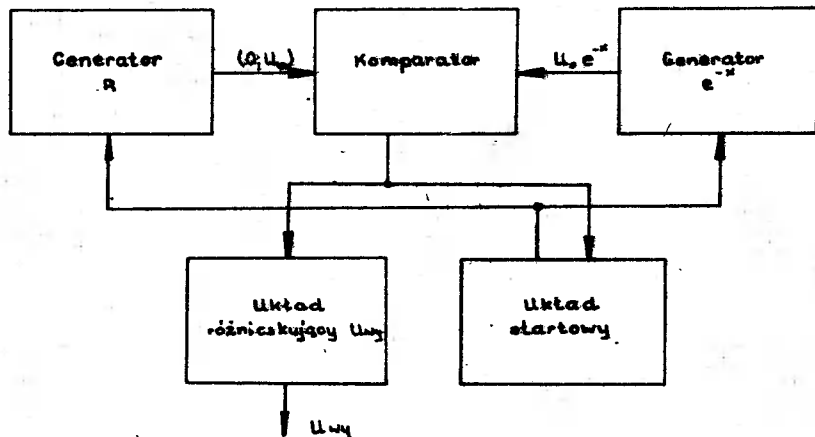
Ponieważ zmienna losowa o rozkładzie jednostajnym przyjmuje wszystkie wartości w przedziale $(0, 1)$ zatem można przyjąć, że:

$$x = -\frac{1}{\lambda} \ln(R) \quad (7)$$

Przedstawiona metoda jest podstawą działania zbudowanego generatora strumienia Poissona. Podobną zasadę wykorzystuje generator opisany w pracy [6], gdzie w miejsce rozkładu wykładniczego przyjęto rozkład geometryczny zbiegający asymptotycznie do rozkładu wykładniczego.

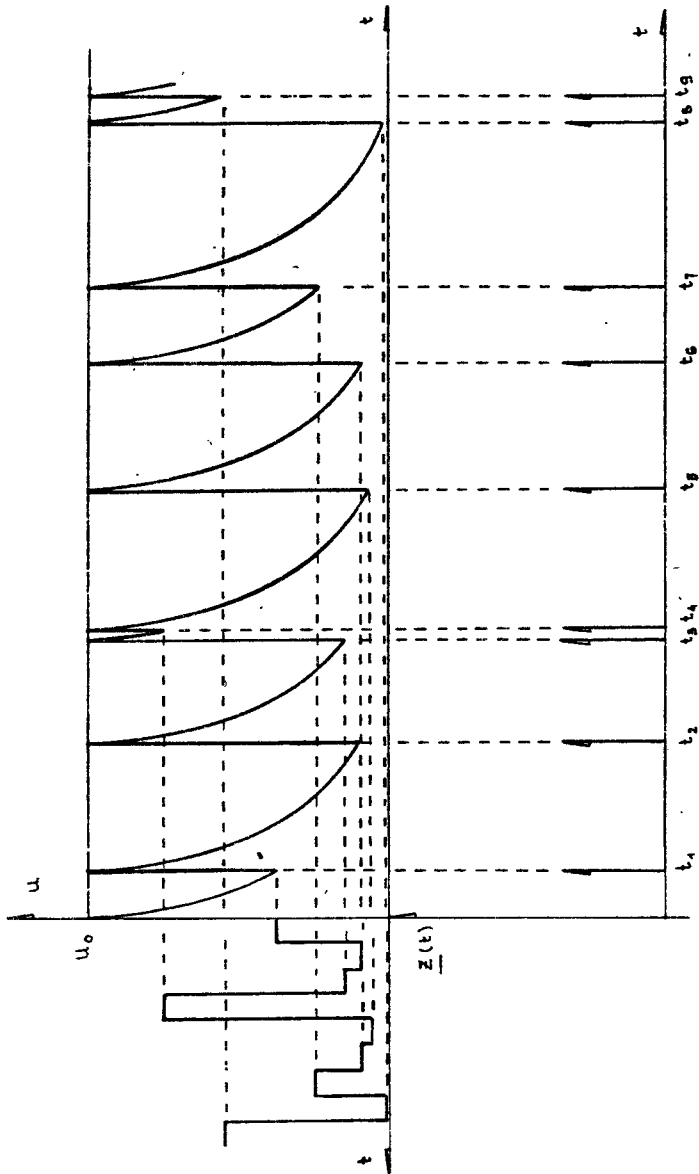
3. REALIZACJA UKŁADU

Jak wynika ze wzoru (7) dla każdej wygenerowanej liczby R należy wyznaczyć jej logarytm i tak obliczoną wartość traktować jako czas. Ponieważ taki sposób generacji zmiennej losowej o rozkładzie wykładniczym jest uciążliwy, dlatego też zasada jaką przyjęto przy budowie generatora polega na wykorzystaniu generatora funkcji $A e^{-\lambda t}$. Schemat blokowy takiego generatora ilustrujący podstawową ideę rozwiązania przedstawiony został na rys.2, zaś przykładowy przebieg czasowy przedstawiono na rys.3.



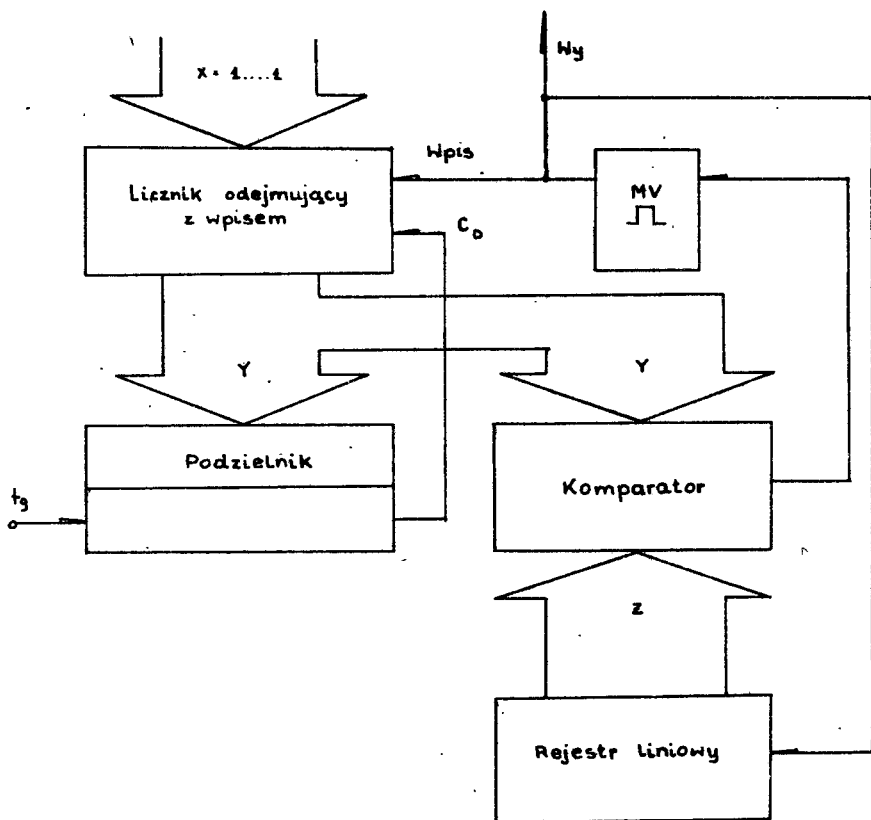
Rys.2. Schemat blokowy generatora prostego strumienia Poissona z wykorzystaniem metody generacji czasów o rozkładzie wykładniczym

Ponieważ przyjęto zasadę, że generator powinien być zbudowany na elementach cyfrowych, stąd jako generator rozkładu jednostajnego przyjęto



Rys. 3. Schemat blokowy cyfrowego poissonowskiego generatora zdarzeń

rejestr liniowy. Rejestr liniowy składa się z n -stopniowego rejestru przesuwanego z odpowiednio dobraną pętlą sprzężenia zwrotnego zrealizowaną na sumatorze mod.2. Doboru pętli sprzężenia zwrotnego dokonuje się tak, aby generator liczb pseudolosowych generował sekwencję o maksymalnej długości, tzn. o okresie powtarzania $T = 2^n - 1$. Taka sekwencja spełnia trzy warunki zwane cechami losowości, a mianowicie: właściwość zrównoważenia, właściwość serii oraz właściwość korelacji. Można ją zatem uznać jako sekwencję pseudolosową [9]. Dla generacji funkcji wykładniczej zastosowano w modelu standardowy układ wykorzystujący podzielnik [5]. Podstawowy schemat cyfrowego układu generatora przedstawiono na rys.4.



Rys.4. Schemat blokowy cyfrowego poissonowskiego generatora impulsów

-Podzielnik z licznikiem odejmującym generuje na wyjściu wielkość

$$Y = X e^{-\lambda t}$$

gdzie:

$$\lambda = \frac{f_g}{(2^n - 1)} \quad / f_g - \text{częstotliwość generatora podstawowego/}$$

n - długość podzielnika

$$X = 111 \dots 1$$

Rejestr liniowy generuje liczbę R o rozkładzie jednostajnym. Z chwilą zrównania się na komparatorze wielkości Y oraz R generowany jest krótki impuls na uniwibratorze MV, który:

- wpisuje $X = 111 \dots 1$ do licznika odejmującego,
- przesuwa rejestr liniowy o jedną pozycję /generuje nową liczbę R/,
- rozpoczyna cykl generacji funkcji wykładniczej.

W trakcie badania modelu okazało się, że istotny wpływ na pracę urządzenia posiada zakres dobrego przybliżenia funkcji $A e^{-\lambda t}$, a ściślej czas t mierzony w jednostkach stałej λ , dla której uzyskana wartość $A e^{-\lambda t}$ nie odbiega od wyliczonej. Czas ten jest tym większy im więcej pozycji binarnych zawiera licznik odejmujący i podzielnik. Dlatego też w praktycznej realizacji przyjęto, że układ generacji funkcji wykładniczej jest 20-bitowy, zaś rejestr liniowy jest 16-bitowy. Komparacja liczb odbywała się z wykorzystaniem 16 najbardziej znaczących bitów podzielnika. Ważnym czynnikiem mającym również wpływ na właściwą pracę badanego modelu okazał się sposób generacji liczb pseudolosowych na rejestrze liniowym. Jak wykazano w pracy Morgana [7], aby uzyskać na rejestrze liniowym nieskorelowane ciągi binarne należy generować rozłączne sekwencje takich ciągów. Dlatego też w trakcie realizacji modelu przyjęto, że rejestr liniowy pracuje cały czas z częstotliwością $f_g/10^4$, zaś w momencie wygenerowania impulsu wyjściowego liczba zawarta w rejestrze liniowym jest przepisywana do rejestru buforowego i tam przechowywana do momentu wygenerowania kolejnego impulsu wyjściowego. Tak zrealizowany układ poddano testowaniu.

4. WYNIKI POMIARÓW

Generator prostego strumienia Poissona testowano w następujący sposób: w ustalonej jednostce czasu mierzono liczbę impulsów pojawiającą się na wyjściu generatora. Następnie znajdowano rozkład liczby impulsów w jednostce czasu i sprawdzano za pomocą testu χ^2 słuszność hipotezy, że liczba impulsów wychodzących z generatora w ustalonej jednostce czasu podlega rozkładowi Poissona. Pomiar czasu odbywał się poprzez zliczanie impulsów wejściowych z generatora sterującego /generatora częstotliwości f_g /, a następnie blokowaniu tego generatora na czas odczytu. W ten sposób można było dokonać pomiaru w rzeczywistych warunkach pracy generatora.

Otrzymano następujące wyniki:

1/ Podzielnik 16-bitowy

Liczba impulsów	Liczba wyników	Częstość pomierzona	λ	Prawdopodobieństwo obliczone
0	150	0,3106	1,1076	0,3303
1	190	0,3934		0,3659
2	93	0,1925		0,2026
3	42	0,0869		0,0748
4	7	0,0145		0,0207
5	1	0,0021		0,0046
	483			

$\chi^2 = 4,3267$; χ_{kr}^2 przy poziomie istotności 0,05 i trzech stopniach swobody /zaniedbujemy wynik dla pięciu impulsów/ jest równy 7,81.

W n i o s e k : hipoteza o rozkładzie Poissona uzyskanych wyników nie może zostać odrzucona.

2/ Podzielnik 20-bitowy

a/ czas równy jednej jednostce

Liczba impulsów	Liczba wyników	Częstość pomierzona	λ	Prawdopodobieństwo obliczone
0	260	0,3714	1,0226	0,3597
1	251	0,3586		0,3678
2	120	0,1714		0,1881
3	54	0,0771		0,0641
4	12	0,0172		0,01638
5	3	0,0043		0,00335
	700			

$\chi^2 = 3,7024$; χ_{kr}^2 przy poziomie istotności 0,05 i czterech stopniach swobody jest równe 9,49.

W n i o s e k : hipoteza o rozkładzie Poissona nie może zostać odrzucona.

b/ czas równy dwóm jednostkom

Liczba impulsów	Liczba wyników	Częstość pomierzona	λ	Prawdopodobieństwo obliczone
0	77	0,15127	2,0196	0,13270
1	130	0,25540		0,26801
2	131	0,25737		0,27064
3	95	0,18664		0,18220
4	47	0,09234		0,091996
5	18	0,03536		0,037160
6	7	0,01375		0,012508
7	2	0,00392		0,003609
8	1	0,00196		-
9	1	0,00196		-
	509			

$\chi^2 = 2,13355$; χ_{kr}^2 przy poziomie istotności 0,05 i sześciu stopniach swobody /zaniedbujemy wyniki zawierające osiem i dziewięć impulsów/ jest równe 12,59.

Wniosek: hipoteza o rozkładzie Poissona nie może zostać odrzucona.

Uzyskane wyniki potwierdzają słuszność wybranej koncepcji, gdyż wynika z nich, że nie można odrzucić hipotezy o tym, że badany model jest generatorem prostego strumienia Poissona.

Zbudowany model generatora strumienia poissonowskiego służyć będzie dla symulacji zakłóceń impulsowych przy badaniach jakościowych łącz telekomunikacyjnych. Należy zauważyć, że niewielka modyfikacja układu pozwala utworzyć też generator impulsów o bardziej złożonych rozkładach czasów pomiędzy pojawiającymi się impulsami.

L i t e r a t u r a

- [1] de Coulon F., Rumley R.A.: Pseudorandom Poisson - Process Generator. Electronics Letters vol.9 No 20, 1973, s.466-467
- [2] Hartley M.G.: Development Design and Test Procedures for Random Generators Using Chaincodes. Proc. IEEE No 116, 1969, s.22-26
- [3] Hartley M.G.: Evaluation and Performance of Random Generators Employing Chaincodes. ibid s.27-34
- [4] Kopociński B.: Zarys teorii odnowy i niezawodności. PWN, Warszawa 1979
- [5] Misiurewicz P., Grzybek M.: Półprzewodnikowe układy logiczne. WNT, Warszawa 1975

- [6] Moisin L.H.: Generator de impulsuri aleatoare simulator de canal de transmisie perturbat. Elektrotehnica, Electronica, Automatica vol.24, Nr 3, 1980, s.117-120
- [7] Morgan D.R.: Autocorrelation Function of Sequential M-Bit Words Taken from N-Bit Shift Register /PN/ Sequence. IEEE Trans. on Comp. No 5, 1980, s.408-410
- [8] Papoulis A.: Prawdopodobieństwo, zmienne losowe i procesy stochastyczne. WNT, Warszawa 1972
- [9] Zieliński R.: Generatory liczb losowych. WNT, Warszawa 1972

PSEUDORANDOM POISSON PULSE GENERATOR

Summary

The idea and a block diagram of random pulse generator with exponentially distributed time between pulses are given in the paper. Such a random generator is called a stochastic Poisson process generator. The results obtained from the measurements of the model confirm theoretical considerations.

ПСЕВДОСЛУЧАЙНЫЙ ГЕНЕРАТОР ПОТОКА ПУАССОНА

Резюме

В статье представлен принцип разработки и блок - схема генератора последовательностей случайных импульсов с экспоненциальным распределением времени между очередными импульсами. Такой генератор называется генератором простого потока Пуассона. Полученные результаты полностью подтверждают правильность принятой идеи.

Marian Molski
Bożydar Dubalski

KONCEPCJA GENERATORA CYFROWEGO WEKTORÓW LOSOWYCH O SKŁADOWYCH NIEZALEŻNYCH I DOWOLNYCH ROZKŁADACH

W artykule przedstawiono koncepcję generatora wektorów losowych. Składowe wektora losowego są binarne i niezależne. Generator taki może mieć zadawane prawdopodobieństwa co 0.01 z błędem procentowym 0.1%.

Omawiany generator ma możliwość zmiany trybu pracy tak, aby spełniał funkcję generatora dyskretnej /lecz nie binarnej/ zmiennej losowej.

W artykule wskazano także na szereg możliwych zastosowań takiego generatora.

1. WSTĘP

Generatory wektorów losowych /zwane dalej w skrócie GWL/, mających binarne i niezależne składowe, mogą być stosowane jako:

- symulatory uszkodzeń w węzłach modelowanej sprzętowo sieci teleinformatycznej,
- symulatory uszkodzeń chwilowych w projektowanych urządzeniach cyfrowych, przy ich analizie niezawodnościowej [7],
- bloki niezbędne przy symulowaniu automatów probabilistycznych [7],
- symulatory zakłóceń w kanałach transmisji danych [2],
- model środowiska losowego do badania zachowania automatów uczących się w tym środowisku [4],
- blok służący wraz z konwerterem c/a do konstruowania wybranych typów ciągłej zmiennej losowej.

Generatory takie - jak wiadomo - mogą być realizowane w sposób programowy bądź sprzętowy. W pierwszym przypadku stanowią one niezbędny składnik programów służących do symulacji badanego systemu na komputerze. Programy takie bazują na algorytmie generowania liczb pseudolosowych o rozkładzie równomiernym w przedziale $[0, 1]$ i stanowią element oprogramowania użytkowego komputerów. Są one obszernie opisywane w literaturze, np. [1, 11].

Poziom współczesnej technologii układów scalonych umożliwia sprzętowo realizację GWL o zadawanych rozkładach. W literaturze znane są rozwiązania dotyczące generatorów binarnej zmiennej losowej X o rozkładzie prawdopodobieństwa $P[X = 1] = P[X = 0] = \frac{1}{2}$ realizowane w oparciu o rejestr przesuwany z wprowadzonym odpowiednio sprzężeniem zwrotnym poprzez bramki typu

„suma modulo 2” /rejestr liniowy/ [8]. Znane są także sprzętowe realizacje generatorów zmiennych losowych o danym konkretnie typie rozkładu, np. Poissona [10].

Celem niniejszego opracowania jest przedstawienie koncepcji generatora wielu binarnych zmiennych losowych niezależnych X_i o zadawanych prawdopodobieństwach $P[X_i = 1] = p_i$ /czyli generatora wektora losowego $X = \langle X_1, X_2, \dots, X_n \rangle$ o binarnych składowych niezależnych/. Jednocześnie generator taki ma mieć możliwość, będącą niejako naturalnym efektem ubocznym, generowania jednej dyskretnej /lecz nie binarnej/ zmiennej losowej X przyjmującej jedną z n wartości również z zadawanymi prawdopodobieństwami - czyli $P[X = i] = p_i^*$, gdzie $i = 1, 2, \dots, n$. W tym drugim przypadku oczywiście też jest generowany wektor losowy, lecz w odróżnieniu od pierwszego typu generatora jego składowe tworzą kod „1” z n^* . Tak rozumiany generator dalej nazywać będziemy w skrócie GZL /generator zmiennej losowej/. Równość $X = i$ dla GZL rozumieć zaś będziemy jako odpowiednik tego, że na i -tym wyjściu GWL jest jedynka, a na wszystkich innych wyjściach jest zero /czyli $X_i = 1, X_j = 0$ dla $j \neq i$ /. Tak rozumiany GWL wraz z możliwością jego modyfikacji stanowi propozycję autorów, bowiem nie są znane w dostępnej literaturze opracowania dotyczące tego typu generatora.

Proponowany generator cechuje się dużą prostotą i przejrzystością rozwiązania, możliwością zmian zadawanych prawdopodobieństw co 0.01 z błędem 0.1% oraz tym, że może być zaprojektowany na elementach produkcji krajowej.

2. OKREŚLENIA I STOSOWANE SYMBOLE

Uznając definicje takich pojęć jak zmienna losowa, rozkład zmiennej losowej, binarna zmienna losowa, dyskretna zmienna losowa i jej rozkład za ogólnie znane i elementarne /patrz [9]/ ograniczymy się tu do wyszczególnienia występujących dalej w tekście oznaczeń:

$X = \langle X_1, X_2, \dots, X_n \rangle$ - wektor losowy o składowych binarnych i niezależnych,

X_i - i -ta składowa wektora losowego X ,

$P[X_i = 1] = p_i$ - prawdopodobieństwo tego, że i -ta składowa X przyjmie wartość równą jeden /dla GWL/,

$P[X_i = 0] = 1 - p_i$ - prawdopodobieństwo tego, że i -ta składowa X przyjmie wartość równą zero /dla GWL/,

$P[X = i] = p_i^*$ - prawdopodobieństwo tego, że tylko na i -tym wyjściu generatora /GZL/ pojawi się jedynka /gdzie

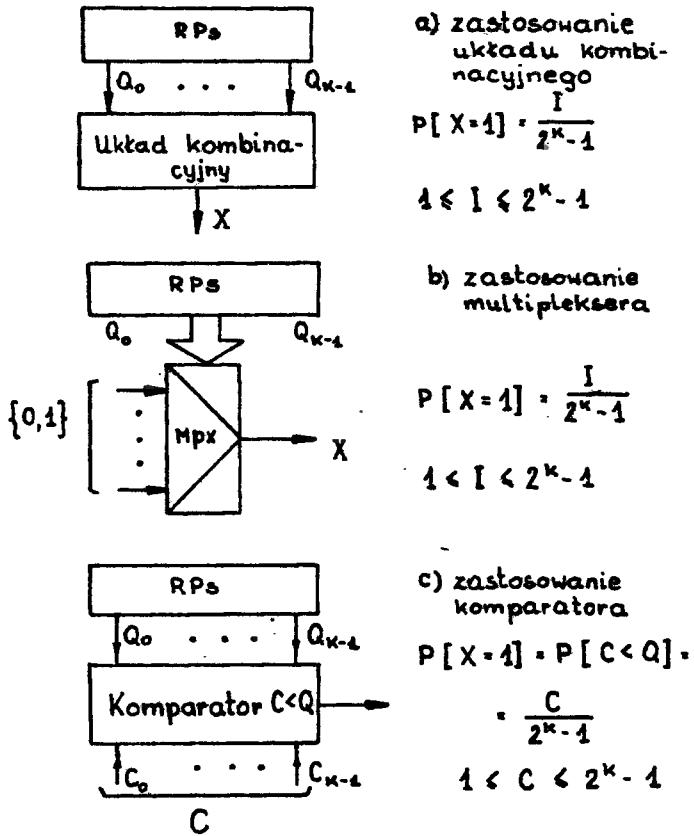
$$\sum_{i=1}^n p_i^* = 1 /,$$

- Q - rejestr liniowy /RL/, a także jego zawartość binarna,
 Q_1 - i-te wyjście rejestru liniowego Q ,
 k - liczba przerzutników tworzących rejestr liniowy /długość rejestru liniowego Q /,
 n - liczba składowych wektora losowego X ,
 r - wymagana liczba symulacji /obserwacji/ w zastosowaniach generatora,
 m - liczba wejść komparatora współpracującego z rejestrem liniowym Q ,
 C - liczba całkowita /pamiętana w pamięci RAM/, pomocnicza przy programowaniu wartości p_1 /patrz dalej wzór 2/,
RPS - rejestr przesuwany ze sprzężeniem,
 ξ - najmniejsza możliwa wartość p_1 ,
 Δ - błąd procentowy p_1 .

3. WYBÓR I UZASADNIENIE KONCEPCJI UKŁADOWEGO ROZWIĄZANIA GWL

Podobnie jak w przypadku programowej generacji zmiennych losowych, układem bazowym w realizacji sprzętowej jest z reguły generator zmiennej losowej o rozkładzie równomiernym w zbiorze liczb całkowitych z przedziału $[1, 2^k - 1]$, jeśli obserwowane są równoległe wyjścia Q_i RL [8]. Istnieją oczywiście inne metody generowania zmiennych losowych wykorzystujące naturalne procesy fizyczne, np. szum biały, lecz tutaj rezygnujemy z nich zachowując celowo „cyfrową jednorodność” proponowanego rozwiązania.

Podstawowym problemem przy przekształcaniu rozkładu równomiernego na dowolnie żądany inny rozkład zmiennej losowej jest wybór metody. Możliwości jakimi dysponujemy w technice cyfrowej zobrazowano na rys.1. Na rysunku tym podano wzory na $P[X_1 = 1]$ dla każdej wersji przekształcenia. Symbol „I” oznacza we wzorach liczbę składowych jedności funkcji boolewskiej realizowanej przez układ kombinacyjny /a/ lub liczbę jedynek na wejściach multipleksera jako uniwersalnego układu kombinacyjnego /b/, natomiast C jest liczbą kodowaną binarnie na wejściach komparatora. Układ /a/ nie posiada możliwości zadawania prawdopodobieństw i może być stosowany jedynie w prostych generatorach zmiennych losowych o zadanym na stałe rozkładzie prawdopodobieństw. Układ /b/ ma możliwość zadawania wielkości I, lecz przy dużych wartościach k /a tak jest najczęściej/, konieczne jest łączenie multiplekserów standardowych w struktury multiplekserowe bardzo rozbudowane, co jest zabiegiem nieopłacalnym. Układ /c/ ma możliwość zadawania stałej C , a ponadto proste jest iteracyjne łączenie typowych czterobitowych komparatorów. Tak więc w proponowanym rozwiązaniu zdecydowano się na przyjęcie wersji /c/ jako bazowej dla projektowanego GWL /GZL/.



Rys.1. Sposoby przekształceń typów rozkładów

4. SCHEMAT BLOKOWY I PARAMETRY GWL

Bazując na koncepcji /c/ wymienionej w punkcie 3 można zaproponować rozwiązanie konstrukcyjne GWL. Przy technicznej realizacji GWL konieczne jest zapewnienie:

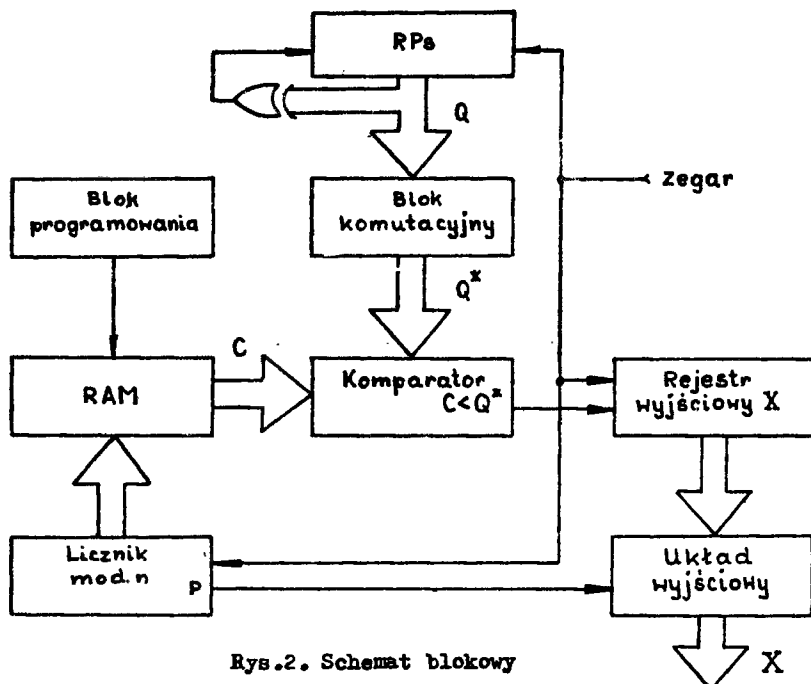
- a/ niezależności składowych X_i wektora X i możliwie małej autokorelacji dla poszczególnych składowych,
- b/ odpowiedniej długości cyklu podstawowego generatora liczb pseudolosowych zrealizowanego na RL,
- c/ możliwości zadawania dowolnych prawdopodobieństw $p_i / p_i^{\#}$,
- d/ żądanej dokładności zmian prawdopodobieństw $/\xi = 0.01/$ z określonym błędem $/\Delta = 0.1 \%/$.

Uwzględniając powyższe warunki należy wyznaczyć następujące parametry GWL: k, n, r, m, C .

Podstawowym założeniem jakie przyjęto jest $n = 16$. Wielkość ta będąca potęgą dwójki jest typową dla sprzętu cyfrowego i w zasadzie wystarczającą np. do laboratoryjnych zastosowań podczas modelowania bądź w procesie dydaktycznym. Poza tym wielkość ta jest wtórnie zdeterminowana pojemnością dostępnych na rynku modułów pamięci RAM /UCY 7498/, tj. 16 słów czterobitowych.

Ad. a/

Ponieważ każdy podciąg ciągu losowego jest ciągiem losowym, możemy na bazie RL generować kolejno $\dots, X_1, X_2, \dots, X_n; X_1, X_2, \dots, X_n, \dots$. Oczywiście kolejne X_1 generuje się po wykonaniu operacji porównania przez układ komparatora wygenerowanej przez RL liczby pseudolosowej Q z liczbą C_1 . Wykorzystanie co n -tej liczby pseudolosowej Q zapewnia najmniejszą autokorelację X_1 [6]. Największa niezależność pomiędzy poszczególnymi składowymi wektora X jest uzyskiwana przez zastosowanie bloku komutacji /rys.2/.



Rys.2. Schemat blokowy

Zadaniem bloku komutacji jest dokonanie odpowiednich połączeń wyjść RL, tj. $\langle Q_0, Q_1, \dots, Q_{k-1} \rangle$ z m wejściami komparatora w taki sposób, że do tych wejść dołączyć należy kolejno $Q_{1_0}, Q_{1_1}, \dots, Q_{1_{m-1}}$, gdzie $k > m$. Naj-

większą niezależność kolejno generowanych liczb pseudolosowych uzyskuje się przy obserwacji pierwszych i ostatnich $\frac{n}{2}$ wyjść RL [3]. Blok komutacji jest tutaj rozumiany po prostu jako system stałych połączeń rozumianych jak wyżej. Rozwiązaniem konstrukcyjnym mogą być dwie łączówki /do jednej dochodzą wyjścia RL, a do drugiej wejścia komparatora/ odpowiednio wymiennie krosowane, co ułatwiać może eksperymenty mające na celu badanie wpływu systemu połączeń między RL a komparatorem na stopień niezależności generowanych liczb pseudolosowych [5].

Ad. b/

Każdy generator liczb pseudolosowych zbudowany na RL ma określony cykl, stąd oczywisty jest warunek:

$$nr < 2^k - 1 \quad (1)$$

Gdy $n = 16$, natomiast $r = 50\ 000$ /co wystarcza na ogół w każdym procesie symulacji/, wówczas łatwo obliczyć, że $k = 20$ spełnia nierówność (1). Aby zapewnić maksymalny cykl, tj. $2^{20} - 1 = 1\ 048\ 575$ trzeba zastosować sprzężenie zwrotne typu $Q_{19} \oplus Q_2$ [8] /zestawy takich sprzężeń podawane są tabela - rycznie w podręcznikach z techniki cyfrowej, a teoretycznie wynikają one z rozważań dotyczących wielomianów konstruowanych w ciele Galois/.

Ad. c/

Możliwość zadawania prawdopodobieństw może być zapewniona poprzez pamiętanie w pamięci RAM tablicy stałych C_1 , gdzie /patrz też rys.1c/

$$C_1 = (2^m - 1) \cdot p_1 \quad (2)$$

Na rys.2 nie uwzględniono szczegółów dotyczących bloku programującego pamięć RAM.

Ad. d/

Najmniejszą wartością p_1 jest $\epsilon = 0.01$. Można przykładowo pokazać, że $n = 8$ nie gwarantuje wymaganej dokładności Δ dla zadawanego prawdopodobieństwa p_1 , ponieważ (2)

$$C_1 = 0.01 \cdot 255 = 2.55 \approx 3$$

Błąd procentowy wynikający z zaokrąglenia C_1 do najbliższej liczby całkowitej wynosi wówczas $\approx 20\%$. Jak łatwo sprawdzić $n = 12$ gwarantuje nam żądaną dokładność Δ dla p_1 , gdyż $C_1 = 40.95 \approx 41$. Błąd procentowy wynikający z zaokrąglenia wynosi w tym przypadku

$$\Delta = \frac{41 - 40.95}{40.95} \cdot 100\% = \frac{0.05}{40.95} \cdot 100\% \approx 0.1\%$$

Przy tak zaproponowanym rozwiązaniu GWL możliwe jest również generowanie jednej zmiennej losowej dyskretnej X /dotyczy GZL/ przyjmującej jedną

z szesnastu wartości, również z zadawanymi prawdopodobieństwami p_i^* /patrz punkt 2/, gdzie

$$p_i^* = P [X = i] ; \quad i = 1, 2, \dots, 16$$

oraz

$$\sum_{i=1}^n p_i^* = 1$$

Prawdopodobieństwo przyjęcia przez zmienną losową jednej z szesnastu wartości jest określone następująco:

$$P [X = i] = P \left[X = \min_j \left\{ j : Q < \sum_{l=1}^j C_l \right\} \right] = p_i^* \quad (3)$$

Interpretacja wzoru (3) jest następująca: wygenerowaną liczbę pseudolosową Q /po bloku komutacji/, a więc liczbę z przedziału $[1, 2^m - 1]$ porównywać należy kolejno z liczbami $C_1, C_1 + C_2, C_1 + C_2 + C_3, \dots, C_1 + C_2 + C_3 + \dots + C_j$. Najmniejsza wartość j / $j = 1, 2, \dots, n$ /, przy której

$$Q < \sum_{l=1}^j C_l$$

jest wygenerowaną wartością dyskretnej zmiennej losowej. Takie postępowanie jest niczym innym jak metodą odwracania dystrybuanty stosowaną przy zamianie typu rozkładu dyskretnej zmiennej losowej z równomiernego na dowolny inny. W rozwiązaniu układowym oznacza to potrzebę sumowania stałych C_l przy wprowadzaniu do pamięci. W praktyce sprowadza się to do minimalnej modyfikacji bloku programowania pamięci.

Zauważmy, że w przypadku GWL możliwe jest pojawienie się jednocześnie jedynek na kilku spośród n wyjść, bowiem każde wyjście realizuje fizycznie inną niezależną składową wektora binarnego. W przypadku GZL jedynka może pojawić się tylko na jednym /i-tym/ wyjściu generatora symbolizując w ten sposób fakt, że dyskretna obecnie /a nie binarna/ zmienna losowa przyjmuje jedną z szesnastu /i-tą/ możliwych wartości. Tak więc wektor losowy /fizycznie/ na wyjściu generatora jest wtedy interpretowany /matematycznie/ jako pojedyncza dyskretna zmienna losowa przyjmująca jedną z szesnastu wartości. Zaprojektowany generator winien mieć możliwość przełączania trybu pracy z GWL na GZL. Poza tym należy pamiętać przy wprowadzaniu stałych do pamięci /faza programowania generatora/, że stałe C_l /w przypadku GZL/ powinny spełniać warunek normalizacyjny, tj.

$$\sum_{i=1}^n p_i^* = 1$$

natomiast w przypadku GWL taki warunek po prostu nie ma sensu, bowiem jest on spełniony dla każdej składowej binarnej oddzielnie $P [X_i = 1] = p_i^*$,
 $P [X_i = 0] = 1 - p_i^*$.

5. WNIOSKI

W pracy tej nie zamieszczono szczegółowego schematu logicznego i montażowego, ponieważ założeniem było opracowanie koncepcji konstrukcji GWL. Szczegóły konstrukcyjne opisano dokładnie w innej publikacji autorów [5]. Nadmienimy jedynie tutaj, że w sumie potrzeba do realizacji technicznej GWL około czterdziestu dostępnych na rynku cyfrowych elementów scalonych małej i średniej skali generacji. Z tego faktu wypływa istotny wniosek, że nieopłacalne jest zastosowanie mikroprocesora w konstrukcji omawianego generatora. Sygnalizowane we wstępie aplikacje takiego generatora jak i doświadczenie zdobyte przez autorów przy różnych eksperymentach badawczych mogą stanowić uzasadnienie celowości zajmowania się opisaną problematyką.

Literatura

- [1] Fishman G.S.: Symulacja komputerowa, pojęcia i metody. PWE, Warszawa 1981
- [2] Jabłoński M.: Generator ciągów pseudolosowych do modelowania kanałów, Elementy półprzewodnikowe i układy scalone, Zastosowania, Układy cyfrowe. Unitra CEMI PIE, Warszawa, nr 3/1981
- [3] Krause J.: Generatory pseudolosowe stacjonarnego procesu stochastycznego. Praca dyplomowa, ITiE ATR Bydgoszcz, 1980
- [4] Molski M.: Model automatowy uczącej się struktury sterującej. Materiały V Sympozjum „Metody matematyczne w elektrotechnice”. PTETiS i PW, Zakopane - Warszawa 1977
- [5] Molski M., Dubalski B.: Generator binarnych zmiennych losowych niezależnych o dowolnych rozkładach. Elementy półprzewodnikowe i układy scalone, Zastosowania, Układy cyfrowe. Unitra CEMI PIE, Warszawa Nr 4/1982 /w druku/
- [6] Morgen D.R.: Autocorrelation function of sequential M-bit words taken from an N-bit shift register /PN/ sequence. IEEE Transaction on Computer, No 5, 1980
- [7] Pospiełowski D.A.: Automaty probabilistyczne. WNT, Warszawa 1972
- [8] Pieńkos J., Turczyński J.: Układy scalone TTL serii UCY 74 i ich zastosowanie. WKŁ, Warszawa 1976
- [9] Sobczak W.: Metody statystyczne w elektronice. WNT, Warszawa 1971
- [10] Zabłudowski A., Dubalski B.: Pseudolosowy generator prostego strumienia Poissona. Wiadomości Telekomunikacyjne, nr 2/1982, Warszawa
- [11] Zieliński R.: Generatory liczb pseudolosowych. WNT, Warszawa 1979

A CONCEPT OF A RANDOM VECTOR GENERATOR

Summary

The author presents a concept of a random vector generator circuit. The components of the random vector are binary and statistically independent. The generator has a variable probability with $0,01^{\text{th}}$ step and $0,1\%$ accuracy. There is a possibility of changing the operation mode in such a way that functions of the non-continuous /but not binary/ random variable are fulfilled. Several examples of the application of the generator are described.

КОНЦЕПЦИЯ ЦИФРОВОГО ГЕНЕРАТОРА СЛУЧАЙНЫХ ВЕКТОРОВ С НЕЗАВИСИМЫМИ СОСТАВЛЯЮЩИМИ И ПРОИЗВОЛЬНЫМИ РАСПРЕДЕЛЕНИЯМИ

Резюме

В статье представлен принцип построения генератора случайных векторов. Компоненты случайного вектора бинарные и независимые. В таком генераторе можно устанавливать вероятность с распределением $0,01$ и точность $0,1\%$. Он может также выполнять работу генератора дискретной /но не бинарной/ случайной величины. В статье приведено также несколько возможных применений этого генератора.

Jerzy Piotr Pawlak

FUNKCJONALNA STRUKTURA PROCESU
AWARYJNEGO ZAMYKANIA RUCHU WEWNĘTRZNEGO
W KONCENTRATORZE CSADz CENTRALI ELEKTRONICZNEJ E-10

W artykule przedstawiono metodę zapewnienia łączności abonentom przyłączonym do odległego koncentratora elektronicznej centrali telefonicznej E-10 w przypadku awarii całkowicie uniemożliwiającej realizację połączeń przez centralę. Praktyczne rozwiązanie problemu zamykania ruchu zawiera dodatkowe wyposażenie sprzętowe i programowe tylko wewnątrz koncentratora CSAD przy minimalnej ingerencji w jego dotychczasową normalną pracę. Awaryjną łączność wewnętrzną realizuje się przy pomocy dodatkowego, specjalnego oprogramowania ELS sterującego pracą dodatkowych układów rozpoznawania stanu awarii, odbioru, analizy i przeliczania numerów, polem czasowej komutacji lokalnej oraz zapewniających wykonywanie większości normalnych funkcji koncentratora.

Artykuł jest funkcjonalnym opisem zrealizowanej konstrukcji przedstawionej w [2] i opracowanej przy współpracy z Instytutem łączności. Stosowane w artykule skróty i symbole pochodzą z języka francuskiego.

1. WSTĘP

1.1. Miejsce i funkcje koncentratora CSAD w centrali E-10

W systemie komutacji elektronicznej E-10 można wyróżnić dwa rodzaje zespołów:

- zespoły przeznaczone do łączenia abonentów z kanałami czasowymi VT /komutacja przestrzenno-czasowa/,
- zespoły przeznaczone do łączenia kanałów rozmównych VT między sobą /komutacja czasowa/.

Pierwszą grupę zespołów stanowią koncentratory CSA /lokalne - CSAL i odległe - CSAD/, drugą zaś stanowi sama centrala elektroniczna ze swoim polem komutacji czasowej CX, procesorami MQ i MR, przelicznikiem, zespołem sygnałowym ETA itp.

W koncentratorze CSA dokonuje się koncentracji 512 abonentów do 60 łączy. Całkowity ruch z 60 zespołów połączeniowych JCT załatwiany jest przez 2 trakty PCM. Koncentrator CSAD jest urządzeniem fizycznie oddalonym od centrali na odległość do kilkudziesięciu km i w tym przypadku sygnały mowy /30 kanałów czasowych/ jak i informacje sygnalizacyjne /w dwóch po-

zostających kanałach czasowych/ są transmitowane po tych samych torach metalicznych /po tej samej parze przewodów/.

W koncentratorach, poza podstawowym zadaniem, jakim jest koncentracja przestrzenna abonentów, realizowane są następujące funkcje:

- poszukiwanie i zestawienie dróg połączeniowych w polach koncentracji od łączy abonenckich do kanałów PCM,
- nadzór nad stanem łączy abonenckich, polegający na ciągłym ich przepatrywaniu i wykrywaniu nowych wywołań oraz rezygnacji z połączenia,
- konwersja analogowo-cyfrowa sygnałów rozmównych,
- zasilanie łączy abonenckich i wysyłanie prądu dzwonienia do abonenta wywoływanego oraz zmiana biegunowości baterii,
- wymiana sygnalizacji i meldunków z centralą.

Wykrywanie nowych wywołań, wyszukiwanie fałszywych wywołań, wybór drogi połączeniowej itp. są dokonywane przez układ logiczny sterowania stopniem koncentracji, który jest specjalizowanym procesorem pracującym w oparciu o stałe programy zapisane w pamięci ROM, nazywany standardowym wyposażeniem logicznym /ELS/.

1.2. Kryteria awarii

System sygnalizacyjny stosowany w centralach E-10 dla koncentratorów odległych polega na przesyłaniu „wiadomości” po tych samych łączach transmisyjnych, które są wykorzystywane do przesyłania sygnałów rozmównych, w odpowiednich kanałach czasowych. Jednocześnie w tych wybranych i ściśle określonych kanałach przesyłane są kody synchronizacji ramki /MVTR/ i wieloramki /MVMTR/. Dla zabezpieczenia przed przekłamaniami przesyłanych informacji stosuje się tzw. kod poprawnej transmisji wykorzystując kod Hamminga. W celu zapewnienia korekcji zniekształceń tłumieniowych i fazowych wprowadzanych przez tor transmisyjny PCM, co około 1,8 km włącza się w tor wzmacniacze - regeneratory, których ilość, dla przypadku koncentratora CSAD, może sięgać kilkunastu.

Awaria łączności koncentratora z centralą może wystąpić w przypadkach:

- fizycznego przerwania lub uszkodzenia traktów,
- uszkodzenia regeneratorów,
- utraty synchronizacji,
- wystąpienia błędów w transmisji informacji objawiających się niezgodnością kodu poprawnej transmisji.

Błędne słowa synchronizacji lub ich brak jest objawem awarii określonej przez pierwsze trzy przypadki. Specjalny układ liczenia błędów synchronizacji generuje sygnały stanów niesprawności poszczególnych traktów. Podobny układ zliczający błędne kody Hamminga informuje zespół sterujący koncentratorom o wystąpieniu niepoprawnej transmisji meldunków. Przełączenie na pracę awaryjną i zamykanie ruchu wewnętrznego dokonuje się wówczas, gdy obydwa trakty jednocześnie nie mają prawidłowej synchronizacji albo

występuje, świadcząca o błędnej transmisji informacji, określona liczba błędów kodu Hamminga.

Możliwe jest ponadto przełączenie /np. w celach diagnostycznych i innych/ koncentratora na pracę tylko w zamkniętym ruchu wewnętrznym oraz przywrócenie pracy normalnej za pomocą specjalnych rozkazów zdalnie przekazanych przez operatora z CTI.

2. PROCES AWARYJNEGO ZAMYKANIA RUCHU WEWNĘTRZNEGO W CSADz

2.1. Funkcje koncentratora CSADz

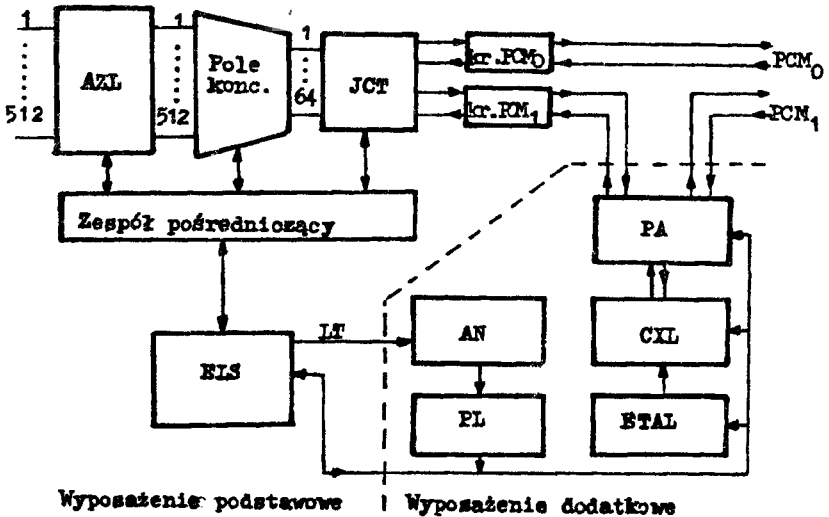
Koncentrator odległy /CSAD/ z możliwością zamykania ruchu wewnętrznego /CSADz/ musi wykonywać w czasie pracy awaryjnej dodatkowe funkcje przede wszystkim związane z komutacją czasową, odbiorem i analizą wybieranego numeru, które w pracy normalnej wykonuje centrala. Analizując proces nawiązywania i realizacji połączenia wewnętrznego można wyróżnić następujące czynności koncentratora CSADz:

- a/ po podniesieniu mikrotelefonu przez abonenta wywołującego /ab A/ poszukiwanie tzw. drogi wewnętrznej między odpowiednim zespołem połączeniowym JCT /z zakresu przeznaczonych do łączności wewnętrznej/ a wyposażeniem indywidualnym abonenta generującego nowe wywołanie;
- b/ zestawienie tej drogi i przesłanie do ab A sygnału imitującego sygnał zgłoszenia centrali z dodatkowego lokalnego zespołu sygnałowego ETAL;
- c/ odbiór, rejestracja i analiza numeru wybieranego przez ab A;
- d/ w przypadku, gdy numer abonenta żądanego /ab B/ jest różny od numerów obsługiwanych przez koncentrator z zamykaniem ruchu wewnętrznego, wysłanie do ab A sygnału nieosiągalności z ETAL;
- e/ dla ab B podłączonego do CSADz przeliczenie wybranego przez ab A numeru katalogowego na numer wyposażenia indywidualnego i poszukiwanie drogi wewnętrznej między ab B i odpowiednim JCT;
- f/ testowanie stanu ab B - gdy jest wolny i gdy istnieje do niego droga wewnętrzna, zestawienie tej drogi połączeniowej i wysłanie prądu dzwonienia do ab B oraz sygnału zwrotnego dzwonienia do ab A, jeżeli natomiast ab B jest zajęty lub nie istnieje do niego aktualnie wolna droga wewnętrzna - wysłanie do ab A sygnału zajętości;
- g/ po podniesieniu mikrotelefonu przez ab B, połączenie kanałów czasowych abonentów A i B w lokalnym cyfrowym polu komutacyjnym CXL;
- h/ nadzór nad połączeniem, a w szczególności czynności związane z obsługą przeniesienia aparatu przez ab B /na czas około 1 minuty/ i odłożeniem mikrotelefonu przez ab A /zakończenie rozmowy/;
- i/ rozłączenie abonentów w polu CXL i uwolnienie dróg połączeniowych /zachowane są te same zasady jak w przypadku rozłączania przez centralę/.

Przyjęto zasadę, że połączenia lokalne nawiązywane w czasie awarii łączności z centralą /z powodu znacznego ograniczania usług telefonicznych dla abonentów/ nie są taryfikowane.

2.2. Układy wyposażenia dodatkowego CSADz

Koncentrator CSADz wyposażono dodatkowo w układy realizujące te z wyżej wymienionych funkcji, których nie spełnia odległy standardowy koncentrator CSAD. Powiązanie najistotniejszych układów dodatkowych z blokami koncentratora przedstawiono na rys.1.



Rys.1. Schemat blokowy CSADz

Układami realizującymi zamykanie ruchu wewnętrznego są:

- PA - układ przełączania awaryjnego,
- CXL - lokalne cyfrowe pole komutacyjne,
- ETAL - lokalny zespół sygnałowy,
- AN - analizator numeru abonenta,
- PL - przelicznik lokalny.

Układ przełączania awaryjnego służy do dołączenia traktu LR₁ do pola CXL /kanałów wychodzących z krotnicy PCM do pola CXL i z CXL do krotnicy PCM/ lub do toru transmisyjnego do centrali.

Pole komutacyjne CXL służy do połączenia par kanałów czasowych abonentów A i B dla rozmowy lokalnej. O zestawionym w ten sposób połączeniu i rozmowie abonentów A i B nie ma żadnych informacji w pamięciach centrali, gdyż proces połączeniowy realizowany był bez udziału centrali i w przypadku braku jakiegokolwiek z nią prawidłowej łączności.

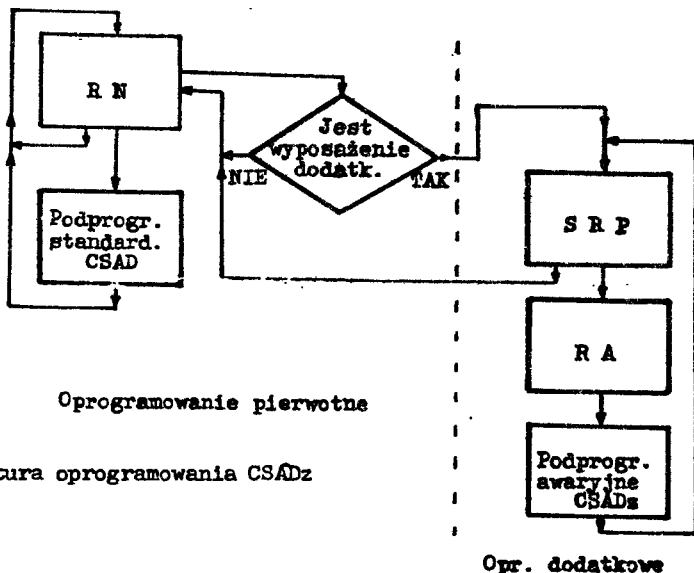
Zespół sygnałowy ETAL generuje wszystkie potrzebne sygnały tonowe zastępując w CSADz podczas awarii odpowiedni zespół centrali.

Analizator numeru abonenta i przelicznik lokalny służą do realizacji wszystkich czynności związanych z odbiorem numeru wybieranego przez ab A, rozpoznaniem czy ab B jest podłączony do CSADz oraz przeliczeniem dziesiętnego numeru katalogowego na binarny numer wyposażenia indywidualnego.

Sprzęt dodatkowy obejmuje również /nie uwzględnione na rys.1/ układy analizy stanów traktów współpracujące ściśle z układami odbioru sygnalizacji i synchronizacji koncentratora.

2.3. Struktura oprogramowania CSADz

Funkcje sterowania zespołami koncentratora realizuje wyspecjalizowany układ logiczny, tzw. procesor ELS, pracujący z programem zapisanym w pamięci ROM. Wprowadzenie wielu dodatkowych funkcji rozpoznawania stanu awarii i zamykania ruchu lokalnego wiąże się z rozbudową oprogramowania i powiększeniem pamięci stałej. Nadrzędną zasadą przy tworzeniu programów koncentratora w wersji z zamykaniem ruchu wewnętrznego było dokonanie jak najmniejszych zmian w istniejącym oprogramowaniu koncentratora odległego. Zastosowanie więc dodatkowego bloku pamięci programu daje możliwość uniknięcia produkcji dwóch różnych typów koncentratorów /z zamykaniem ruchu lub bez zamykania/, ponieważ każdy typowy koncentrator po wyposażeniu w dodatkowy sprzęt i oprogramowanie może pracować jako koncentrator z zamykaniem ruchu wewnętrznego. Jedyną zmianą w standardowym oprogramowaniu jest wstawienie warunku sprawdzającego czy CSAD wyposażono w dodatkowy program i sprzęt dla zamykania ruchu /rys.2/.

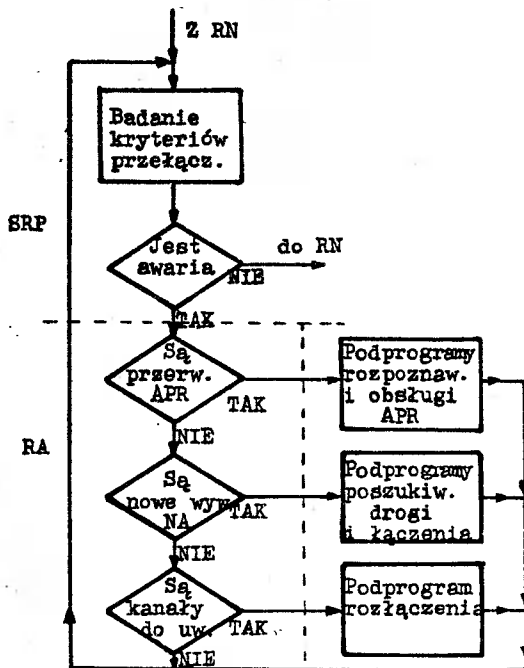


Rys.2. Struktura oprogramowania CSADz

W przypadku, gdy koncentrator nie jest wyposażony w dodatkowe układy, praca jego przebiega według harmonogramu ustalonego przez standardowe oprogramowanie /Rutyna "normalna" /RN/ i podprogramy istniejące/. Obecność sprzętu dodatkowego powoduje przejście do realizacji podprogramu Sterowania Rodzajem Pracy /SRP/, który na podstawie ustalonych kryteriów ma możliwość przełączenia koncentratora na pracę awaryjną lub kontynuowanie pracy "normalnej". W pierwszym przypadku wykonywany jest program RUTYNA AWARYJNA /RA/ i podprogramy awaryjne, w drugim - realizacja programu RN. Z dotychczasowych wyjaśnień wynika, że programy RN i RA pracują alternatywnie nie wykorzystując nawzajem swoich podprogramów. Jedynie podprogram zerowania jest wspólny, co w tym przypadku nie zakłóca pracy normalnej lub awaryjnej.

Dodatkowe, awaryjne oprogramowanie składa się z następujących elementów /rys.3/:

- programu sterowania rodzajem pracy /SRP/,
- programu przepatrywania cyklicznego /rutyna awaryjna - RA/,
- podprogramów rozpoznawania i obsługi wywołań priorytetowych, tzw. APR,
- podprogramów użytkowych /poszukiwania drogi połączeniowej, łączenia i rozłączenia/.



Rys.3. Oprogramowanie pracy awaryjnej

SRP spełnia rolę analizatora kryteriów przełączania na pracę awaryjną, wprowadzania stanu poawaryjnego i powrotu do pracy normalnej oraz wydaje odpowiednie decyzje.

RA jest awaryjnym odpowiednikiem standardowego programu przepatrywania cyklicznego, którego zadaniami są:

- przepatrywanie abonentkich zespołów liniowych w celu wykrycia nowych zgłoszeń /NA/,
- przepatrywanie kanałów czasowych VT do uwolnienia,
- umożliwienie przerwania cyklicznej pracy programu koncentratora przez wywołania priorytetowe /APR/.

Zadaniem podprogramu analizy i obsługi wywołań priorytetowych jest określenie rodzaju APR, zidentyfikowanie jego źródła, wyznaczenie właściwego podprogramu do realizacji zadań związanych z obsługą przerw i wywołanie, w zależności od potrzeb, odpowiedniego programu użytkowego.

Wśród podprogramów użytkowych można wyróżnić szczególnie podprogramy łączenia i rozłączenia kanałów czasowych w polu CXL, podprogram przeliczania numeru abonenta i podprogram poszukiwania drogi połączeniowej w zakresie kanałów wydzielonych do połączeń wewnętrznych.

3. UWAGI KOŃCOWE

Zamykanie ruchu wewnętrznego w koncentratorze odległym w przypadku awarii obydwu traktów stanowi praktyczne rozwiązanie problemu podtrzymania ruchu telefonicznego w obrębie np. osiedla lub miejscowości znacznie oddalonej od większej aglomeracji, w której znajduje się elektroniczna centrala E-10. Funkcję realizacji połączeń wewnętrznych spełnia się tylko w przypadku całkowitej awarii jednocześnie obydwu traktów PCM. Fakt realizacji łączności lokalnej może sugerować czytelnikowi przypuszczenie, że funkcja taka mogłaby być przydatna również w czasie normalnej pracy CSADz. Nie stwarza to jednak żadnych korzyści, ponieważ przydzielając cały trakt dla łączności wewnętrznej /a takie rozwiązanie przyjęto/ zmniejsza się o połowę liczbę kanałów przeznaczonych do współpracy z centralą. Nie jest także istotne, czy kanały przydzielone rozmawiającym ze sobą abonentom tego samego CSADz są komutowane w polu centrali czy CXL, gdyż w drugim przypadku liczba kanałów między CSADz i centralą nie zwiększy się.

Ze względu na brak doświadczeń i eksperymentów eksploatacyjnych koncentratora CSADz nie przedstawiono w artykule analizy niezawodnościowej zachowania się modelu w czasie awarii. Rozważania teoretyczne byłyby prawdopodobnie zbyt abstrakcyjne i niecelowe, bowiem rozwiązanie jest receptą na zaistniałą sporadycznie /kilką razy w ciągu 6 lat eksploatacji centrali E-10 w Poznaniu/ sytuację i niezawodność tej koncepcji może być określona i zbadana eksperymentalnie w zależności od natężenia ruchu telefonicznego i zachowania się abonentów w czasie awarii.

Następnym problemem jest podobne zamykanie ruchu lokalnego w grupie kilku CSAD obsługujących abonentów oddalonej od centrali miejscowości lub osiedla. Rozwiązania tego problemu już istnieją i przedstawione będą w innych opracowaniach.

Literatura

- [1] Błaszyk J. i inni: System komutacji elektronicznej E-10. WKŁ, Warszawa 1977
- [2] Dombek G., Żurek M.: Zamykanie ruchu lokalnego w koncentratorze odległym CSAD centrali systemu E-10. Przegląd Telekomunikacyjny nr 5, 1980, s.161-163
- [3] Réseau numérique intégré de télécommunications. Système CITEDIS. Wyd. CIT-Alcatel, SLE

FUNCTIONAL STRUCTURE OF THE PROCESS OF EMERGENCY CLOSING
OF INTERNAL TRAFFIC IN THE ELECTRONIC EXCHANGE E-10 CSADz CONCENTRATOR

Summary

The paper presents a method of providing connexion for subscribers who are joined to the remote concentrator of the electronic exchange E-10 in case of a damage which totally cuts off the cooperation between the exchange and the concentrator. In the time of damage, internal connexions are realized with the aid of an additional special ELS software. It steers the work of newly designed arrangements for the recognition of the state of the damage, of receipt, analysis, and counterchanging numbers. It also controls the temporal commutation local field and executes most of the normal concentrator functions.

ФУНКЦИОНАЛЬНАЯ СТРУКТУРА ПРОЦЕССА АВАРИЙНОЙ ОРГАНИЗАЦИИ СВЯЗИ ВНУТРИ
КОНЦЕНТРАТОРА CSAD₂ ЭЛЕКТРОНИЧЕСКОЙ СТАНЦИИ E-10

Резюме

В статье представлен метод обеспечения связи для абонентов подключенных к отдаленному концентратору электронной телефонной станции E-10, когда авария делает полностью невозможным сотрудничество между телефонной станцией и концентратором. Во время аварии внутренние соединения осуществляются благодаря дополнительному специальному программированию ELS-а. Это программирование управляет работой дополнительных заново запроектированных систем определения аварийного состояния устройств приема, анализа и вычисления добавочных номеров, а также управляет местным коммутатором. Это программирование обеспечивает также выполнение большинства нормальных действий концентратора.

Józef Pieprzyk

AUTENTYFIKACJA INFORMACJI W SYSTEMACH I SIECIACH KOMPUTEROWYCH

W pracy przedstawiono sposoby stwierdzania autentyczności wiadomości przesyłanych w systemach i sieciach komputerowych przy użyciu kryptograficznych metod ochrony informacji. Rozważania przeprowadzono zarówno w przypadku klasycznego systemu kryptograficznego, jak i systemu kryptograficznego z jawnymi kluczami. Na zakończenie podano stosowane w praktyce reguły autentyfikacyjne.

1. WSTĘP

W miarę upowszechniania techniki komputerowej rośnie zainteresowanie - zarówno użytkowników jak i projektantów systemów i sieci komputerowych - metodami ochrony informacji przed przypadkowym lub zamierzonym naruszeniem prywatności, bądź zniszczeniem informacji w nich przechowywanych lub przesyłanych przez nielegalnych użytkowników. Jedną z ważnych metod ochrony informacji przed tymi zagrożeniami jest kryptografia. Może być ona stosowana zarówno do zabezpieczenia wiadomości przesyłanych w niechronionych kanałach, jak też do utajniania treści informacji przechowywanych w pamięci systemu komputerowego. W obu przypadkach elementem składowym kryptograficznych metod ochrony jest autentyfikacja, która polega na dostarczeniu takich środków użytkownikom, aby mogli oni stwierdzić, czy otrzymane wiadomości pochodzą ze źródeł, do których są zaliczane. Problemy autentyfikacji pojawiają się szczególnie wyraźnie w momencie inicjowania seansu łączności, gdy od poprawności procesu autentyfikacji w zasadniczy sposób zależy jakość ochrony informacji przekazywanej pomiędzy dwoma komunikującymi się partnerami. Zatem można tu mówić o dwóch kategoriach autentyfikacji, a mianowicie autentyfikacji wiadomości oraz wzajemnej autentyfikacji partnerów. Autentyfikacja wiadomości polega na ustaleniu po obu stronach kanału zbioru wiadomości autentycznych, np. poprzez przesyłanie kryptogramów dla par /wiadomość, hasło/, przy czym hasło jest znane zarówno nadawcy jak i odbiorcy. Z kolei autentyfikacja partnera jest równoznaczna z ustaleniem, czy użytkownik jest tą osobą, za którą się podaje, a więc czy posiada on pewną wyróżnioną i charakterystyczną dla niego wiadomość. Wynika stąd, że w obu przypadkach autentyfikacja sprowadza się do autentyfikacji wiadomości, ponieważ partnerzy w celu wzajemnej autentyfikacji mogą przesyłać wyłącznie wiadomości.

W pracy tej przedstawimy zarys problemów, które trzeba rozwiązać przy poszukiwaniu reguł autentyfikacji niezbędnych do zapewnienia odpowiedniej ochrony informacji przesyłanych pomiędzy dwoma partnerami /końcówka - komputer, komputer - komputer, użytkownik - komputer, użytkownik - użytkownik, itp./. Zakładamy, że czytelnik posiada podstawowe wiadomości z teorii kryptografii, a w szczególności zna zasadę działania zarówno klasycznych systemów kryptograficznych, jak też systemów kryptograficznych z jawnymi kluczami [1, 4, 6].

2. PODSTAWOWE METODY AUTENTYFIKACJI

W celu sformalizowania rozważań wprowadzamy następujące oznaczenia:

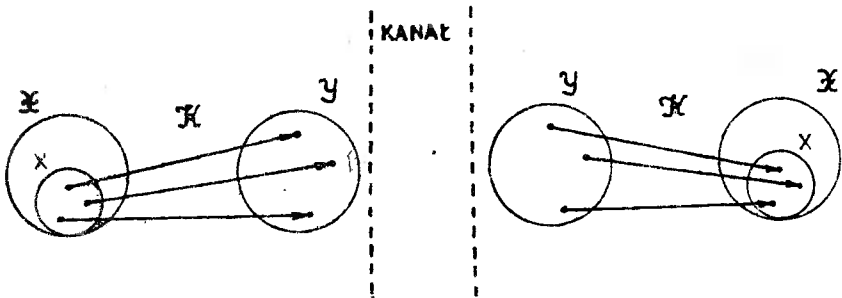
\mathfrak{X} - zbiór wszystkich wiadomości,

X - zbiór wiadomości, które będą uważane przez odbiorcę jako autentyczne / $X \subset \mathfrak{X}$ /,

\mathfrak{Y} - zbiór kryptogramów zaszyfrowanych wiadomości,

\mathfrak{K} - zbiór kluczy kryptograficznych.

Jeżeli teraz założymy, że utworzony został kanał umożliwiający przesyłanie wiadomości pomiędzy dwoma użytkownikami oraz do ich ochrony stosowany jest klasyczny system kryptograficzny /KSK/, to odbiorca zdając sobie sprawę z możliwości nielegalnej ingerencji w kanale będzie stosował następującą strategię odbioru wiadomości /rys.1/.



Rys.1. Graficzne przedstawienie zasady autentyfikacji

W przypadku, gdy $y = f(x, k)$, gdzie $y \in \mathfrak{Y}$, $x \in \mathfrak{X}$, $k \in \mathfrak{K}$ oraz $f(\cdot, *)$ jest funkcją kryptograficzną szyfratora dla KSK, oraz gdy wiadomość wyznaczona z kryptogramu należy do zbioru X , to odbiorca będzie tę wiadomość uważał za autentyczną, w przeciwnym razie $x \in \mathfrak{X} \setminus X$ wiadomość zostanie odrzucona jako podstawiona przez nielegalnego użytkownika /NU/.

W najprostszym przypadku autentyfikacja polega na ustaleniu przed transmisją pewnej pomocniczej informacji zwanej hasłem^{1/} oraz przesyłaniu

^{1/} Hasło jest znane wyłącznie nadawcy i odbiorcy.

każdej wiadomości x za pomocą kryptogramu w postaci

$$y = f [x, h, k]$$

przy czym h jest hasłem oraz k jest znanym przez nadawcę i odbiorcę kluczem kryptograficznym. Taki sposób autentyfikacji został zastosowany w systemie LUCIFER opracowanym przez IBM do ochrony informacji przekazywanej pomiędzy końcówką a centrum obliczeniowym [2]. W odbiorniku po deszyfracji kryptogramu sprawdza się, czy otrzymane hasło jest zgodne z hasłem przechowywanym w odbiorniku i w przypadku zgodności wiadomość x traktuje się jako autentyczną. Opisywany sposób autentyfikacji ma jednak zasadniczą wadę wynikającą z konieczności znajomości przez obu partnerów jawnej postaci hasła i w przypadku nieostrożności jednego z nich może być ono nielegalnie rozpowszechniane.

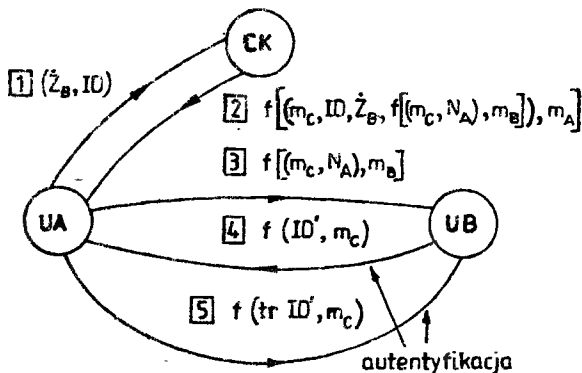
W innym rozwiązaniu autentyfikacji wspomniana wyżej nieautoryzowana propagacja hasła przechowywanego w odbiorniku jest dozwolona, bowiem nadawca dostatecznie wcześniej wytwarza na podstawie wybranego przez siebie hasła jego obraz uzyskany za pomocą tzw. funkcji jednokierunkowej^{1/} /ang. one-way function/. Realizacja tej idei w oparciu o system kryptograficzny z jawnymi kluczami polega na wytworzeniu przez nadawcę i dostarczeniu na stronę odbiorczą pary /klucz jawny, kryptogram/, przy czym nadawca pozostawia sobie jako swój autentyfikator /hasło/ klucz tajny. Zatem odbiorca może stwierdzić autentyczność nadawcy po otrzymaniu klucza tajnego. Jeśli para /wiadomość, klucz tajny/^{2/} umożliwia wytworzenie przechowywanego po stronie odbiorczej kryptogramu, to nadawca jest autentyczny. Z drugiej strony jeśli do ochrony kryptograficznej są stosowane RSK, to w odbiorniku przechowywana jest para /wiadomość, kryptogram/^{3/}, natomiast nadawca jako hasła używa klucza kryptograficznego, który został użyty do wytworzenia tej pary. Zatem odbiorca może stwierdzić autentyczność nadawcy po otrzymaniu klucza przez porównanie, czy on "pasuje" do przechowywanej w odbiorniku pary /wiadomość, kryptogram/. Cechą charakterystyczną takiego postępowania, którą można uznać za jego wadę, jest fakt, że każde hasło /klucz kryptograficzny/ może być używane w celu autentyfikacji tylko raz. Należy dodać, że tak określoną regułą autentyfikacji stosuje się w ogólności : na końcu seansu łączności i można wtedy hasło /klucz kryptograficzny/ przesyłane przez nadawcę traktować jako jego podpis, który umożliwia odbiorcy stwierdzenie autorstwa otrzymanej informacji [5, 7].

-
- 1/ Funkcja taka charakteryzuje się tym, że otrzymanie obrazu hasła wymaga małej liczby elementarnych operacji przetwarzania, natomiast wyznaczenie hasła na podstawie jego obrazu wymaga dużej liczby elementarnych operacji przetwarzania [3].
 - 2/ Odbiorca znając /klucz jawny, kryptogram/ może w prosty sposób wyznaczyć jawną postać wiadomości, natomiast nie jest w stanie w rozsądnym przedziale czasu określić klucza tajnego [3].
 - 3/ Para ta jest wyznaczona przez nadawcę i odpowiednio wcześniej przesyłana do odbiorcy, przy czym zakładamy, że algorytm szyfrowania nie pozwala na wyznaczenie klucza na podstawie tej pary w rozsądnym przedziale czasu.

W celu autentyfikacji wiadomości podczas wielu seansów łączności niezbędne jest ustalenie odpowiednio wcześniej kilku, bądź kilkunastu różnych haseł. Problem takiego rozdzielenia haseł i odpowiadających im par /wiadomość, kryptogram/ dla KSK lub /klucz jawny, kryptogram/ dla systemu kryptograficznego z jawnymi kluczami staje się uciążliwy, szczególnie wtedy, gdy mamy do czynienia z siecią teleinformatyczną. Wyjściem z tej sytuacji jest stworzenie centrum kontrolnego /CK/, które będzie nadzorowało prawidłowość procesu autentyfikacji i w przypadkach wątpliwych odgrywać będzie rolę arbitra [5].

3. PROCES AUTENTYFIKACJI W PRZYPADKU WYKORZYSTYWANIA KSK

Przyjmujemy, że mamy do czynienia z siecią teleinformatyczną z jednym centrum kontrolnym^{1/}, które będzie nadzorowało prawidłowość procesów zarówno autentyfikacji jak i szyfrowania wiadomości. Zakładamy dalej, że jest dwóch użytkowników /końcówki, komputery, itp./ A i B, z których użytkownik A /UA/ żąda przydzielenia kanału, który umożliwi mu konwersację z użytkownikiem B /UB/ - patrz rys.2.



Rys.2. Kolejność przesyłania informacji w przypadku stosowania KSK pomiędzy centrum a użytkownikami w fazie inicjacji seansu łączności /kolejność zaznaczono przez numerowanie poszczególnych informacji/

UA przesyła to żądanie Z_B do centrum kontrolnego /CK/ w jawny sposób wraz ze swoim identyfikatorem /ID/. Centrum na podstawie otrzymanego żądania Z_B nawiązuje łączność z użytkownikiem B oraz identyfikatora użytkownika A wyszukuje w spisie użytkowników klucze kryptograficzne m_A , m_B służące do ochrony informacji przesyłanych pomiędzy odpowiednio UA ↔ CK oraz UB ↔ CK. Z kolei CK wytwarza klucz m_C , który będzie służył do ochrony wiadomości przesyłanych pomiędzy użytkownikami A i B. Na podstawie tych

^{1/} Sieć z organizacją scentralizowaną [5].

wiadomości centrum tworzy kryptogram w postaci

$$f \left[(m_C, ID, Z_B, f((m_C, N_A), m_B)), m_A \right]$$

gdzie N_A jest nazwą użytkownika A oraz $f/\cdot, \cdot/$ jest funkcją kryptograficzną KSK stosowanego w sieci i przesyła go do UA. Użytkownik A znając swój klucz m_A wydobywa jawną postać wiadomości, czyli określa ciąg

$$[m_C, ID, Z_B, f((m_C, N_A), m_B)]$$

oraz sprawdza czy ciąg $/ID, Z_B/$ jest zgodny z oryginałem, a ciąg $f[(m_C, N_A), m_B]$ przesyła do UB, który^{1/} wydobywa jawną postać ciągu $/m_C, N_A/$. Na podstawie nazwy N_A użytkownik B dowiadyuje się, który z użytkowników nawiązuje z nim łączność, natomiast klucz m_C jest przez niego wykorzystywany do szyfrowania identyfikatora ID' , którego postać użytkownik wyznacza w sposób losowy. Użytkownik A po otrzymaniu kryptogramu $f/ID', m_C/$ wyznacza ID' oraz przekształca ją zgodnie z pewną funkcją $tr/\cdot, \cdot/$, która jest znana przez obu użytkowników. W rezultacie tych przekształceń do użytkownika B przesyłany jest kryptogram

$$f [tr(ID'), m_C]$$

Dwa ostatnie kryptogramy przekazywane pomiędzy użytkownikami są używane w celu wzajemnej autentyfikacji. Natomiast dalsza autentyfikacja, tym razem wiadomości, jest przeprowadzana za pomocą:

- hasła,
- dwukrotnej transmisji wiadomości /tworzenie "pociągu" [2] /,
- systemu kryptograficznego ze sprzężeniem zwrotnym wiadomości [1].

Dla zilustrowania tych metod autentyfikacji wiadomości zakładamy, że użytkownik po pozytywnej autentyfikacji będzie przysyłał ciąg kolejnych wiadomości $/x_1, \dots, x_i, \dots/$ za pomocą klucza m_C . W przypadku metody hasła ciąg kryptogramów pojawiających się kolejno w kanale będzie miał postać

$$y_i = f [x_i, h, m_C] \quad ; \quad i = 1, \dots$$

gdzie h jest ciągiem reprezentującym hasło, które jest znane po obu stronach kanału.

Z kolei metoda tworzenia pociągu /ang. train/ polega na przysyłaniu kolejno kryptogramów w formie

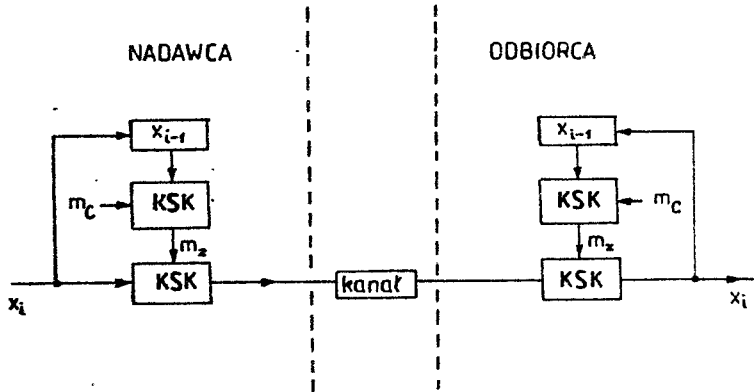
$$y_{2i-1} = f [x'_i, x''_i, m_C] = f [x_i, m_C]$$

$$y_{2i} = f [x''_i, x'_{i+1}, m_C] \quad ; \quad i = 1, 2, \dots$$

^{1/} UB dokonuje tego na podstawie znanego mu klucza m_B .

Łatwo zauważyć, że każdy kolejny kryptogram zawiera powtórzenie już przesłanej połowy ciągu reprezentującego wiadomość, co powoduje dwukrotne wydłużenie ciągu kryptogramów w stosunku do ciągu wiadomości /każda wiadomość jest przesyłana dwukrotnie/. Autentyfikacja w tym rozwiązaniu polega na sprawdzeniu powtórzonych segmentów przesyłanych wiadomości. Jeśli powtórzona część wiadomości nie jest zgodna z oryginałem, to odbiorca kończy odbiór wiadomości i sygnalizuje naruszenie ochrony.

Podobną koncepcję do metody tworzenia pociągu jest stosowanie systemów kryptograficznych ze sprzężeniami zwrotnymi wiadomości /rys.3/.



Rys.3. Autentyfikacja wiadomości w oparciu o system kryptograficzny ze sprzężeniem zwrotnym wiadomości

Na rysunku tym przedstawiono schemat systemu z zastosowaniem dwóch KSK. Łatwo zauważyć, że kolejne kryptogramy są uzależnione od poprzednio nadanych kryptogramów w następujący sposób:

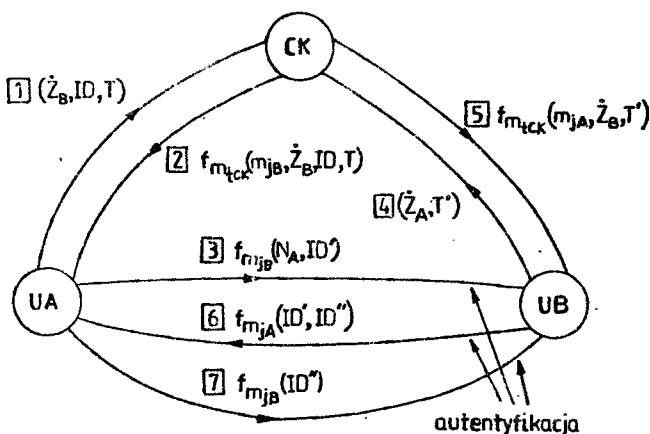
$$y_i = f(x_i, m_c) = f[x_i, f(x_{i-1}, m_c)] ; \quad i = 1, 2, \dots$$

A więc każdy kolejny kryptogram zależy od aktualnie przesyłanej wiadomości x_i oraz od przesłanej już wiadomości x_{i-1} . Zatem NU chcąc podstawić fałszywą wiadomość musi wyznaczyć co najmniej dwie kolejne wiadomości, co z uwagi na budowę systemu kryptograficznego z rys.3 jest praktycznie niemożliwe [4]. Jak można zauważyć, autentyfikacja w tym przypadku polega na używaniu w procesie kryptograficznym /szyfrowanie i deszyfrowanie/ zmiennej wartości klucza m_c . W przypadku gdy kryptogram jest kryptogramem fałszywym otrzymane wiadomości nie mają sensu semantycznego.

4. PROCES AUTENTYFIKACJI Z ZASTOSOWANIEM SYSTEMÓW KRYPTOGRFICZNYCH Z JAWNYMI KLUCZAMI

W sieci teleinformatycznej, w której ochrona informacji została wprowadzona w oparciu o system kryptograficzny z jawnymi kluczami, również ist-

nieje centrum, którego zadaniem jest kontrola, rozdział i generacja kluczy kryptograficznych. Podobnie jak poprzednio zakładamy, że dwóch użytkowników A i B żąda przydzielenia kanału dla wymiany informacji. Użytkownik A, który inicjuje połączenie, przesyła do centrum kontrolnego jako pierwszą wiadomość odpowiednie żądanie \hat{Z}_B wraz ze swoim identyfikatorem ID i aktualnym czasem T /patrz rys.4/.



Rys.4. Kolejność przesyłania informacji w przypadku stosowania systemów kryptograficznych z jawnymi kluczami podczas inicjacji seansu łączności /kolejność zaznaczono przez numerowanie poszczególnych informacji/

Centrum zwrótnie przesyła kryptogram wyznaczony za pomocą klucza tajnego m_{tCK} dla ciągu wiadomości $/m_{jB}, \hat{Z}_B, ID, T/$, przy czym m_{jB} jest jawnym kluczem użytkownika B, natomiast klucz m_{tCK} jest znany wyłącznie w centrum. Odbiorca /UA/ znając jawny klucz centrum m_{jCK} odczytuje jawną wartość klucza m_{jB} oraz sprawdza pozostałe informacje. W celu przeprowadzenia procesu wzajemnej autentyfikacji UA przesyła do użytkownika B swoją nazwę N_A i identyfikator ID' w postaci kryptogramu

$$f_{m_{jB}} /N_A, ID'/$$

Odbiorca /UB/ wydobywa jawną postać wiadomości za pomocą swojego klucza tajnego. Znając nazwę użytkownika A formułuje on żądanie \hat{Z}_B dostarczenia przez centrum jawnego klucza użytkownika A i razem z czasem T' przesyła obie te informacje do CK w jawnej postaci. Centrum przekazuje do UB kryptogram w postaci

$$f_{m_{tCK}} /m_{jA}, \hat{Z}_B, T'/$$

gdzie m_{jA} jest wartością jawnego klucza użytkownika A. Z kolei użytkownik B przesyła wiadomość $/ID', ID''/$ zaszyfrowaną kluczem jawnym użytkownika

ka A. W odpowiedzi UA przekazuje kryptogram $f_{m,B}/ID''$. Zaznaczone strzałkami kolejne etapy wymiany informacji są realizowane w celu wzajemnej autentyfikacji, a przedstawiona procedura inicjowania seansu łączności musi poprzedzać z oczywistych względów każdy seans łączności. Natomiast transmisja kolejnych wiadomości będzie się odbywać jak w punkcie 3.

5. WNIOSKI

Przedstawione metody autentyfikacji są jak gdyby "wtopione" w całość ochrony kryptograficznej. Mają one do spełnienia szczególnie ważne zadania, a mianowicie - ochronę przed przypadkową lub zamierzoną wymianą przesyłanych informacji. W granicznym przypadku poprawnie zaprojektowane metody autentyfikacji powinny umożliwiać wykrywanie podstawienia kryptogramów, które już wcześniej były przesyłane.

Należy zwrócić uwagę, że proces autentyfikacji stosowany podczas seansu łączności składa się z dwóch części. W pierwszej z nich dokonywana jest autentyfikacja użytkownika, natomiast w drugiej części - autentyfikacja przesyłanych ciągów wiadomości. Autentyfikacja użytkowników przeprowadzana jest poprzez autentyfikację wiadomości, przy czym zbiór wiadomości traktowanych jako autentyczne jest zbiorem jednoelementowym. W przypadku autentyfikacji wiadomości tak nie jest. Aby zmniejszyć prawdopodobieństwo sukcesu NU w podstawianiu fałszywych kryptogramów dąży się do ograniczenia liczności zbioru autentycznych wiadomości, co jednak zwiększa redundancję ciągów kryptogramów /np. podczas stosowania metody pociągu stosunek rozwlekłości kryptogramów do wiadomości równa się dwa/.

Z autentyfikacją ściśle wiążą się nowe usługi oferowane w sieciach teleinformatycznych, a mianowicie poczta elektroniczna i przesyłanie podpisów. Ze względu na specyfikę problemów, które pojawiają się podczas wdrażania takich usług, zostały one w pracy pominięte, a zainteresowanych odsyłamy do literatury [5,6].

L i t e r a t u r a

- [1] Diffie W., Hellman M.E.: An Introduction to Cryptography. Proceedings of the IEEE, vol.67, No 3, March 1979, s.397-427
- [2] Feistel H.: Cryptography and Computer Privacy. Scientific Amer., vol. 228, No 5, May 1973, s.15-23
- [3] Lempel A.: Cryptology in Transition. Computing Surveys, vol.11, No 4, December 1979, s.285-303
- [4] Pieprzyk J., Rutkowski D.: Kryptograficzne metody zabezpieczenia informacji. Rozprawy Elektrotechniczne, z.1, 1982, s.285-314

- [5] Popek G.J., Kline Ch.S.: Encryption and Secure Computer Networks. Computing Surveys, vol.11, No 4, December 1979, s.331-356
- [6] Simmons G.J.: Symmetric and Asymmetric Encryption. Computing Surveys, vol.11, No 4, December 1979, s.304-330
- [7] Simmons G.J.: Secure in the Presence of Pervasive Deceit. Proceedings of the 1980 Symposium on Security and Privacy, Oakland, California, April 1980, s.84-93

AUTHENTICATION OF INFORMATION IN COMPUTER SYSTEMS AND NETWORKS

Summary

The paper presents ways of stating the authenticity of the information sent through computer systems and networks by the use of cryptographic methods for information protection. There are considered both classical and public key cryptosystems. The paper includes authentication rules used in practice.

ПРОВЕРКА ИНФОРМАЦИИ В СИСТЕМАХ И СЕТЯХ ЭВМ

Резюме

В работе представлены методы проверки происхождения информации передаваемых в вычислительных системах при использовании криптографических методов защиты информации. Рассмотрены как симметрические криптографические системы, так и криптографические системы с открытыми ключами. В заключение работы представлены практические правила проверки происхождения.

Andrzej Błaszczak
Ireneusz Żmizkiński

OCENA KOMUTACJI ZINTEGROWANEJ W MODELU SIECI TELEKOMUNIKACYJNEJ

W pracy przedstawiony został model sieci telekomunikacyjnej, uwzględniający integrację metody komutacji kanałów i metody komutacji informacji. Dla podanego modelu przeprowadzono ocenę jakości sieci, przy czym jako wskaźnik jakości przyjęto średnie opóźnienie, jakiemu podlega informacja przesyłana przez sieć. Rozważania zostały zilustrowane przykładem obliczeniowym.

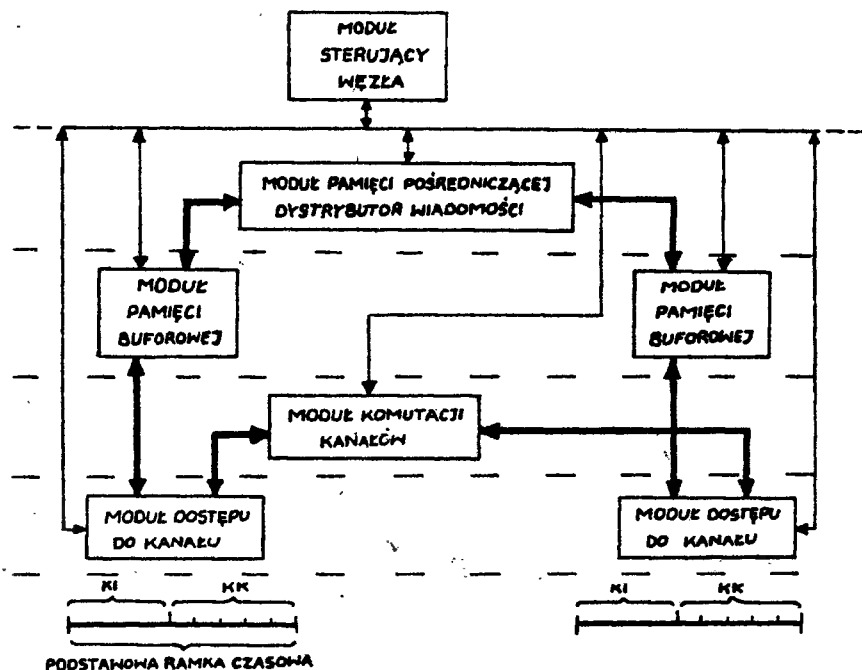
1. WSTĘP I ZAŁOŻENIA

Badania nad usprawnieniem przesyłania informacji mowy i informacji komputerowych przez sieć telekomunikacyjną inspirowane były dążeniem do wykorzystania zalet odmiennych metod komutacji w tej samej sieci dla szerokiego zakresu zmienności parametrów, charakteryzujących strumienie informacji pochodzące od różnych użytkowników sieci. W konwencjonalnych sieciach istnieje wyraźny podział na sieci telefoniczne, w których stosowana jest metoda komutacji kanałów /KK/ oraz sieci komputerowe, w których stosowana jest metoda komutacji informacji /KI/. Poważne zmiany konstrukcyjne i technologiczne, jakie wystąpiły w ostatnich latach w budowie urządzeń spotykanych w sieciach telefonicznych [3, 9], pozwalają na rezygnację z takiego tradycyjnego przyporządkowania metod komutacji. Oznacza to między innymi możliwość zastosowania metody KI do przesyłania informacji mowy oraz metody KK do przesyłania informacji komputerowych. Wyniki takich badań zawierają prace [1, 2, 4, 6, 7, 8, 9], w których podjęto próby oceny sieci z odmiennymi metodami komutacji.

W tej sytuacji wydaje się celowe podjęcie próby podania takiego modelu sieci, który uwzględniać będzie integrację metod komutacji, a ponadto pozwoli na analizę sieci z punktu widzenia oddziaływania jej parametrów na wybrany wskaźnik jakości. Podobnie jak ma to miejsce w pracach [1, 4, 6, 7, 9] za taki wskaźnik przyjmujemy średni ważony czas przesyłania informacji przez sieć, który będziemy nazywali opóźnieniem sieci. Zdefiniujemy je jako przedział czasu zawarty pomiędzy chwilą wprowadzenia do węzła źródłowego /WZ/ żądania połączenia /metoda KK/ lub pierwszego bitu informacji /metoda KI/, a chwilą dostarczenia ostatniego bitu przesyłanej informacji

do węzła docelowego /WD/.

Przyjmujemy, że węzły rozważanej sieci mają proponowaną w pracach [3,8] strukturę modułową /rys.1/ i połączone są kanałami zwielokrotnionymi czasowo, a czas ich użytkowania jest przydzielany /zmiennie lub sztywno/ dla potrzeb transmisji według każdej z metod komutacji. Oczywiście zarówno informacje mowy jak i informacje komputerowe przesyłane są w postaci cyfrowej.



Rys.1. Przykładowa struktura węzła modelowanej sieci

Każdy węzeł sieci może realizować metodę KK lub KI zależnie od tego, która z nich jest korzystniejsza w określonych warunkach dla sprawnego przesyłania informacji. Wybór metody komutacji może w szczególności zależeć od kategorii użytkownika sieci, jak np. w sieci PACUIT [2]. W przypadku metody KK następuje przydzielenie wybranych podkanałów czasowych określonej parze użytkowników na czas trwania seansu łączności^{1/}, a niezbędne dla doboru, zestawiania i rozłączania trasy informacje sterująco-kontrolne^{2/} przesyłane są z wykorzystaniem metody KI. Realizacja metody KI odbywa

^{1/} Seansem łączności będziemy nazywali przedział czasu, w którym zestawiona trasa pozostaje do wyłącznej dyspozycji pary komunikujących się użytkowników sieci.

^{2/} Zwana wiadomościami sygnalizacyjnymi; rozróżnia się na ogół trzy rodzaje wiadomości sygnalizacyjnych, a mianowicie: doboru trasy /ang. inquiry message/, potwierdzenia zestawienia trasy /ang. response message/ i rozłączenia trasy /ang. disconnect message/.

się według zasady "zapamiętaj przejściowo wiadomość a następnie prześlij ją dalej" /ang. "store-and-forward"/, a więc obejmuje między innymi buforowanie, dobór trasy i potwierdzanie poprawnego odbioru wiadomości^{1/}.

Dokonyjąc oceny jakości sieci zależy nam na tym, by z jednej strony model sieci możliwie wiernie odzwierciedlał jej własności a z drugiej, by prowadził do wyników w możliwie prostej i zwartej postaci. Dla potrzeb tak rozumianej analizy przyjmujemy następujące założenia o charakterze upraszczającym:

1. Strumienie wiadomości właściwych i sterująco-kontrolnych opisane są modelem Poissona - wykładniczym, zdefiniowanym w pracy [10] str.150.
2. Spełniona jest hipoteza Kleinrocka, tzn. strumień wiadomości po przebyciu dowolnej liczby węzłów i kanałów podlega rozkładowi Poissona - wykładniczemu.
3. Kanały sieci są niezawodne, nie wnoszą zakłóceń do przesyłanej informacji, a czas propagacji przenoszących je sygnałów jest pomijalnie mały.
4. Urządzenia komutujące w węzłach realizują każde żądanie zestawienia trasy, a czas łączenia pary/podkanałów jest pomijalnie mały.
5. Pojemność pamięci buforowych węzłów jest nieskończenie duża, a czas dostępu do pamięci jest pomijalnie mały.
6. Badana sieć jest siecią zrównoważoną, w której każdy z węzłów i każdy z kanałów sieci jest jednakowo obciążony, a więc współczynniki wykorzystania wszystkich kanałów są jednakowe.

Założenia tego rodzaju są stosowane w większości prac dotyczących oceny jakości i analizy porównawczej sieci z odmiennymi metodami komutacji /patrz np. [2, 4, 5, 7, 9] /.

2. MODEL BADANEJ SIECI

Przyjmujemy, że wszystkie wiadomości w sieci podzielone są na dwie następujące kategorie:

- kategoria A, zgodnie z którą wiadomość wprowadzana do sieci jest obsługiwana za pomocą metody KK. W pierwszej kolejności zostaje wysłana z WZ wiadomość sygnalizacyjna reprezentująca żądanie doboru trasy /ZDT/; po jej dotarciu do WD zostaje w kierunku przeciwnym przesłana wiadomość sygnalizacyjna mająca sens żądania zestawienia trasy /ZZT/ w postaci łańcucha podkanałów łączących węzły uprzednio wybrane przez wiadomość ZDT^{2/}. Odebranie wiadomości ZZT w WZ powoduje rozpoczęcie seansu łączności przez użytkowników sieci, po zakończeniu którego zostaje przesłana wiadomość będąca żądaniem rozłączenia trasy /ZRT/, ma-

^{1/} Przez wiadomość rozumiemy informację tworzącą semantyczną całość z punktu widzenia użytkownika sieci.

^{2/} Jest to tzw. regresywne zestawianie trasy /ang. backward path set-up/.

jąca sens przekazania w każdym węźle podkanałów wykorzystanych w zestawianej trasie do puli podkanałów niezajętych,

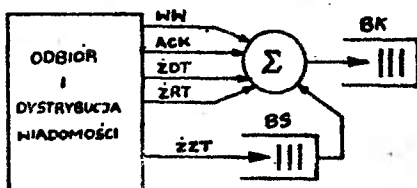
- kategoria B, zgodnie z którą wiadomość wprowadzana do sieci jest obsługiwana za pomocą metody KI. Każda wiadomość wprowadzona do WZ zostaje uzupełniona informacją sterująco-kontrolną /nagłówkiem/, a następnie przesłana zgodnie z obowiązującą regułą doboru trasy do węzła sąsiedniego.

Odbiór wiadomości bez zakłóceń jest potwierdzany za pomocą wiadomości kontrolnej PPO /potwierdzenie poprawnego odbioru, ang.acknowledgement/.

Przyjmując hipotezę Kleinrocka /założenie 2/, interesujące nas opóźnienie sieci T może być wyrażone za pomocą średniego ważonego czasu przechodzenia wiadomości przez poszczególne kanały i węzły [10]. Uwzględniając dodatkowo założenie 6, dotyczące jednakowych współczynników wykorzystania wszystkich kanałów w sieci, będziemy jako jej opóźnienie przyjmowali średni czas przechodzenia informacji wzdłuż trasy o średniej długości [11].

Jako uproszczony model węzła /rys.2/ przyjęliśmy system masowej obsługi, który będziemy analizowali przez dekompozycję na dwa następujące modele:

- M/M/S/∞ [5], reprezentujący obsługę wiadomości z wykorzystaniem metody KK,
- M/G/1/∞ z priorytetem względnym [5], reprezentujący obsługę wiadomości z wykorzystaniem metody KI.



Rys.2. Uproszczony model węzła:

BK - pamięć buforowa dla obsługi metodą KI

BS - pamięć buforowa dla obsługi wiadomości ZDT

W celu podania zależności umożliwiających obliczenie opóźnienia dla poszczególnych metod komutacji wprowadzimy następujące oznaczenia:

\bar{v}_s - średnia długość seansu łączności /w bitach/,

\bar{v}_w - średnia długość wiadomości właściwych /w bitach/,

C_1 - przepustowość pojedynczego podkanału /w bitach na sekundę/,

S - liczba podkanałów przeznaczonych do obsługi wiadomości metodą KK,

K - liczba podkanałów przeznaczonych do obsługi wiadomości metodą KI,

λ - średnie natężenie strumienia wiadomości wprowadzanych do WZ,

\bar{v}_n - średnia długość wiadomości sterująco-kontrolnych /w bitach/,

α - współczynnik charakteryzujący podział wiadomości w sieci na kategorie A i kategorie B,

$$\rho_2 = \frac{\alpha \lambda \bar{v}_s}{SC_1} \quad - \text{współczynnik wykorzystania kanału międzywęzłowego dla metody KK,}$$

$$\rho_2 = \frac{(\bar{v}_w + \bar{v}_n) \lambda (1 - \alpha)}{KC_1} \quad - \text{współczynnik wykorzystania kanału międzywęzłowego dla metody KI,}$$

$$\rho_3 = \frac{(1 + 2\alpha) \lambda \bar{v}_n}{KC_1} \quad - \text{współczynnik wykorzystania kanału międzywęzłowego przez wiadomości sterująco-kontrolne,}$$

$Er_2(S, \rho_1)$ - druga funkcja Erlanga, reprezentująca prawdopodobieństwo czekania w kolejce przez wiadomość ŻT wskutek zajętości wszystkich podkanałów przeznaczonych do realizacji metody KK.

Wprowadzone założenia oraz przyjęte modele matematyczne systemów masowej obsługi pozwalają po niezbędnych przekształceniach podstawowych zależności [1, 5, 11] na otrzymanie dla metody KK wzoru umożliwiającego obliczenie średniego opóźnienia dla sieci KK

$$\bar{T}_{KK} = \left[\frac{Er_2(S, \rho_1) \bar{v}_s}{SC_1(1 - \rho_1)} + \lambda \frac{(1 - \alpha) [E^2(\bar{v}_w) + E^2(\bar{v}_n) + 2\bar{v}_w \bar{v}_n] + (1 + 2\alpha) E^2(\bar{v}_n)}{(KC_1)^2 (1 - \rho_3)} \right] \bar{L} + \frac{\bar{v}_s}{C_1} \quad (1)$$

gdzie $E^2(\bar{v}_w)$ i $E^2(\bar{v}_n)$ są odpowiednio drugimi momentami rozkładu prawdopodobieństwa wiadomości właściwych i sterująco-kontrolnych. Pierwszy składnik sumy w nawiasie kwadratowym wzoru (1) reprezentuje tu średni czas czekania wskutek zajętości wszystkich podkanałów, natomiast drugi jest średnim czasem zestawiania trasy w jednym węźle.

Podobnie, dla metody KI otrzymamy średnie opóźnienie sieci \bar{T}_{KI} jako sumę czasu buforowania i przesyłania wiadomości przez trasę liczącą \bar{L} kanałów

$$\bar{T}_{KI} = \left[\lambda \frac{(1 - \alpha) [E^2(\bar{v}_w) + E^2(\bar{v}_n) + 2\bar{v}_w \bar{v}_n] + (1 + 2\alpha) E^2(\bar{v}_n)}{2(KC_1)^2 (1 - \rho_3) (1 - \rho_2 - \rho_3)} + \frac{\bar{v}_w + \bar{v}_n}{KC_1} \right] \bar{L} \quad (2)$$

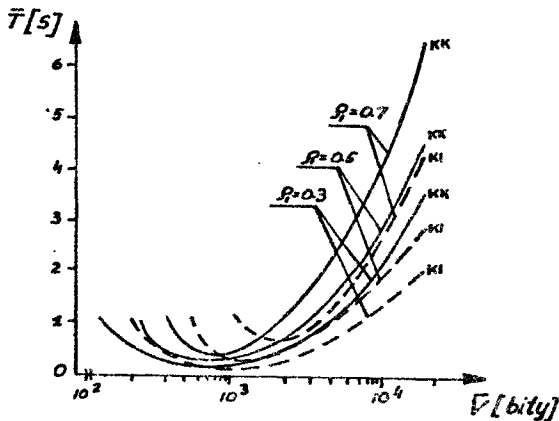
3. PRZYKŁAD

W celu ustalenia oddziaływania parametrów sieci na wielkość opóźnienia dla omawianego modelu sieci przyjmujemy:

1. Kanały międzywęzłowe o przepustowości 50 kb/s podzielone są na $S = 5$ podkanałów o przepustowości $C_1 = 5$ kb/s dla metody KK oraz podkanał o przepustowości $KC_1 = 25$ kb/s dla metody KI.
2. Wszystkie wiadomości sterująco-kontrolne posiadają stałą długość $V_n = 100$ bitów.

Obliczenia przeprowadzono korzystając z zależności (1) i (2) dla trzech wartości $\rho_1 = 0,3; 0,5; 0,7$ oraz trzech różnych wariantów podziału strumienia wiadomości w WZ, tzn. $\alpha = 0,3; 0,5; 0,7$ i średniej długości trasy $\bar{L} = 3$ jako parametrów sieci. Opóźnienie dla poszczególnych parametrów sieci zostało wyznaczone jako funkcja długości wiadomości właściwej przy założeniu, że w trakcie seansu łączności przesyłana jest jedna wiadomość $\bar{V}_s = \bar{V}_w = \bar{V}$. Dodatkowo przeprowadzono dla wybranego \bar{V} badania w zakresie wpływu średniej długości trasy na średnie opóźnienie sieci.

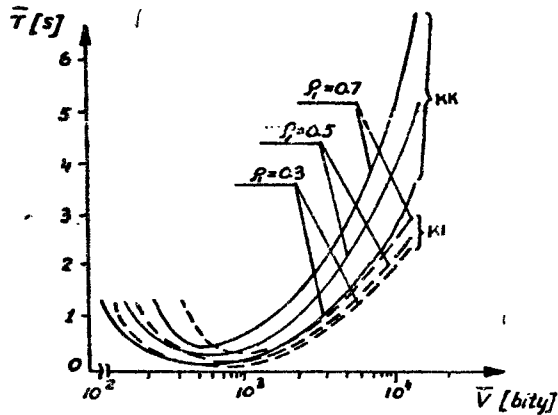
Rys.3 przedstawia zależność średniego opóźnienia dla obu metod komutacji w funkcji V dla trzech wartości współczynnika ρ_1 przy założeniu $\alpha = 0,5$.



Rys.3. Średnie opóźnienie sieci w funkcji długości wiadomości dla $\bar{L} = 3$ i $\alpha = 0,5$

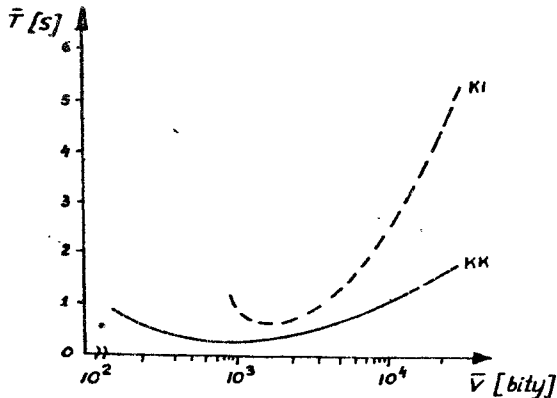
Mając na uwadze fakt, że \bar{T}_{KK} i \bar{T}_{KI} jest sumą średnich czasów oczekiwania wiadomości na przesłanie oraz jej przesyłania w sieci, to dla przyjętego modelu sieci minimalną wartość średniego opóźnienia \bar{T} otrzymujemy dla wiadomości o średniej długości około 1 kb. Przyrost \bar{T} dla wiadomości o mniejszej długości wynika z tego, że przy założeniu stałej długości i wyższego priorytetu wiadomości sterująco-kontrolnych, średni czas oczekiwania wiadomości właściwych na obsługę staje się dominującą składową \bar{T} dla obu rozważanych metod komutacji. Dla wiadomości właściwych o długości $\bar{V} > 10$ kb przyrost wartości \bar{T} wynika bezpośrednio ze wzrostu czasu przesyłania tych wiadomości i w konsekwencji prowadzi do zwiększenia czasu zestawiania trasy dla metody KK /wiadomości sygnalizacyjne posiadają priorytet względny/.

Zwiększenie wartości współczynnika /rys.4/ nie zmienia zasadniczo kształtu otrzymywanych charakterystyk prowadząc do przyrostu \bar{T}_{KK} w wyniku wydłużenia średniego czasu zestawiania trasy.



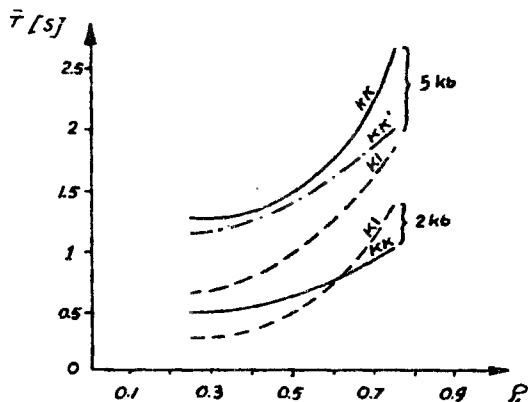
Rys.4. Zależność średniego opóźnienia sieci od długości wiadomości przy $\bar{L} = 3$ i $\alpha = 0,7$

Dla metody KI daje się zauważyć jedynie niewielką fluktuację wartości średniego opóźnienia dla różnych wartości ρ_1 , co jest efektem zmniejszania się wartości współczynników ρ_2 i ρ_3 , które występują we wzorze (2). Zmniejszenie współczynnika α /rys.5/ prowadzi do przyrostu \bar{T}_{KI} w wyniku zwiększania ρ_2 przy założonym ρ_1 . Z tego też względu dla zmniejszenia \bar{T}_{KI} konieczne jest spełnienie warunku $\rho_2 + \rho_3 < 1$, który wynika z założenia o stacjonarności kolejki w pamięci buforowej. Ponadto dla tego przypadku występuje w szerokich granicach zmienności \bar{V} mała wartość \bar{T}_{KK} .



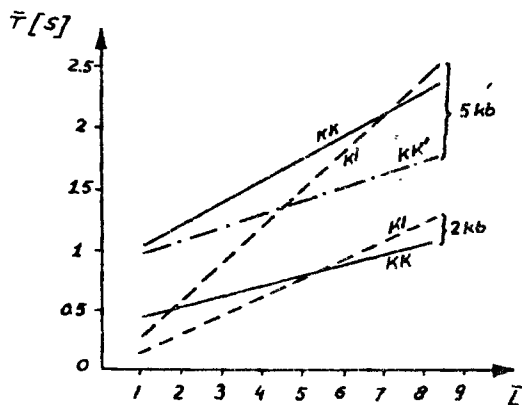
Rys.5. Średnie opóźnienie sieci w funkcji długości wiadomości dla $\bar{L} = 3$ i $\alpha = \rho_1 = 0,3$

Interesującą zależność \bar{T} w funkcji ρ_1 , przedstawia rys.6, z którego wynika, że dla krótkich wiadomości $|\bar{V}| < 2 \text{ kb}$ opóźnienie dla metody KK jest mniejsze niż dla metody KI.



Rys.6. Zależność średniego opóźnienia sieci od ρ_1 dla $\bar{L}=3$ i $\alpha=0,5$

Wynika z tego konieczność usprawnienia przesyłania ZZT, a tym samym zmniejszenia \bar{T}_{KK} dla długich wiadomości $|\bar{V}| > 2 \text{ kb}$, co reprezentuje krzywa KK' na rys.6. Wprowadzono tu modyfikację polegającą na tym, że wiadomości długie przesyłane są za pomocą metody KK, natomiast wiadomości krótkie za pomocą metody KI. Korzyści takiej modyfikacji wyraźnie ilustruje rys.7, przedstawiający zależność \bar{T} w funkcji średniej długości trasy \bar{L} .



Rys.7. Zmiany średniego opóźnienia sieci w funkcji średniej długości trasy dla $\alpha = \rho_1 = 0,5$

4. UWAGI KOŃCOWE

Przeprowadzone w pracy badania pozwalają na stwierdzenie, że podział wiadomości w sieci na kategorie jest możliwy nie tylko z uwagi na żądanie

określonego rodzaju usługi /metoda KK lub KI/, ale również ze względu na parametry strumienia informacji, jak np. średnia długość wiadomości. Podany w pracy model sieci umożliwia przeprowadzenie tego rodzaju badań. Wydaje się, że istotnym w dążeniu do zmniejszenia opóźnienia byłoby zmodyfikowanie modelu dla obu metod komutacji. Dla metody KK należałoby rozważyć zmniejszenie czasu zestawiania trasy przez wprowadzenie priorytetu bezpośredniego dla wiadomości sterująco-kontrolnych, bądź wydzielenie kanału sygnalizacyjnego. Z kolei w metodzie KI, zrealizowanej w pracy przez komutację wiadomości, należałoby dodatkowo wprowadzić metodę komutacji pakietów.

L i t e r a t u r a

- [1] Błaszczak A.: Analiza opóźnienia przesyłania informacji w sieciach teleinformatycznych z komutacją kanałów. Przegląd Telekomunikacyjny, Nr 2-3, 1982, s.53-58
- [2] Gerla M., Müller D.: Pacuit: The Integrated Packet and Circuit Switching Alternative to Packet Switching. Proc. of COMPCOM Conf. New York, March 1978, s.1-4
- [3] Gitman I., Frank H., Occhiogrosso B., Hsieh W.: Issues in Integrated Network Design. ICC 77 Conf. Record Chicago, June 1977, s.38.1.36-43
- [4] Kermani O., Kleinrock L.: A Tradeoff Study of Switching Systems. ICC 79 Conf. Boston, June 1979
- [5] König D., Stoyan D.: Metody teorii obsługi masowej. WNT, Warszawa 1979
- [6] Kümmerle K., Rudin H.: Packet and Circuit Switching: Cost/Performance Boundaries. Computer Networks, vol.2, No 1, February 1978, s.3-17
- [7] Miyahara H. et al.: Delay and Throughput Evaluation of Switching Methods in Computer Communication Networks. IEEE Trans. on Communications, vol. Com-26, No 3, March 1978, s.337-344
- [8] Rudin H.: Studies on the Intergration of Circuit and Packet Switching. ICC 78 Conf. Proc., Toronto 1978, s.22.1-22.7
- [9] Rudrapatna A.N.: Evaluation of Advanced Switching Technologies for Integrated Voice/Data Applications. IEEE Communications Magazine, No 2, March 1981, s.38-44
- [10] Seidler J.: Analiza i synteza sieci łączności dla systemów teleinformatycznych. PWN, Warszawa 1979
- [11] Tobagi F.A. et al.: Modeling and Measurement Techniques in Packet Communication Networks. Proc. of the IEEE, vol.66, No 11, November 1978, s.1423-1447

EVALUATION OF INTEGRATED SWITCHING METHOD
IN COMMUNICATION NETWORK MODEL

Summary

The paper presents a description of a communication network model integrating the circuit - and message - switching methods. For the model, the performance evaluation is made. The average network delay is taken as a criterion. The considerations are exemplified.

ОЦЕНКА КОММУТАЦИИ ИНТЕГРИРОВАННОЙ В МОДЕЛИ СЕТИ СВЯЗИ

Резюме

Представлена модель сети связи учитывающая интеграцию метода коммутации каналов и метода коммутации информации. Для этой модели указаны результаты оценки качества сети по критерию среднего опоздания во времени перехода информации через сеть. Рассуждения проиллюстрированы примером, причем за критерий качества принято среднее опоздание, которому подвержена информация передаваемая через сети.

Tadeusz Uhl

WYZNACZANIE OKRESU UAKTUALNIANIA TABLIC DOBORU TRAS W SIECIACH Z KOMUTACJĄ PAKIETÓW

W pracy przedstawiono metodę umożliwiającą wyznaczenie przedziału, w którym powinna znaleźć się optymalna wielkość okresu uaktualniania. Jako kryterium przyjęto obciążenie kanałów informacją pomocniczą oraz jej aktualność.

1. WPROWADZENIE

W literaturze poświęconej sterowaniu przesyłaniem wiadomości w sieci teleinformatycznej można znaleźć wiele reguł, które wpływają w nader złożony i na ogół przeciwstawny sposób na podstawowe wskaźniki jakości pracy sieci [1+3, 7+11]. Szczególnie interesująca i ważna w zastosowaniach jest reguła deterministyczna zmiennego doboru trasy, która oparta jest na definicji długości trasy w sensie opóźnień wnoszonych przez poszczególne kanały. W myśl tej reguły, w każdym węźle oddzielnie podejmowane są decyzje o doborze trasy, przy czym uwzględniane są obciążenia własnych buforów wyjściowych i informacje pomocnicze o stanie najbliższych węzłów sąsiednich lub bardziej oddalonych węzłów.

Informacje pomocnicze mogą być przekazywane wydzielonymi do tego celu kanałami, lub, co jest najczęściej spotykane w praktyce, kanałami wykorzystywanymi bezpośrednio do przesyłania pakietów przenoszących informację właściwą. Oczywiście pierwszy sposób przesyłania informacji pomocniczej zwiększa koszt budowy sieci, natomiast drugi zmniejsza efektywną przepustowość kanałów sieci.

Podstawowym warunkiem, jaki musi spełniać informacja pomocnicza, jest jej aktualność. W praktyce, z uwagi na opóźnienia przesyłania informacji przez kanały sieci, informacja pomocnicza po dotarciu do miejsca jej przeznaczenia jest w pewnej mierze zdeaktualizowana. Aby stopień dezaktualizacji informacji pomocniczej był możliwie mały, należy przysyłać ją z możliwie dużą częstotliwością, a zatem okres uaktualniania T_c powinien być z tego punktu widzenia jak najmniejszy. Natomiast biorąc pod uwagę efektywną przepustowość kanałów, należałoby dobierać okres uaktualniania T_c możliwie duży, by zmniejszyć natężenie strumienia informacji pomocniczej. Oczywiście odpowiednia obróbka informacji pomocniczej może wpływać na zwiększenie efektywności jej wykorzystania /np. zastosowanie analizy regresyjnej do pre-

dykcji stanu buforów wyjściowych wężła [7] /. Z kolei zmniejszenie natężenia strumienia informacji pomocniczej jest możliwe poprzez odpowiedni dobór postaci tej informacji [8]. Należy zatem przypuszczać, że zagadnienie doboru okresu uaktualniania jest zagadnieniem optymalizacyjnym. Jednakże jego rozwiązanie, z uwagi na brak odpowiedniej miary jakości oraz problemy wyboru postaci informacji pomocniczej, jej zasięgu i skuteczności jej wykorzystania, jest złożone. W niektórych pracach podejmowane były próby określenia optymalnej wielkości okresu uaktualniania [4], lecz wartość uzyskanych rozwiązań była niewielką z uwagi na pomijanie istotnych czynników. Z tego względu najczęściej w praktyce wartość T_c ustalana była na drodze eksperymentalnej. Ponieważ rozważane zagadnienie jest ważne z praktycznego punktu widzenia, dlatego też w dalszym ciągu przedstawione zostanie nowe do niego podejście analityczne oraz uzyskane wyniki.

2. ROZKŁAD PRAWDOPODOBIEŃSTWA STANU KOLEJKI DLA SYSTEMU OBSŁUGI TYPU M/M/1/K

Rozważmy sieć składającą się z N węzłów. Każdy węzeł posiada J w buforów wyjściowych. Liczbę wiadomości w poszczególnych buforach wyjściowych w dowolnej chwili czasu określa proces stochastyczny $K_j(t)$, $j = 1, 2, \dots, J(w)$. Załóżmy, że strumień zgłoszeń do j -tego bufora jest strumieniem Poissona o intensywności λ_j pakietów/s, a czas obsługi wiadomości ma rozkład wykładniczy o średniej intensywności μ pakietów/s.

Właściwości statystyczne kolejki wiadomości oczekujących w tym buforze na obsługę są określone przez chwilowy rozkład prawdopodobieństwa długości kolejki. Jeżeli przez k oznaczymy długość kolejki w chwili t , to chwilowy rozkład prawdopodobieństwa długości kolejki w buforze o ograniczonej pojemności jest dany wzorem

$$P_k(t) = P [K(t) = k], \quad k = 0, 1, 2, \dots, K \quad (1)$$

gdzie

$$\sum_{k=0}^K P_k(t) = 1$$

a K oznacza pojemność bufora.

Przy uczynionych założeniach chwilowy rozkład prawdopodobieństwa jest rozwiązaniem równania różniczkowego

$$\frac{d\vec{P}(t)}{dt} = A \vec{P}(t) \quad (2)$$

przy warunku początkowym $\vec{P}(0)$, gdzie A jest macierzą trójprzekątniową o wymiarach $(K+1) \times (K+1)$ i przyjmuje postać

$$A = \mu \begin{bmatrix} -\varrho & 1 & 0 & \cdot & 0 & 0 \\ \varrho & -(1+\varrho) & 1 & \cdot & 0 & 0 \\ 0 & \varrho & -(1+\varrho) & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & -(1+\varrho) & 1 \\ 0 & 0 & 0 & \cdot & \varrho & -1 \end{bmatrix} \quad (3)$$

a ponadto^{1/}:

$$\vec{P}(t) = [P_0(t), P_1(t), \dots, P_K(t)]^* \quad (4)$$

natomiast $\varrho = \frac{\lambda}{\mu}$ oznacza współczynnik obciążenia buforu.

Równanie (2) definiuje skończony łańcuch Markowa, którego przebiegi chwilowe $\vec{P}(t)$ można wyrazić w postaci skończonej sumy składników eksponencjalnych o rzeczywistych wykładnikach. W celu otrzymania takiej postaci wektora $\vec{P}(t)$ dokonamy następującej zamiany zmiennych

$$\vec{B}(t) = C \vec{P}(t) \quad (5)$$

gdzie

$$C = \begin{bmatrix} -\frac{1}{2} & & & & & \\ c_0 & 0 & 0 & \cdot & 0 & 0 \\ & c_1 & 0 & \cdot & 0 & 0 \\ & 0 & c_2 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & c_{K-1} & 0 \\ 0 & 0 & 0 & \cdot & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & \cdot & 0 & c_K \end{bmatrix} \quad (6)$$

$$\text{natomiast } c_k = \frac{\varrho^k (1-\varrho)}{1-\varrho^{K+1}} ; \quad k = 0, 1, \dots, K \quad (7)$$

Łatwo zauważyć, że wielkości określone zależnością (7) oznaczają prawdopodobieństwo stamu kolejki M/M/1/K w przypadku równowagi statystycznej.

^{1/} Zapis $\vec{P}(t) = [P_0(t), P_1(t), \dots, P_K(t)]^*$ oznacza wektor transponowany.

Po zamianie zmiennych określonej zależnością (5) równanie (2) przyjmuje postać

$$\frac{d\vec{B}(t)}{dt} = -\mu D \vec{B}(t) \quad (8)$$

gdzie

$$D = \begin{pmatrix} \varrho & -\sqrt{\varrho} & 0 & \cdot & 0 & 0 \\ -\sqrt{\varrho} & (1+\varrho) & -\sqrt{\varrho} & \cdot & 0 & 0 \\ 0 & -\sqrt{\varrho} & (1+\varrho) & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & (1+\varrho) & -\sqrt{\varrho} \\ 0 & 0 & 0 & \cdot & -\sqrt{\varrho} & 1 \end{pmatrix} \quad (9)$$

Biorąc pod uwagę twierdzenia zawarte w [6], możemy rozwiązanie równania (8) przedstawić w formie

$$\vec{B}(t) = e^{-\mu D t} \vec{B}(0) \quad (10)$$

Wykorzystując spektralną postać macierzy $e^{-\mu D t}$ [5] możemy rozwiązanie to określić następującą zależnością

$$\vec{B}(t) = \left[\sum_{i=0}^K e^{-\mu \xi_i t} \vec{B}_i \vec{B}_i^* \right] \vec{B}(0) \quad (11)$$

gdzie:

- ξ_i - i-ta wartość własna macierzy D,
- \vec{B}_i - wektor własny macierzy D odpowiadający wartości własnej ξ_i ,
- $\vec{B}_i \vec{B}_i^*$ - wektory ortonormalne.

Rozwiązanie układu równań różniczkowych (2) przy uwzględnieniu (4) i (1) możemy przedstawić w postaci

$$\vec{P}(t) = \vec{P}_0 + \sum_{i=1}^K h_i \vec{P}_i e^{-\mu \xi_i t} \quad (12)$$

przy czym składowe wektora

$$\vec{P}_0 = [P_{00}, P_{01}, \dots, P_{0K}]^* \quad (13)$$

określają prawdopodobieństwo długości kolejki w warunkach równowagi statystycznej, natomiast

$$\vec{P}_i = C^{-1} \vec{B}_i \quad (14 a)$$

$$h_i = \overline{B}_i C P(0) \quad (14 b)$$

Aby wzór (12) uczynić praktycznie użytecznym, należy określić wartości własne ξ_i . W literaturze [9, 10] wykazano, że szukanymi wartościami własnymi macierzy D są

$$\begin{aligned} \xi_0 &= 0 \\ \xi_i &= 1 + \rho - 2\sqrt{\rho} \cos \theta_i \end{aligned} \quad (15)$$

gdzie

$$\theta_i = i \frac{\pi}{K+1}, \quad i=1,2,\dots,K \quad (16)$$

Natomiast składowe wektora własnego odpowiadającego określonej wartości własnej są dane wzorem

$$B_k = a \cos k\theta + b \sin k\theta, \quad k=0,1,\dots,K \quad (17)$$

gdzie:

$$\begin{aligned} a &= 1 \\ b &= \frac{\cos \theta - \sqrt{\rho}}{\sin \theta} \end{aligned} \quad (18)$$

W rezultacie wzór (12) wraz z wzorami (15 + 18) opisują chwilowy rozkład prawdopodobieństwa długości kolejki w buforze o ograniczonej pojemności.

Znając już teraz chwilowy rozkład prawdopodobieństwa długości kolejki $\overline{P}(t)$ możemy wyznaczyć wartość średnią $\overline{K}(t)$ i wariancję $G_K^2(t)$ zmiennej losowej $K(t)$. Mamy więc

$$\overline{K}(t) = \overline{\alpha}_1^K \overline{P}(t) \quad (19)$$

gdzie

$$\overline{\alpha}_1 = [0, 1, 2, \dots, K]^K$$

oraz

$$G_K^2(t) = \overline{K}^2(t) - [\overline{K}(t)]^2 \quad (20)$$

przy czym

$$\overline{K}^2(t) = \overline{\alpha}_2^K \overline{P}(t) \quad (21)$$

natomiast

$$\overline{\alpha}_2 = [0, 1^2, 2^2, \dots, K^2]^K$$

Znając długość kolejki w chwili t_0 tzn. $K(t_0)$, można średnią $\overline{K}(t)$ opisać za pomocą najwolniej zanikających składników rozkładu (12), a więc

tych, które odpowiadają najmniejszym wartościom własnym $\xi_1, i=1,2,\dots,K$. W pracy [9] podana została aproksymacja 1-go rzędu, tzn. aproksymacja uwzględniająca tylko najwolniej zanikający składnik. Porównanie wyników uzyskanych dla wybranych przykładów w oparciu o wzór uwzględniający wszystkie wartości własne oraz w oparciu o zależności aproksymujące skłania nas do zastosowania w praktyce następującego wzoru aproksymującego dla $\bar{K}(t)$

$$\bar{K}(t) \approx n_0 + [K(t_0) - n_0] e^{-\mu \xi_1 (t-t_0)} \quad (22)$$

gdzie:

$$n_0 = \frac{\alpha_1}{\alpha_1} \frac{K}{P_0} = \sum_{i=1}^K i c_i = \sum_{i=1}^K i \frac{\rho^i (1-\rho)}{1-\rho^{K+1}} \quad (23)$$

oraz wzoru aproksymującego dla $\bar{K}^2(t)$ w postaci

$$\bar{K}^2(t) \approx m_0 + [K^2(t_0) - m_0] e^{-\mu \xi_1 (t-t_0)} \quad (24)$$

gdzie:

$$m_0 = \frac{\alpha_2}{\alpha_2} \frac{K}{P_0} = \sum_{i=1}^K i^2 c_i = \sum_{i=1}^K i^2 \frac{\rho^i (1-\rho)}{1-\rho^{K+1}} \quad (25)$$

Zatem wariancja może być ostatecznie wyrażona zależnością

$$\sigma_K^2(t) = m_0 + [K^2(t_0) - m_0] e^{-\mu \xi_1 (t-t_0)} - \left\{ n_0 + [K(t_0) - n_0] e^{-\mu \xi_1 (t-t_0)} \right\}^2 \quad (26)$$

Wyznaczone wyżej zależności posłużą nam do określenia optymalnej wielkości okresu uaktualniania.

3. DOBÓR OKRESU UAKTUALNIANIA NA PODSTAWIE PRZEWIDYWANEGO STANU BUFORÓW WYJŚCIOWYCH

W celu określenia optymalnej wielkości okresu uaktualniania założymy, że strumień informacji pomocniczej przesyłanej między węzłami sieci stanowi dodatkowy^{1/} strumień zgłoszeń do każdego bufora wyjściowego o intensywności λ_d określonej wzorem

$$\lambda_d = \frac{I}{T_c} \quad (27)$$

^{1/} W odniesieniu do strumienia informacji właściwych.

gdzie:

- I - ilość informacji pomocniczej mierzonej liczbą pakietów przesyłanych w pojedynczym okresie uaktualniania,
- T_c - okres uaktualniania.

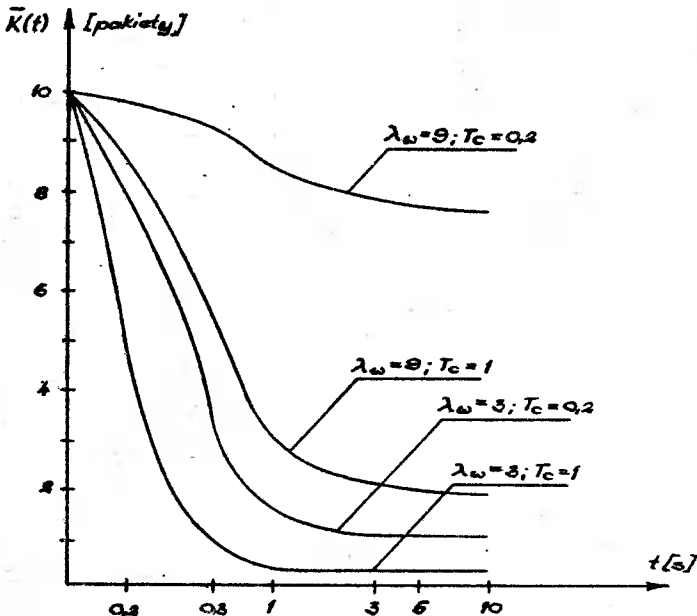
Takie podejście umożliwia uzależnienie ilości informacji pomocniczej od T_c . Jest oczywiste, że z jednej strony wraz z małym okresem uaktualniania T_c wzrastać będzie intensywność strumienia dodatkowego λ_d , co z kolei będzie prowadziło do wzrostu średniej długości kolejek w buforach wyjściowych. Z drugiej strony jednak przy wzroście T_c informacja pomocnicza będzie ulegała dezaktualizacji. W celu znalezienia T_c założymy, że znane jest średnie obciążenie określonego bufora wyjściowego, które określa średnia liczba pakietów $\bar{K}(t)$. Wielkość tę możemy wyznaczyć ze wzoru (22) przy założeniu, że

$$\lambda = \lambda_w + \lambda_d \quad (28)$$

gdzie:

- λ_w - intensywność zgłoszeń do danego bufora pakietów przynoszących informację właściwą,
- λ_d - intensywność zgłoszeń pakietów przynoszących informację pomocniczą.

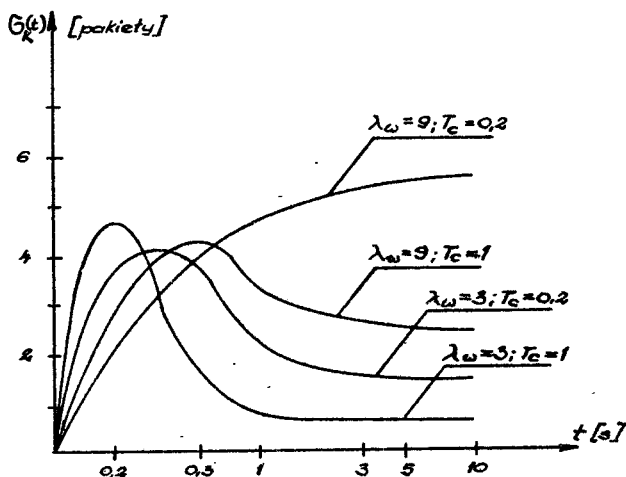
Dla ilustracji graficznej zależności $\bar{K}(t) / T_c = \text{const}$ wykonano symulację dla następujących danych: $\mu = 15$ pakietów/s, $K = 20$ pakietów, $K(t_0) = 10$ pakietów, $T_c = 0,2; 1; 5$ s, $\lambda_w = 3; 9$ pakietów/s. Otrzymane wyniki przedstawiono na rys.1.



Rys.1. Zależność średniej liczby wiadomości w buforze od czasu t dla ustalonych wartości okresu uaktualniania

Z rysunku tego wynika, że wraz z maleniem okresu T_c maleje wartość średnia $\bar{K}(t)$ osiagając wartość n_0 wynikającą z równowagi statystycznej systemu dla dużych wartości t . Łatwo zauważyć, że wartość n_0 zależy również od T_c zgodnie z zasadą: im mniejsze T_c tym większa wartość n_0 . Możemy zatem, znając stan początkowy $K(t_0)$ oraz intensywność strumienia głównego λ_w , wyznaczyć taką wartość T_{c1} , aby współczynnik wykorzystania kanału nie przekraczał narzuconej wartości. Będzie to wybór wartości okresu uaktualniania dokonany ze względu na obciążalność kanałów.

W celu zobrazowania przedziału, w którym mogą znaleźć się stany buforu należy przeanalizować wariancję procesu $K(t)$. Jak wynika z zależności (20) wariancja ta zależy od wartości współczynnika φ poprzez wartość własną ξ_1 i jest również funkcją czasu, osiagając dla $t=0$ wartość $\sigma_K^2(0) = 0$, a dla $t \rightarrow \infty$ wariancja $\sigma_K^2(\infty) = n_0 - n_0^2$. Charakter zmian tej funkcji dla danych z poprzedniego przykładu pokazano na rys.2.



Rys.2. Zależność odchylenia standardowego liczby wiadomości w buforze od czasu t dla ustalonych okresów uaktualniania

Jak wiemy dla zachowania właściwości adaptacyjnych reguły doboru tras informacja pomocnicza powinna być możliwie aktualna. Obserwując zależności przedstawione na rys.2 możemy przyjąć przebieg odchylenia standardowego w czasie jako pewną miarę aktualności informacji pomocniczej. Oczywiście interesować nas będzie przebieg odchylenia standardowego tylko w pewnym początkowym przedziale czasu, w którym podlega ono dużym zmianom. Znając więc $\sigma_K(t)$ oraz intensywność λ_w możemy wyznaczyć okres T_{c2} uaktualniania tablic doboru tras w ten sposób, aby nie została przekroczona przez $\sigma_K(t)$ pewna zadana wartość σ_p . Uwzględniając przy tym wyniki rozważań dotyczące wyboru T_{c1} możemy ostatecznie określić przedział wielkości czasu uaktualniania, z którego wybierzemy ostateczną wielkość.

4. PODSUMOWANIE

Przedstawione w pracy podejście porusza tylko niektóre aspekty złożonego problemu, jakim jest wyznaczanie wielkości okresu uaktualniania tablic doboru tras. Zagadnienie to rozpatrywane było przy uwzględnieniu obciążenia kanałów informacją pomocniczą oraz aktualności informacji pomocniczej, za miarę której przyjęto odchylenie standardowe przewidywanej wartości średniej liczby pakietów w buforze wyjściowym węzła. Można tu wysunąć pewne zastrzeżenia dotyczące założenia, że średnia długość informacji pomocniczej jest równa średniej długości wiadomości. Należy jednak stwierdzić, że otrzymane wyniki umożliwiają wyznaczenie przedziału, w którym powinny się znajdować wartości okresu uaktualniania tablic doboru tras. Ponadto w sytuacji, gdy obciążenie buforów wyjściowych znacznie powiększy się wskutek wzrostu natężenia strumienia wiadomości - otrzymane wyniki umożliwiają racjonalną zmianę okresu uaktualniania, aby zwiększyć efektywną przepustowość kanałów.

Z uwagi na złożoność poruszanego problemu niezbędne są dalsze prace na ten temat, które w rezultacie prowadzić będą do zwiększenia efektywności protokołu komunikacyjnego.

L i t e r a t u r a

- [1] Agnew E.: On Quadratic Adaptive Routing Algorithms. Communications of the ACM, vol.19, No 1, January 1976, s.18-22
- [2] Boorstyn R., Livne A.: A Technique for Adaptive Routing in Networks. IEEE Trans. on Commun., vol.29, No 4, April 1981, s.474-480
- [3] Chou W., Bragg A., Nilsson A.: The Need for Adaptive Routing in the Chaotic and Unbalanced Traffic Environment. IEEE Trans. on Commun., vol.29, No 4, April 1981, s.481-490
- [4] Holtzman J.: Routing Updates in Packet Switched Networks. Computer Performance, North Holland Publishing Company, 1977, s.537-543
- [5] Kaczorek T.: Teoria układów regulacji automatycznej. PWN, Warszawa 1972
- [6] Kowalczyk B.: Macierze i ich zastosowania. WNT, Warszawa 1976
- [7] Rutkowski D., Uhl T.: Adaptive Routing Based on the Linear Prediction of Queue Length. Proc. Dalkovy Prensos Dat, Karlovy Vary 1981, s.II.46-II.48
- [8] Rutkowski D., Uhl T.: Adaptive Routing Algorithm Using an Auxiliary Information Concerning the State of Neighboring Nodes. Artykuł złożony w IEEE Trans. on Commun.

- [9] Stern T.: Aproximations of Queue Dynamics and their Application to Adaptive Routing in Computer Communication Networks. NTC Birmingham, Alabama, December 1978, s.4.4.1 -4.4.5
- [10] Uhl T.: Algorytmy lokalnego doboru tras w sieci teleinformatycznej i ich implementacja. Praca doktorska, PG Gdańsk 1981
- [11] Yum Tak-Shing P.: The Design and Analysis of a Semidynamic Deterministic Routing Rule. IEEE Trans. on Comm., vol.29, No 4, April 1981, s.498-504

EVALUATION OF ROUTING TABLES UPDATING PERIOD
IN PACKET - SWITCHED NETWORKS

Summary

There is presented in the paper a method of finding the time interval in which the optimal value of routing table updating period should be detected. The channel overhead associated with an auxiliary information and its topicality are taken as a criterion.

ПОДБОР ПЕРИОДА КОРРЕКЦИИ МАТРИЦ МАРШРУТОВ В СЕТЯХ С КОММУТАЦИЕЙ ПАКЕТОВ

Реюме

В работе предлагается метод определения диапазона оптимальных значений периода коррекции. В качестве критериев принимается с одной стороны загрузка каналов дополнительной информацией, а с другой - актуальность этой информации.

Krzysztof Chmara

CHARAKTERYSTYKI WĘZŁA SIECI TELEINFORMATYCZNEJ O OGRANICZONEJ PAMIĘCI BUFOROWEJ

W pracy wyznaczone zostały w sposób analityczny charakterystyki natężenia strumienia wiadomości przekazywanych przez węzeł, średniego opóźnienia wnoszonego przez węzeł oraz mocy węzła w funkcji natężenia zgłoszeń dla węzła rozważanego w izolacji, jak też przy uwzględnieniu oddziaływania następnego węzła na trasie przesyłania wiadomości i gdy nie jest stosowana reguła przeciwdziałania przeciążeniu.

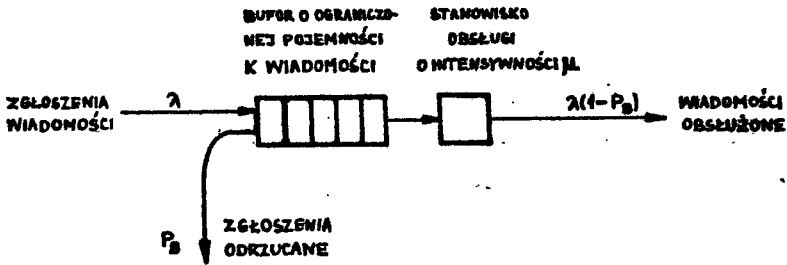
1. WSTĘP

W celu uzyskania sprawnego działania sieci teleinformatycznej konieczne jest stosowanie odpowiednich reguł sterowania przesyłaniem w niej wiadomości, w tym zwłaszcza reguł doboru tras i reguł przeciwdziałania przeciążeniom. Jakość pracy sieci pozwalają określić jej podstawowe charakterystyki, takie jak całkowite natężenie strumienia wiadomości przesyłanych w sieci oraz ich średnie opóźnienie w funkcji natężenia zgłoszeń wiadomości, przy czym pożądane jest uzyskanie możliwie dużego natężenia strumienia wiadomości przesyłanych w sieci dla danego natężenia zgłoszeń przy możliwie małym średnim opóźnieniu. Możliwe jest też określenie charakterystyki syntetyzującej wymienione wymagania. Charakterystyka ta, zwana mocą, została wprowadzona przez Glesslera i innych [3], a rozwinięta przez Kleinrocka [4, 5] i jest zdefiniowana jako iloraz całkowitego natężenia strumienia wiadomości przesyłanych w sieci i średniego opóźnienia.

W pracy tej przedstawione zostaną podstawowe charakterystyki węzła sieci teleinformatycznej o ograniczonej pamięci buforowej. Zwrócona zostanie uwaga na pogorszenie jakości pracy węzła wraz ze wzrostem natężenia zgłoszeń i na potrzebę stosowania odpowiednich reguł sterowania przesyłaniem informacji. Najpierw rozważony będzie węzeł izolowany, a następnie uwzględnione zostanie oddziaływanie węzła następnego na trasie przesyłania wiadomości. Należy tu zaznaczyć, że znajomość wspomnianych charakterystyk węzła pozwala na wyciągnięcie w prosty sposób uogólniających wniosków o charakterze jakościowym dotyczących właściwości całej sieci i może służyć do określania założeń na poszczególne węzły przy projektowaniu sieci.

2. CHARAKTERYSTYKI WĘZŁA IZOLOWANEGO

Weźmy pod uwagę izolowany węzeł sieci, do którego dopływa z intensywnością λ poissonowski strumień wiadomości o wykładniczym rozkładzie długości. Przyjmijmy, że węzeł znajduje się pula buforów o ograniczonej pojemności K oraz nadajnik^{1/} o intensywności obsługi μ . Załóżmy na razie, że nie występuje oddziaływanie następnego węzła, do którego przesyłane są wiadomości z interesującego nas węzła. Mamy zatem do czynienia z systemem obsługi $M/M/1/K$ i możemy wykorzystać znane wyniki analizy tego typu systemów obsługi [6, 9]. Ogólny schemat rozważanego węzła jest przedstawiony na rys.1.



Rys.1. Schemat węzła o ograniczonej pojemności pamięci buforowej
 P_B oznacza prawdopodobieństwo odrzucania wiadomości/

Wiadomo, że prawdopodobieństwo zdarzenia, iż w warunkach równowagi statystycznej w węźle typu $M/M/1/K$ znajduje się k wiadomości, jest określone wzorem [9]:

$$P(k) = \begin{cases} \frac{(1 - \rho) \rho^k}{1 - \rho^{K+2}} & \text{dla } k \leq K+1 \\ 0 & \text{dla } k > K+1 \end{cases} \quad (1)$$

gdzie

$$\rho = \frac{\lambda}{\mu} \quad (2)$$

Prawdopodobieństwo P_B nieprzyjęcia napływających wiadomości do węzła jest równe prawdopodobieństwu całkowitego zapełnienia buforu w węźle, a więc

^{1/} Podobne rozważania możemy przeprowadzić, gdy w węźle będzie więcej stanowisk obsługi, jeśli tylko każdy strumień wiadomości będzie kierowany do odpowiedniego buforu wyjściowego niezależnie od innych.

$$P_B = P(K+1) = \frac{(1-\rho)\rho^{K+1}}{1-\rho^{K+2}} \quad (3)$$

W dalszym ciągu interesować nas będzie charakterystyka natężenia wiadomości opuszczających węzeł w funkcji natężenia zgłoszeń. Charakterystykę tę wyznaczmy ze wzoru [8]

$$S(\lambda) = \lambda(1 - P_B) \quad (4)$$

Po podstawieniu (3) do (4) otrzymamy

$$S(\lambda) = \lambda \left[1 - \frac{(1-\rho)\rho^{K+1}}{1-\rho^{K+2}} \right] = \frac{\lambda(1-\rho^{K+1})}{1-\rho^{K+2}} \quad (5)$$

a po uwzględnieniu (2) mamy

$$S(\lambda) = \frac{\lambda\mu(\mu^{K+1} - \lambda^{K+1})}{\mu^{K+2} - \lambda^{K+2}} \quad (6)$$

W celu wyznaczenia średniego opóźnienia wnoszonego przez węzeł musimy znaleźć średnią liczbę wiadomości przebywających w buforze. Jest ona określona następująco

$$\bar{k}_b = \sum_{k=2}^{K+1} (k-1) P(k) \quad (7)$$

Po podstawieniu (1) do (7) dostaniemy

$$\bar{k}_b = \sum_{k=2}^{K+1} (k-1) \frac{(1-\rho)\rho^k}{1-\rho^{K+2}} = \frac{\rho^2 [1 - \rho^K (1 + K - K\rho)]}{(1-\rho)(1-\rho^{K+2})} \quad (8)$$

Na mocy zależności Little'a średni czas oczekiwania w buforze jest równy

$$\bar{T}_b = \frac{\bar{k}_b}{\lambda(1 - P_B)} = \frac{\bar{k}_b}{S(\lambda)} \quad (9)$$

Po podstawieniu (8) oraz (5) do (9) otrzymamy

$$\bar{T}_b = \frac{\rho^2 [1 - \rho^K (1 + K - K\rho)]}{\lambda(1-\rho)(1-\rho^{K+1})} \quad (10)$$

Ponieważ całkowite średnie opóźnienie wnoszone przez węzeł $\bar{T}(\rho)$ jest równe sumie średniego czasu oczekiwania wiadomości w buforze i średniego

czasu nadawania wiadomości, więc

$$\bar{T}(\rho) = \frac{\rho^2 [1 - \rho^K (1 + K - K\rho)]}{\lambda(1-\rho)(1-\rho^{K+1})} + \frac{1}{\mu} = \frac{\rho^2 [1 - \rho^K (1 + K - K)]}{\lambda(1-\rho)(1-\rho^{K+1})} + \frac{\rho}{\lambda} = \frac{\rho \{1 - \rho^{K+1} [K+2 - (K+1)\rho]\}}{\lambda(1-\rho)(1-\rho^{K+1})} \quad (11)$$

Aby sprawdzić, czy funkcja $\bar{T}(\rho)$ jest ciągła i posiada ciągłą pochodną w całym zakresie zmienności ρ , zapiszemy ją w postaci

$$\bar{T}(\rho) = \frac{(K+1)\rho^{K+2} - (K+2)\rho^{K+1} + 1}{\mu(1-\rho)(1-\rho^{K+1})} \quad (12)$$

Łatwo można zauważyć, że wielomian w liczniku zależności (12) przyjmuje wartość równą zero dla $\rho = 1$, a zatem wielomian ten jest podzielny przez $(\rho - 1)$. Wobec tego licznik i mianownik funkcji (12) można podzielić przez $(\rho - 1)$ i w rezultacie otrzymamy

$$\begin{aligned} \bar{T}(\rho) &= \frac{(K+1)\rho^{K+1}(\rho-1) - (\rho^{K+1} - 1)}{\mu(\rho-1)(\rho^{K+1} - 1)} = \\ &= \frac{(K+1)\rho^{K+1}(\rho-1) - (\rho-1)(\rho^K + \rho^{K-1} + \dots + \rho + 1)}{\mu(\rho-1)(\rho-1)(\rho^K + \rho^{K-1} + \dots + \rho + 1)} = \\ &= \frac{(K+1)\rho^{K+1} - \rho^K - \rho^{K-1} - \dots - \rho - 1}{\mu(\rho-1)(\rho^K + \rho^{K-1} + \dots + \rho + 1)} \quad (13) \end{aligned}$$

Po kolejnym podzieleniu licznika i mianownika funkcji (13) przez $(\rho - 1)$ dostaniemy

$$\begin{aligned} \bar{T}(\rho) &= \frac{(K+1)\rho^K + K\rho^{K-1} + (K-1)\rho^{K-2} + \dots + 2\rho + 1}{\mu(\rho^K + \rho^{K-1} + \dots + \rho + 1)} = \\ &= \frac{\sum_{k=0}^K (k+1)\rho^k}{\mu \sum_{k=0}^K \rho^k} \quad (14) \end{aligned}$$

Z postaci (14) funkcji $\bar{F}(\rho)$ wynika, że jest ona ciągła i posiada ciągłą pochodną dla $\rho > 0$. Funkcję tę możemy też zapisać w formie jawnej zależności od λ i w rezultacie mamy

$$\bar{F}(\lambda) = \frac{\sum_{k=0}^K (k+1) \left(\frac{\lambda}{\mu}\right)^k}{\mu \sum_{k=0}^K \left(\frac{\lambda}{\mu}\right)^k} \quad (15)$$

Znając $S(\lambda)$ (wzór [5]) oraz $\bar{F}(\lambda)$ (wzór [15]), możemy wyznaczyć wskaźnik zwany mocą węzła, który ujmujemy w formie syntetycznej podstawowe właściwości węzła. Wskaźnik ten jest określony w następujący sposób [3, 4, 5]

$$V(\lambda) = \frac{S(\lambda)}{\bar{F}(\lambda)} = \frac{\lambda(1 - \rho^{K+1})}{(1 - \rho^{K+2})} \cdot \frac{\mu \sum_{k=0}^K \rho^k}{\sum_{k=0}^K (k+1) \rho^k} \quad (16)$$

a po podzieleniu licznika i mianownika przez $(\rho - 1)$ przyjmuje postać

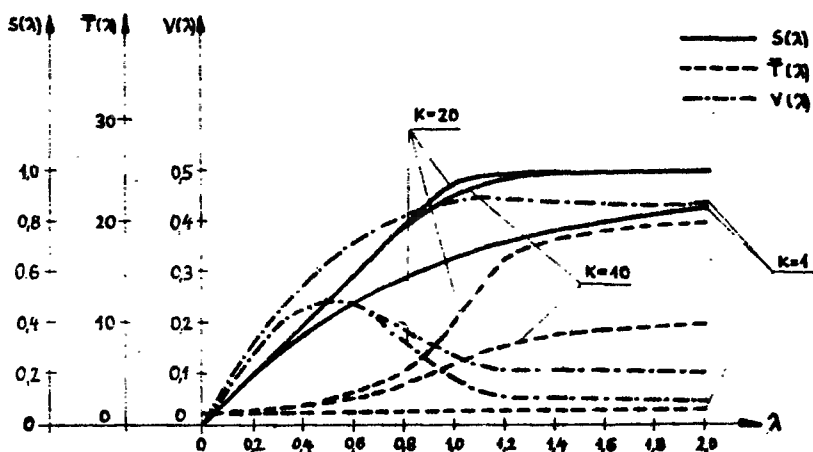
$$V(\lambda) = \frac{\lambda \mu \left(\sum_{k=0}^K \rho^k\right)^2}{\sum_{j=0}^{K+1} \rho^j \sum_{k=0}^K (k+1) \rho^k} \quad (17)$$

Charakterystyki natężenia strumienia wiadomości przekazywanych przez węzeł, średniego opóźnienia wnoszonego przez węzeł oraz mocy węzła o ograniczonej pamięci buforowej w zależności od intensywności zgłoszeń wiadomości, wyznaczone na podstawie wzorów (5), (15) oraz (17) przy założeniu $\mu = 1$, są pokazane na rys. 2.

Uzyskane wyniki badań właściwości węzła o ograniczonej pamięci buforowej nie mogą być bezpośrednio przenoszone na łańcuch węzłów i kanałów. Dla takiej sieci nie można bowiem przyjąć hipotezy Kleinrocka o niezależności [7], gdyż w rozważanym łańcuchu występują powiązania statystyczne pomiędzy liczbami wiadomości w poszczególnych węzłach, a zatem powiązania te dotyczą również opóźnień wnoszonych przez te węzły. Dla łańcucha węzłów i kanałów można natomiast wyznaczyć charakterystykę $S_c(\lambda)$ natężenia strumienia wiadomości przesyłanych przez ten łańcuch w funkcji natężenia zgłoszeń. Jeśli więc przyjmujemy, że łańcuch jest utworzony z $(N+1)$ węzłów typu M/M/1/K oraz N łączących je kanałów, to

$$S_c(\lambda) = \lambda(1 - P_{B_1})(1 - P_{B_2}) \dots (1 - P_{B_N}) \quad (18)$$

gdzie P_{B_i} , $i=1,2,\dots,N$, jest prawdopodobieństwem nieprzyjęcia napływających wiadomości do i -tego węzła.



Rys.2. Charakterystyki natężenia strumienia wiadomości przekazywanych przez węzeł, średniego opóźnienia wnoszonego przez węzeł oraz mocy węzła w funkcji natężenia zgłoszeń dla węzła rozważanego w izolacji

3. CHARAKTERYSTYKI WĘZŁA PRZY UWZGLĘDNIENIU ODDZIAŁYWANIA WĘZŁA NASTĘPNEGO

Z punktu widzenia przeciwdziałania przeciążeniu węzła ważne jest zbadanie oddziaływania następnego węzła, do którego przesyłane są wiadomości. Rozważmy ten przypadek szczególnie i weźmiemy pod uwagę łańcuch utworzony z dwóch węzłów i łączącego je kanału oraz wyznaczmy podstawowe charakterystyki opisujące właściwości określonego węzła współpracującego z węzłem sąsiednim. Fakt współpracy dwóch sąsiednich węzłów można uwzględnić w analizie poprzez uzależnienie intensywności wyprowadzania wiadomości z interesującego nas węzła od intensywności wprowadzania wiadomości do węzła następnego [8]. W tym celu wprowadzamy efektywną intensywność obsługi

$$\mu_e = \mu(1 - P_B^c) \quad (19)$$

gdzie P_B^c jest prawdopodobieństwem nieprzyjęcia wiadomości do następnego węzła.

Zmniejszenie intensywności obsługi z μ do μ_e wpływa na zmianę współczynnika obciążenia interesującego nas węzła, który przyjmuje teraz wartość

$$\rho_1 = \frac{\rho}{1 - \rho_B'} \quad (20)$$

Wobec tego we wzorach określających podstawowe charakterystyki węzła należy zamienić ρ na ρ_1 . Jeżeli następny węzeł jest węzłem docelowym, to

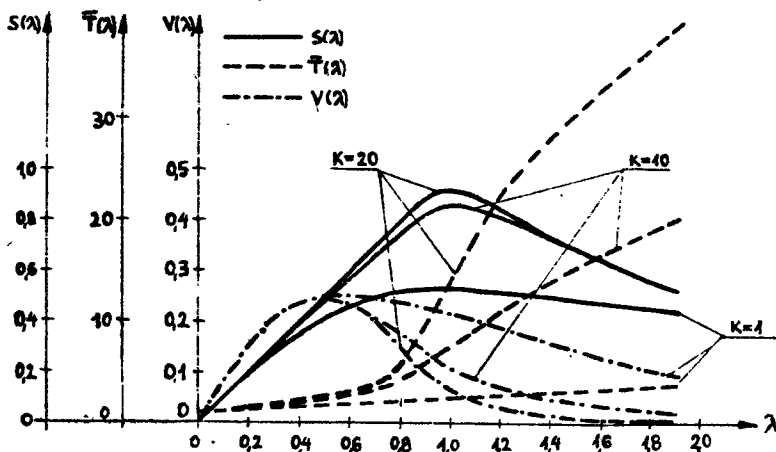
$$\rho_1 = \frac{\rho}{1 - \frac{(1-\rho)\rho^{K+1}}{1-\rho^{K+2}}} = \frac{\rho(1-\rho^{K+2})}{1-\rho^{K+1}} \quad (21)$$

Natomiast charakterystyka strumienia wiadomości przekazywanych przez węzeł przyjmuje w tym przypadku postać

$$s(\lambda) = \frac{\lambda(1-\rho_1^{K+1})}{1-\rho_1^{K+2}} \quad (22)$$

przy czym ρ_1 jest określone wzorem (21).

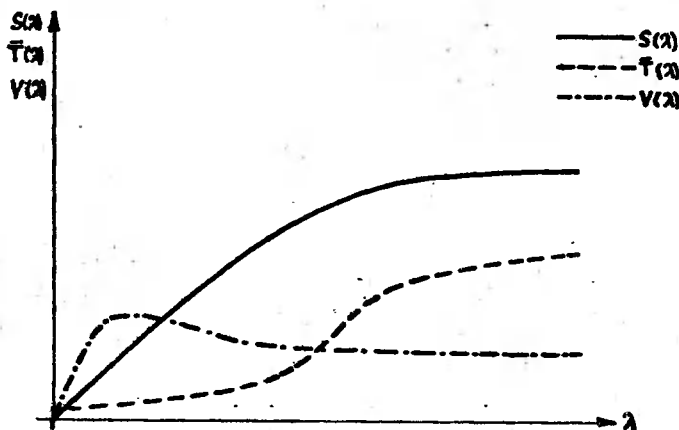
Podobnie całkowite średnie opóźnienie oraz moc węzła można łatwo określić na podstawie wzorów odpowiednio (15) i (17) po podstawieniu w nich ρ_1 zamiast ρ . Otrzymane w ten sposób zależności umożliwiają wyznaczenie podstawowych charakterystyk /patrz rys.3/ opisujących właściwości węzła współpracującego z węzłem sąsiednim.



Rys.3. Charakterystyki natężenia strumienia wiadomości przekazywanych przez węzeł, średniego opóźnienia wnoszonego przez węzeł oraz mocy węzła w funkcji natężenia zgłoszeń przy uwzględnieniu oddziaływania węzła następnego

4. PODSUMOWANIE

Przedstawione charakterystyki węzła o ograniczonej pamięci buforowej, a szczególnie charakterystyka całkowitego natężenia strumienia przesyłanych wiadomości, wskazują przy dużym obciążeniu na pogarszanie jakości pracy węzła ze wzrostem intensywności zgłoszeń. Wynika stąd potrzeba stosowania odpowiednich reguł sterowania przesyłaniem informacji, pozwalających w pożądanym sposób zmieniać opisywane charakterystyki. Pożądane przebiegi tych charakterystyk uzyskane w wyniku zastosowania reguły przeciwdziałania przeciążeniu są przedstawione na rys.4, natomiast poszczególne reguły przeciwdziałania przeciążeniu węzła omówione są w pracach [1, 2].



Rys.4. Przykładowe charakterystyki natężenia strumienia wiadomości przekazywanych przez węzeł, średniego opóźnienia wnoszonego przez węzeł oraz mocy węzła w funkcji natężenia zgłoszeń w przypadku zastosowania reguły przeciwdziałania przeciążeniu

Znajomość opisanych jak też poświadczonych charakterystyk pojedynczego węzła pozwala wyciągać wnioski dotyczące właściwości całej sieci, a ponadto umożliwia na etapie projektowania sieci formułowanie założeń na poszczególne węzły, określających ich współczynniki obciążenia oraz pojemności pamięci buforowej przy wymaganych wskaźnikach jakościowych.

Literatura

- [1] Chmara K.: Przeciwdziałanie przeciążeniom lokalnym sieci teleinformatycznej. Praca doktorska, Wydział Elektroniki PŁ, Gdańsk 1981
- [2] Chmara K., Rutkowski D.: Przeciwdziałanie przeciążeniom sieci teleinformatycznej, rozdz. VI w książce "Projektowanie sieci przesyłania informacji dla systemów teleinformatycznych", która ukaże się nakładem Wydawnictwa Politechniki Wrocławskiej

- [3] Giessler A., Haenle J., Koenig A., Pade E.: Free Buffer Allocation- An Investigation by Simulation. Computer Networks, vol.2, 1978, /s.191-208
- [4] Kleinrock L.: On Flow Control in Computer Networks. International Conference on Communications 78, Toronto, June 1978, Conference Record, s.27.2.1-5
- [5] Kleinrock L.: Power and Deterministic Rules of Thumb for Probabilistic Problems in Computer Communications. International Conference on Communications 79, Boston, June 1979, Conference Record, s.43.1.1-10
- [6] Kleinrock L.: Queueing Systems. Vol.I: Theory. John Wiley and Sons, New York, 1975
- [7] Kleinrock L.: Queueing Systems. Vol.II: Computer Applications. John Wiley and Sons, New York, 1976
- [8] Schwartz M., Saad S.: Analysis of Congestion Control Techniques in Computer Communication Networks. J.L.Grangé, M.Gien, Eds, Flow Control in Computer Networks, North-Holland, Amsterdam, 1979, s.113-130
- [9] Seidler J.: Analiza i synteza sieci łączności dla systemów teleinformatycznych. PWN, Warszawa 1979

CHARACTERISTICS OF A COMPUTER COMMUNICATION NETWORK NODE
WITH FINITE BUFFER STORAGE

Summary

In the paper the node throughput, the average time delay and the power versus average arrival rate characteristics for isolated node as well as the influence of the next node on the message route and when no congestion control rule is used, are analytically determined.

ХАРАКТЕРИСТИКА УЗЛА ВЫЧИСЛИТЕЛЬНОЙ СЕТИ С ОГРАНИЧЕННОЙ БУФЕРНОЙ ПАМЯТЬЮ

Резюме

В работе аналитическим способом определены зависимости фактической пропускной способности узла, средней задержки сообщений и коэффициента мощности узла от входного потока сообщений для изолированного узла, одновременно учитывая воздействие следующего узла на маршруте передачи сообщений и когда не применяется принцип управления потоками данных.

Tadeusz Wysocki jr

O PEWNYM SPOSOBIE
MODULACJI PRZEBIEGU HARMONICZNEGO
SYGNAŁEM CYFROWYM

W artykule przedstawiony jest sposób cyfrowej modulacji przyspieszenia kąowego /AASK/. Przeprowadzona jest szczegółowa analiza sygnału zmodulowanego w dziedzinie czasu. Ponadto przedstawione są wyniki obliczeń gęstości widmowej mocy sygnału AASK.

W celu wykorzystania istniejących łączy analogowych do transmisji danych, trzeba w procesie modulacji przetworzyć sygnał cyfrowy na sygnał analogowy. W niniejszym opracowaniu zostanie przedstawiony jeden z możliwych sposobów modulacji nośnego przebiegu harmonicznego sygnałem cyfrowym.

Rozważmy nośny przebieg harmoniczny $a(t)$ opisany równaniem

$$a(t) = A \cos(\omega_0 t + \varphi_0) \quad (1)$$

gdzie:

A - amplituda przebiegu

ω_0 - pulsacja przebiegu

φ_0 - faza początkowa; $\varphi_0 \in \langle 0, 2\pi \rangle$

Sygnał modulujący $f(t)$ może oddziaływać na amplitudę przebiegu lub na argument funkcji trygonometrycznej w równaniu (1), a więc sygnał zmodulowany możemy zapisać jako

$$g(t) = A(t) \cos[\varphi(t)] \quad (2)$$

gdzie $\varphi(t)$ jest wartością chwilową fazy.

Jeżeli wartość chwilowa fazy w równaniu (2) zależna jest od sygnału modulującego $f(t)$, a amplituda $A(t)$ jest stała w czasie, tj.

$$A(t) = A = \text{const.} \quad (3)$$

zachodzi przypadek modulacji kąowej. Gdy wartość chwilowa fazy przebiegu zmodulowanego zależna jest w sposób liniowy od sygnału modulującego, tj.

$$\varphi(t) = \omega_0 t + \varphi_0 + C_1 f(t) ; C_1 = \text{const.} \quad (4)$$

mamy do czynienia z modulacją fazy [2]. Modulacja częstotliwości [2] zachodzi natomiast w przypadku, gdy pochodna $\frac{d\varphi(t)}{dt}$ zależy w sposób liniowy od sygnału modulującego, tj. gdy

$$\frac{d}{dt} [\varphi(t)] = C_2 f(t) + \omega_0 ; \quad C_2 = \text{const.} \quad (5)$$

Oba wymienione typy modulacji są powszechnie stosowane w systemach transmisji informacji cyfrowych.

W niniejszym opracowaniu zajmiemy się sposobem modulacji kątowej, dla której zachodzi zależność:

$$\frac{d^2}{dt^2} [\varphi(t)] = C_3 f(t) ; \quad C_3 = \text{const.} \quad (6)$$

Sposób ten będziemy nazywać modulacją przyspieszenia kątowego, a w przypadku, gdy sygnał modulujący jest sygnałem cyfrowym będziemy stosować skrót AASK /ang. Angular Acceleration Shift Keying/.

Przebieg sygnału z modulacją przyspieszenia kątowego opisany jest wzorem

$$g(t) = A \cos \left[\omega_0 t + \varphi_0 + C_3 \int_0^t \int_0^{t_2} f(t_1) dt_1 dt_2 \right] \quad (7)$$

a pulsacja chwilowa $\omega(t)$ takiego sygnału zależnością

$$\omega(t) = \frac{d}{dt} [\varphi(t)] = \omega_0 + C_3 \int_0^t f(t_1) dt_1 \quad (8)$$

Jeżeli sygnał modulujący $f(t)$ jest sygnałem izochronicznym o odstępie jednostkowym T , to sygnał zmodulowany można przedstawić w postaci:

$$g(t) = A \cos \left[\omega_0 t + \varphi_0 + C_3 \sum_{n=0}^N b_n \int_0^t \int_0^{t_2} h(t_1 - nT) dt_1 dt_2 \right] \quad (9)$$

gdzie:

$$h(t) = \begin{cases} 1 & \text{gdy } 0 \leq t < T \\ 0 & \text{w przeciwnym przypadku} \end{cases} \quad (10)$$

b_n - wartość n -tego impulsu w sygnale modulującym

N - liczba impulsów w sygnale modulującym, zakończonych do chwili t

Uwzględniając (10) w wyrażeniu (9), możemy wartość chwilową fazy przebiegu zmodulowanego opisać zależnością:

$$\begin{aligned}
 \varphi(t) &= \omega_0 t + \varphi_0 + c_3 \sum_{n=0}^N b_n \int_0^t \int_0^{t_2} h(t_1 - nT) dt_1 dt_2 = \\
 &= \omega_0 t + \varphi_0 + c_3 \sum_{n=0}^{N-1} b_n \int_0^t \int_{nT}^{nT+T} dt_1 dt_2 + c_3 b_N \int_0^t \int_{NT}^{t_2+NT} dt_1 dt_2 = \\
 &= \frac{1}{2} \xi_N t^2 + \omega_N t + \varphi_0
 \end{aligned} \tag{11}$$

gdzie:

$$\xi_N = c_3 b_N - \text{przyspieszenie kątowe}^1/ \tag{12}$$

$$\omega_N = c_3 T \sum_{n=0}^{N-1} b_n + \omega_0 \tag{13}$$

Przebieg zmodulowany opisany jest zatem zależnością

$$g(t) = A \cos \left(\frac{1}{2} \xi_N t^2 + \omega_N t + \varphi_0 \right) \tag{14}$$

a pulsacja chwilowa jest równa

$$\omega(t) = t \xi_N + \omega_N \tag{15}$$

Z zależności (13) i (15) wynika, że sygnał modułujący musi być sygnałem o skończonej bieżącej sumie cyfrowej σ [3]

$$\sigma = \sum_{n=0}^{N-1} b_n \tag{16}$$

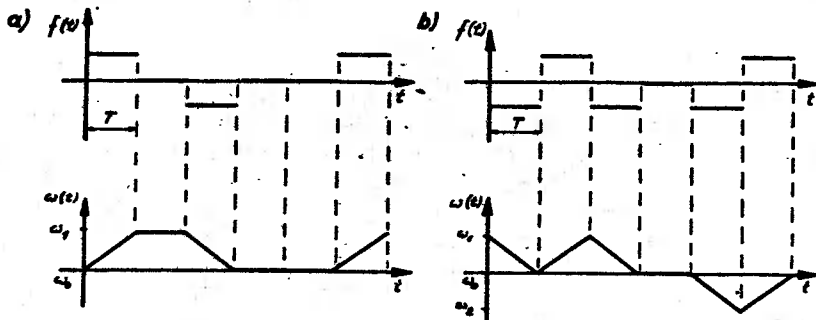
W przeciwnym przypadku pulsacja chwilowa przebiegu zmodulowanego mogłaby przyjmować wartości nieskończone. Ponieważ sygnał modułujący musi być ciągiem słów kodowych kodu o skończonej bieżącej sumie cyfrowej, więc w sygnale zdemodulowanym możemy wykrywać pojedyncze przekłamania [3].

Na podstawie zależności (13) i (15) stwierdza się ponadto, że wahania pulsacji chwilowej sygnału z modulacją AASK są tym mniejsze, im mniejsze są wahania bieżącej sumy cyfrowej σ . Wahania te są najmniejsze, gdy sygnał modułujący jest ciągiem słów kodowych kodu AMI, dla którego bieżąca suma cyfrowa może przyjmować tylko dwie wartości /0 lub 1 albo 0 lub -1 [3]/.

Na rys.1a pokazano zależność pulsacji chwilowej sygnału zmodulowanego od sygnału modułującego, gdy sygnał modułujący jest ciągiem słów kodowych kodu AMI, a na rys.1b, gdy sygnał modułujący jest ciągiem słów kodowych

^{1/} W dalszej części opracowania nazywać je będziemy krótko przyspieszeniem.

kodu HDB2 / $\sigma \in \{-1, 0, 1\}$ [3] /.



Rys.1. Zależności zachodzące między sygnałem modulującym a pulsacją chwilową przebiegu zmodulowanego, gdy sygnał modulujący jest:

- a/ ciągiem słów kodowych kodu AMI
b/ ciągiem słów kodowych kodu HDB2

Ze względu na to, że wahania pulsacji chwilowej przebiegu zmodulowanego są najmniejsze, gdy sygnał modulujący jest ciągiem słów kodowych kodu AMI, w dalszej części opracowania będziemy się zajmować tylko tym przypadkiem. Większość rozważań daje się jednak uogólnić dla dowolnego kodu skończonej bieżącej sumy cyfrowej.

Jeżeli sygnał modulujący jest ciągiem słów kodowych kodu AMI, to pulsacja chwilowa przebiegu zmodulowanego /zob. rys.1a/ zmienia się w granicach od ω_0 do ω_1 . Wprowadźmy współczynnik modulacji zdefiniowany wzorem

$$m = \frac{\omega_1 - \omega_0}{2\pi} T \quad (17)$$

W kodzie AMI występują trzy rodzaje impulsów: 0, +1, -1. Z zależności (12) wynika, że w sygnale zmodulowanym mogą wystąpić również trzy różne wartości przyspieszenia: 0, ξ_+ , ξ_- . Zachodzi ponadto zależność

$$|\xi_+| = |\xi_-| \stackrel{\text{def}}{=} k \quad (18)$$

Korzystając z własności kodu AMI oraz zależności (12), (13), (15) i (18) możemy zapisać

$$\omega_1 - \omega_0 = kT \quad (19)$$

Podstawiając (19) do (17) otrzymujemy

$$m = \frac{kT^2}{2\pi} \quad (20)$$

skąd mamy

$$k = \frac{2\pi m}{T^2} \quad (21)$$

Przy założeniu, że sygnał modulujący jest obustronnie nieskończonym ciągiem słów kodowych kodu AMI, możemy sygnał zmodulowany zapisać w postaci sumy

$$x(t) \stackrel{\text{def}}{=} \sum_n g^{(n)}(t - nT) ; \quad n = 0, \pm 1, \pm 2, \dots, \quad (22)$$

w której $g^{(n)}(t)$ oznacza element przestrzeni G , będącej zbiorem funkcji czasu opisanych, na podstawie wzorów (12), (13), (14) następującymi zależnościami

$$g_{0^-}(t) = A \cos(\omega_0 t + \varphi_1^{0^-}) h(t) \quad (23)$$

$$g_{+1}(t) = A \cos(\omega_0 t + \varphi_1^{+1} + \frac{1}{2} kt^2) h(t) \quad (24)$$

$$g_{0^+}(t) = A \cos[(\omega_0 + kT)t + \varphi_1^{0^+}] h(t) \quad (25)$$

$$g_{-1}(t) = A \cos[(\omega_0 + kT)t + \varphi_1^{-1} - \frac{1}{2} kt^2] h(t) \quad (26)$$

gdzie:

$g_{0^-}(t)$ - impuls AASK odpowiadający zero w kodzie AMI, jeżeli ostatni niezerowy impuls w ciągu kodowym równy był -1 ,

$g_{+1}(t)$ - impuls AASK odpowiadający $+1$ w kodzie AMI,

$g_{0^+}(t)$ - impuls AASK odpowiadający zero w kodzie AMI, jeżeli ostatni niezerowy impuls w ciągu kodowym równy był $+1$,

$g_{-1}(t)$ - impuls AASK odpowiadający -1 w kodzie AMI,

$\varphi_1^{0^-}$ - faza początkowa impulsu $g_{0^-}(t)$, element zbioru

$$\Phi^{0^-} = \{\varphi_1^{0^-}, \varphi_2^{0^-}, \dots, \varphi_1^{0^-}, \dots\}$$

φ_1^{+1} - faza początkowa impulsu $g_{+1}(t)$, element zbioru

$$\Phi^{+1} = \{\varphi_1^{+1}, \varphi_2^{+1}, \dots, \varphi_1^{+1}, \dots\}$$

$\varphi_1^{0^+}$ - faza początkowa impulsu $g_{0^+}(t)$, element zbioru

$$\Phi^{0^+} = \{\varphi_1^{0^+}, \varphi_2^{0^+}, \dots, \varphi_1^{0^+}, \dots\}$$

φ_1^{-1} - faza początkowa sygnału $\xi_{-1}(t)$, element zbioru

$$\Phi^{-1} = \left\{ \varphi_1^{-1}, \varphi_2^{-1}, \dots, \varphi_1^{-1}, \dots \right\}$$

Funkcja $\cos(x)$ jest funkcją okresową o okresie 2π , a więc zachodzi następująca zależność

$$\cos(x + \varphi) = \cos(x + \bar{\varphi}) \quad (27)$$

gdzie

$$\bar{\varphi} \stackrel{\text{def}}{=} \varphi - 2\pi \varepsilon \left(\frac{\varphi}{2\pi} \right), \quad (28)$$

ε - oznacza symbol funkcji entier.

Wielkość $\bar{\varphi}$ będziemy nazywać wartością istotną fazy φ .

Na podstawie (27) i (28) wyrażenia (23) do (26) możemy bez zmiany ogólności zapisać jako

$$\xi_{0^-}(t) = A \cos(\omega_0 t + \bar{\varphi}_1^{0^-}) h(t) \quad (29)$$

$$\xi_{+1}(t) = A \cos(\omega_0 t + \bar{\varphi}_1^{+1} + \frac{1}{2} k t^2) h(t) \quad (30)$$

$$\xi_{0^+}(t) = A \cos[(\omega_0 + kT)t + \bar{\varphi}_1^{0^+}] h(t) \quad (31)$$

$$\xi_{-1}(t) = A \cos[(\omega_0 + kT)t + \bar{\varphi}_1^{-1} - \frac{1}{2} k t^2] h(t) \quad (32)$$

gdzie: $\bar{\varphi}_1^{0^-}, \bar{\varphi}_1^{+1}, \bar{\varphi}_1^{0^+}, \bar{\varphi}_1^{-1}$ oznaczają, odpowiednio, wartości istotne faz początkowych $\varphi_1^{0^-}, \varphi_1^{+1}, \varphi_1^{0^+}, \varphi_1^{-1}$. Wielkości $\bar{\varphi}_1^{0^-}, \bar{\varphi}_1^{+1}, \bar{\varphi}_1^{0^+}, \bar{\varphi}_1^{-1}$ są, odpowiednio, elementami zbiorów $\bar{\Phi}^{0^-}, \bar{\Phi}^{+1}, \bar{\Phi}^{0^+}, \bar{\Phi}^{-1}$.

Ze względu na problemy realizacji systemu z modulacją AASK, np. detekcja z zastosowaniem optymalnej reguły odbioru [4], szczególnie ważne jest wykazanie, jakie warunki muszą być spełnione, aby przestrzeń G była przestrzenią skończoną.

Z zależności (29) do (32) wynika, że jest ona skończona, jeżeli zbiory $\bar{\Phi}^{0^-}, \bar{\Phi}^{+1}, \bar{\Phi}^{0^+}, \bar{\Phi}^{-1}$ są zbiorami skończonymi. W systemie z modulacją AASK, w którym nie jest wymagany warunek ciągłości fazy przebiegu zmodulowanego, warunek skończoności zbiorów $\bar{\Phi}^{0^-}, \bar{\Phi}^{+1}, \bar{\Phi}^{0^+}, \bar{\Phi}^{-1}$ jest następujący

$$\bar{\Phi}^{0^-} = \bar{\Phi}^{+1} = \bar{\Phi}^{0^+} = \bar{\Phi}^{-1} = \left\{ \varphi \right\} \quad (33)$$

Znacznie bardziej skomplikowany jest przypadek, gdy w systemie wymagana jest ciągłość fazy w sygnale zmodulowanym. Ciągłość fazy oznacza, że faza początkowa impulsu $g^{(n)}$ /patrz wzór (29)/ jest równa fazie początkowej impulsu $g^{(n-1)}$ zwiększonej o przesunięcie fazowe wnoszone przez impuls $g^{(n-1)}$.

Niech funkcja $g_0(t)$ wprowadza przesunięcie fazowe φ_0 , wówczas

$$\omega_0 T = \varphi_0 = 2\pi n_1 + \psi_0; \quad 0 \leq \psi_0 < 2\pi; \quad n_1 = 0, 1, 2, \dots \quad (34)$$

Na mocy (21) funkcje $g_{+1}(t)$, $g_{0+}(t)$, $g_{-1}(t)$ wprowadzają, odpowiednio, przesunięcia fazowe φ_1 , φ_2 , φ_3 spełniające zależności

$$\varphi_1 = 2\pi n_1 + \pi m + \psi_0 = 2\pi \left(n_1 + \frac{m}{2} + \frac{\psi_0}{2\pi} \right) \quad (35)$$

$$\varphi_2 = 2\pi n_1 + 2\pi m + \psi_0 = 2\pi \left(n_1 + m + \frac{\psi_0}{2\pi} \right) \quad (36)$$

$$\varphi_3 = 2\pi n_1 + \pi m + \psi_0 = 2\pi \left(n_1 + \frac{m}{2} + \frac{\psi_0}{2\pi} \right) \quad (37)$$

Z zależności (34) do (37) wynika, że zbiory Φ^{0-} , Φ^{+1} , Φ^{0+} , Φ^{-1} , a więc i przestrzeń G są skończone, gdy dla konkretnego systemu współczynnik modulacji m i wielkość $\frac{\psi_0}{2\pi}$ są liczbami wymiernymi. Istnieje wówczas taka liczba naturalna n_{\max} , że po n krokach sumaryczne przesunięcie fazowe wnoszone przez n_{\max} impulsów AASK będzie całkowitą wielokrotnością 2π . Wartość istotna przesunięcia fazowego będzie wówczas równa zero.

Powyższe rozważania zostaną zilustrowane następującymi przykładami.

Przykład 1.

Na wejście modulatora AASK podano sygnał w kodzie AMI; współczynnik modulacji $m = 1$, pierwszy impuls w ciągu kodowym AMI jest równy 1, faza początkowa pierwszego impulsu wychodzącego z modulatora równa jest zero, czas trwania pojedynczego impulsu jest równy T , pulsacja początkowa ω_0 , a przesunięcie fazowe wnoszone przez impuls $g_0(t)$ jest równe $2\pi n$; $n = 1, 2, \dots$. Należy wyznaczyć przestrzeń impulsów wyjściowych z modulatora, tj. G .

Na podstawie zależności (21) wyznaczamy wartość bezwzględna przyspieszenia

$$k = \frac{2\pi m}{T^2} = \frac{2\pi}{T^2}$$

Z zależności (34) do (37) wyznaczamy przesunięcia fazowe $\varphi_0, \varphi_1, \varphi_2$

φ_3

$$\varphi_0 = 2\pi n$$

$$\varphi_1 = 2n\pi + \pi = (2n + 1)\pi$$

$$\varphi_2 = 2n\pi + 2\pi = 2(n + 1)\pi$$

$$\varphi_3 = 2n\pi + \pi = (2n + 1)\pi$$

Wartości istotne przesunięć fazowych φ_0 i φ_2 są równe zero, a φ_1 i φ_3 równe π ; $\bar{\varphi}_1 = \bar{\varphi}_3 = \pi$. Pojawienie się po sobie dwóch impulsów wyjściowych z modulatora odpowiadających $+1$ i -1 w kodzie ANI spowoduje, że wartość istotna sumarycznego przesunięcia fazowego wnoszonego przez te impulsy będzie równa zero. Pojawienie się między nimi dowolnej liczby impulsów odpowiadających zeru w kodzie ANI również nie spowoduje zmiany wartości istotnej sumarycznego przesunięcia fazowego. Zbiory $\bar{\Phi}^{0^-}$, $\bar{\Phi}^{+1}$, $\bar{\Phi}^{0^+}$, $\bar{\Phi}^{-1}$ są zatem zbiorami jednoelementowymi, równymi

$$\bar{\Phi}^{0^-} = \bar{\Phi}^{+1} = \{0\}$$

$$\bar{\Phi}^{0^+} = \bar{\Phi}^{-1} = \{\pi\}$$

Przestrzeń G impulsów wyjściowych z modulatora jest zatem równa

$$G = \{s_1(t), s_2(t), s_3(t), s_4(t)\}$$

gdzie:

$$s_1(t) = s_{0^-}(t) = A \cos(\omega_0 t) h(t)$$

$$s_2(t) = s_{+1}(t) = A \cos\left(\omega_0 t + \frac{\pi t^2}{T^2}\right) h(t)$$

$$s_3(t) = s_{0^+}(t) = A \cos\left[\left(\omega_0 + \frac{2\pi}{T}\right) t + \pi\right] h(t)$$

$$s_4(t) = s_{-1}(t) = A \cos\left[\left(\omega_0 + \frac{2\pi}{T}\right) t + \pi - \frac{\pi t^2}{T^2}\right] h(t)$$

Przykład 2.

Warunki i polecenia jak w przykładzie 1, ale współczynnik modulacji $m = \frac{2}{3}$.

Tak jak poprzednio wyznaczamy wartość bezwzględną przyspieszenia

$$k = \frac{4\pi}{3T^2}$$

Przesunięcia fazowe φ_0 , φ_1 , φ_2 , φ_{3m} są, odpowiednio, równe

$$\varphi_0 = 2n\pi$$

$$\varphi_1 = 2n\pi + \frac{2}{3}\pi$$

$$\varphi_2 = 2n\pi + \frac{4}{3}\pi$$

$$\varphi_3 = 2n\pi + \frac{2}{3}\pi$$

W tym przypadku jedynie wartość istotna przesunięcia fazowego φ_0 jest równa zero. Przeprowadzając rozumowanie podobne jak w przykładzie 1 wyznaczamy zbiory $\bar{\Phi}^{0^-}, \bar{\Phi}^{0^+}, \bar{\Phi}^{-1}$. Są one wszystkie równe, tj.

$$\bar{\Phi}^{0^-} = \bar{\Phi}^{0^+} = \bar{\Phi}^{-1} = \left\{0, \frac{2}{3}\pi, \frac{4}{3}\pi\right\}$$

Przestrzeń G jest natomiast przestrzenią dwumastocelową, a jej elementy możemy zapisać w następujący sposób

$$s_{0^-}^{(1)}(t) = A \cos\left(\omega_0 t + \bar{\varphi}^{(1)}\right) h(t)$$

$$s_{+1}^{(1)}(t) = A \cos\left(\omega_0 t + \frac{2\pi t^2}{3T^2} + \bar{\varphi}^{(1)}\right) h(t)$$

$$s_{0^+}^{(1)}(t) = A \cos\left[\left(\omega_0 + \frac{4\pi}{3T}\right)t + \bar{\varphi}^{(1)}\right] h(t)$$

$$s_{-1}^{(1)}(t) = A \cos\left[\left(\omega_0 + \frac{4\pi}{3T}\right)t - \frac{2\pi t^2}{3T^2} + \bar{\varphi}^{(1)}\right] h(t)$$

gdzie:

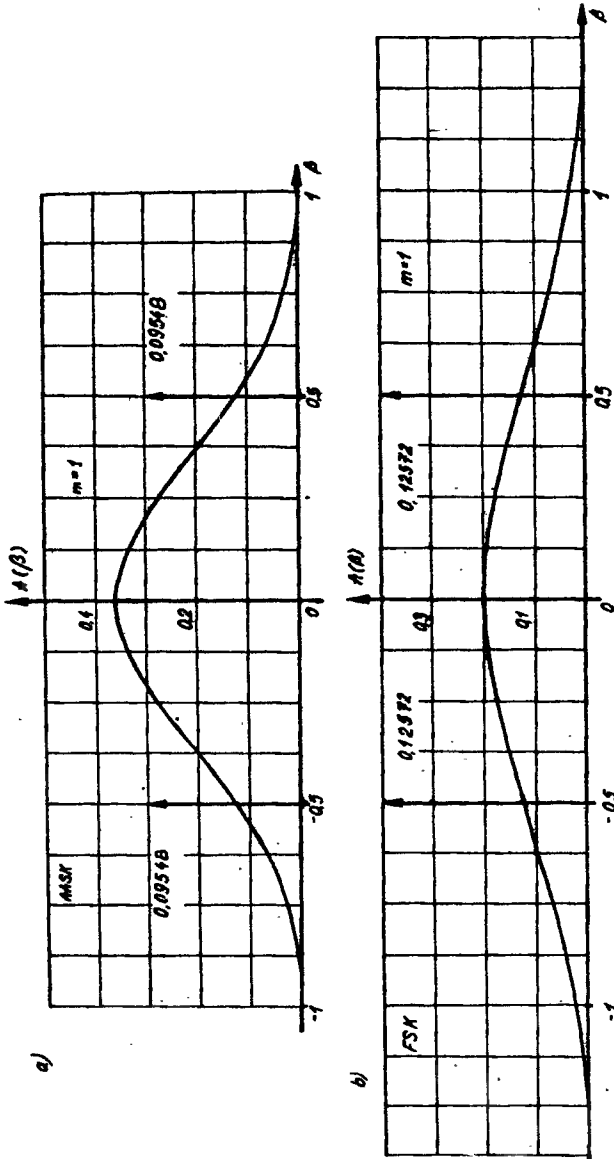
$$\bar{\varphi}^{(1)} \in \left\{\bar{\varphi}^{(1)}, \bar{\varphi}^{(2)}, \bar{\varphi}^{(3)}\right\} = \left\{0, \frac{2}{3}\pi, \frac{4}{3}\pi\right\}$$

Jeżeli spełniony jest warunek skłócenności przestrzeni G impulsów wyjściowych z modulatora, to możemy, w oparciu o metodę opisaną w pracy [3], wyznaczyć widmową gęstość mocy przebiegu AASK.

Do obliczeń przyjmujemy, że pierwotny ciąg binarny jest stacjonarym ciągiem losowym o wyrazach niezależnych. Na rys.2 przedstawiono wykresy znormalizowanej gęstości widmowej mocy $A(\beta)$

$$\beta = \left[\frac{\omega - 0,5(\omega_1 + \omega_0)}{2} \right] T \quad (38)$$

sygnałów AASK 1, dla porównania, binarnej FSK o takich samych współczynnikach modulacji m ; $m=1$. Prawdopodobieństwa stacjonarne wystąpienia jedynki



Rys.2. Znormalizowane systości widmowe mocy sygnałów
 a/ z modulacją ASK b/ z modulacją FSK
 /współczynniki modulacji m jest w obu przypadkach równy 1/

i zera w ciągu binarnym są sobie równe, tj. równe 0,5.

Uwagi końcowe

1. Niosący informację sygnał binarny musi zostać wstępnie zakodowany przy użyciu kodu o skończonej bieżącej sumie cyfrowej i dopiero takim sygnałem możemy dokonywać modulacji przyspieszenia kątownego nośnego sygnału harmonicznego. W wyniku kodowania uzyskuje się sygnał o redundancji równej 1, co umożliwia wykrywanie pojedynczych przekłamań w sygnale zdemodulowanym w czasie rzeczywistym.
2. Najkorzystniejszy przypadek zachodzi, gdy pierwotny sygnał binarny zostaje zakodowany przy użyciu kodu AMI, tj. kodu należącego do klasy kodów transmisyjnych o skończonej bieżącej sumie cyfrowej, charakteryzującego się najmniejszymi wahaniami tej sumy.
3. Z zależności (34) do (37) wynika interesująca własność sygnałów AASK. Przesunięcia fazowe wnoszone przez impulsy wychodzące z modulatora odpowiadające zeru oraz jedynce w pierwotnym sygnale binarnym, różnią się między sobą. Fakt ten można wykorzystać do demodulacji sygnału AASK z jednoczesnym dekodowaniem.
4. Przestrzeń impulsów wyjściowych modulatora jest przestrzenią skończoną, gdy współczynnik modulacji m jest liczbą wymierną oraz wartość istotna przesunięcia fazowego /patrz (28)/ wnoszonego przez impuls odpowiadający zeru w kodzie AMI / 0^+ lub 0^- / jest wymierną podwielokrotnością 2π .
5. Z rys.2 wynika, że pasmo potrzebne do przesłania sygnału AASK przy zakodowaniu pierwotnej informacji binarnej w kodzie AMI jest nieszersze niż pasmo potrzebne do przesłania sygnału binarnej PSK o takim samym współczynniku modulacji. Stanowi to, wraz z innymi, przytoczonymi własnościami modulacji AASK, o przydatności tego typu modulacji w systemach transmisyjnej informacji cyfrowych.

L i t e r a t u r a

- [1] Anderson R.R., Salz J.: Spectra of Digital FM. B.S.T.J., vol.XLIV, VII-VIII.1965, s.1165-1189
- [2] Lathi B.P.: Teoria sygnałów i układów telekomunikacyjnych. PWN, Warszawa 1970
- [3] Rydel M.: Analiza sygnałów kodowych. Prace naukowe - Elektronika, z.39, WFW, Warszawa 1979
- [4] Seidler J.: Optymalizacja adaptacyjnych systemów informacyjnych. PWN, Warszawa 1971

A SINUSOIDAL WAVE DIGITAL MODULATION METHOD

Summary

The article deals with an angular acceleration shift keying /AASK/. The detailed time analysis of the modulated signal is given. The article also includes a presentation of the power spectral density of an AASK signal.

О НЕКОТОРОМ СПОСОБЕ МОДУЛИЦИИ СИНОСОИДАЛЬНОЙ НЕСУЩЕЙ ЦИФРОВЫМ СИГНАЛОМ

Резюме

В статье представлен способ цифровой модуляции углового ускорения /AASK/. Проведен подробный анализ промодулированного сигнала в домене времени и представлены итоги расчетов спектральной плотности мощности сигнала AASK.

Lech J. Weiss
Tadeusz Wysocki

SPOSÓB CYFROWEGO POMIARU MAKSYMALNEJ RÓŻNICY CHWILOWYCH WARTOŚCI PRZEBIEGÓW ODKSZTAŁCONYCH

W artykule opisano dwa przypadki cyfrowego pomiaru wartości międzyszczytowej napięć odkształconych i to dla przebiegów wykazujących dwa ekstrema na okres oraz dla przebiegów wykazujących więcej niż dwa ekstrema na okres. Pomiar wykonywany jest metodą próbkowania.

1. WSTĘP

Dla przebiegu sinusoidalnie zmiennego zachodzą ściśle określone zależności pomiędzy wartością skuteczną, średnią, szczytową i międzyszczytową. Wynikają one z rodzaju funkcji opisującej przebieg. W takim przypadku wystarczy pomierzyć jedną z tych wartości, a pozostałe można obliczyć. Natomiast w przypadku przebiegów odkształconych, nie znając funkcji opisującej przebieg lecz jedynie kształt przebiegu wskazany na ekranie oscyloskopu, na podstawie pomiaru jednej z wymienionych wartości nie można wyznaczyć pozostałych. Dlatego zachodzi potrzeba mierzenia każdej z wartości charakteryzujących przebieg.

Przebieg czasowy wielkości fizycznej $a(t)$ można przetworzyć na odpowiadający mu przebieg napięcia $u(t)$. Przetó ograniczymy się do rozpatrywania tylko pomiaru napięć. Z punktu widzenia możliwości dokonywania pomiarów maksymalnej różnicy chwilowych wartości odkształconego przebiegu napięcia, nałożonego na składową stałą, wyróżnia się trzy charakterystyczne przypadki zilustrowane rys.1:

1. Na jeden okres zmian napięcia, tak jak dla przebiegu sinusoidalnie zmiennego, przypadają dwa ekstrema, a pomiędzy ekstremami funkcja narasta, względnie maleje monotonicznie /rys.1a/.
2. Na jeden okres przypadają więcej niż dwa ekstrema /rys.1b/ oraz istnieje przynajmniej jedna wartość napięcia U_0 , dla której spełnione są warunki

$$u(t_1) = u(t'_1) = U_0 \quad (1)$$

1

$$\text{sign} \frac{du}{dt} \Big|_{t=t_1} = \text{sign} \frac{du}{dt} \Big|_{t=t'_1} \neq 0 \quad (2)$$

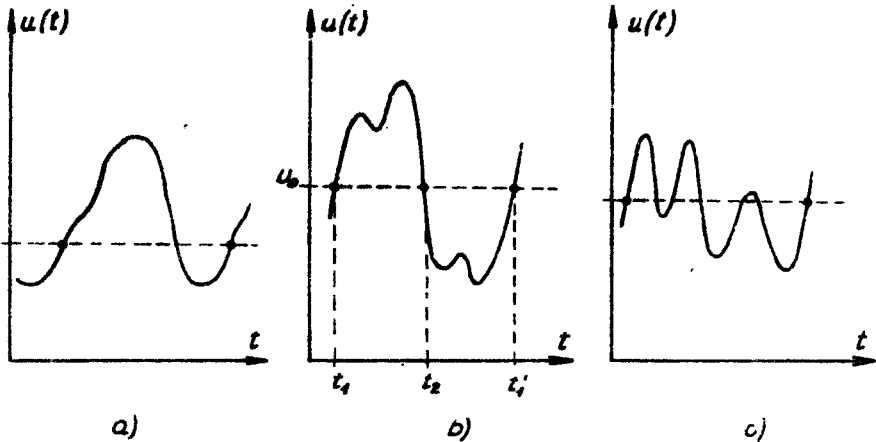
gdy różnica czasów $t_1' - t_2$ jest równa okresowi zmian, oraz tylko dla jednej wartości czasu t_2

$$t_1 < t_2 < t_1' \quad (3)$$

zachodzi zależność

$$u(t_2) = U_0 \quad (4)$$

3. Na jeden okres przypadają więcej niż dwa ekstrema, a nie istnieje taka wartość napięcia U_0 , dla której, spełnione byłyby warunki charakteryzujące przypadek 2.

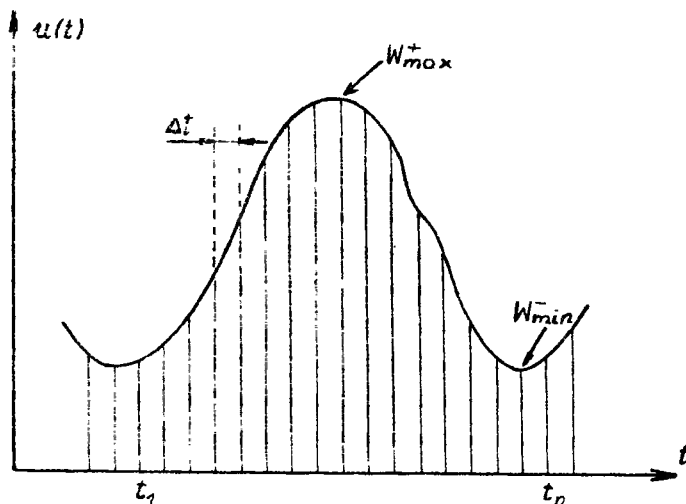


Rys.1. Trzy przypadki przebiegów odkształconych

Opisane w dalszym ciągu sposoby pomiaru maksymalnej różnicy chwilowych wartości dotyczą przypadków 1 i 2. Zakłada się przy tym, że widmo przebiegu $u(t)$ jest ograniczone w pasmie od 0 do f_{\max} .

2. POMIAR DLA PRZYPADKU PIERWSZEGO

Sposób pomiaru maksymalnej różnicy chwilowych wartości, gdy na jeden okres zmian napięcia przypadają tylko dwa ekstrema, ilustruje rys.2. Wartości chwilowe napięcia mierzy się metodą próbkowania, a wartości próbek są wyrażane w postaci cyfrowej. Cykl pomiarowy rozpoczyna się od próbki pobranej bezpośrednio po osiągnięciu minimalnej wartości napięcia, tj. w chwili t_1 /rys.2/. Cykl ten składa się z dwu części: w pierwszej części cyklu występuje narastanie mierzonego napięcia, a w drugiej opadanie. Oznaczmy wartości próbek pobranych podczas pierwszej części cyklu jako W^+ , a podczas drugiej jako W^- .



Rys.2. Ilustracja pierwszego przypadku pomiaru napięcia międzyszczytowego

Pobrane podczas pierwszej części cyklu pomiarowego próbki są kierowane do rejestru pierwszego. Próbką o wartości W_i^+ zostaje porównana z poprzednią próbką, której wartość W_{i-1}^+ została zapamiętana w układzie pomiarowym. Jeżeli zachodzi zależność

$$W_i^+ > W_{i-1}^+ \quad (5)$$

co ma miejsce dla narastającego przebiegu, następuje wyzerowanie rejestru pierwszego i wpisanie doń nowej wartości W_i^+ . Proces ten trwa tak długo, póki zachodzi nierówność (5). Jeżeli w wyniku porównywania wartości dwu kolejnych próbek okaże się, że nie zachodzi nierówność (5), tj. gdy

$$W_i^+ \leq W_{i-1}^+ \quad (6)$$

kończy się pierwsza część cyklu pomiarowego. W rejestrze pierwszym pozostaje zapisana wartość poprzedniej próbki, która jest największą wartością W_{max}^+ pomierzoną podczas pierwszej części cyklu, a wchodzący w skład miernika układ sterujący powoduje przełączenie i kierowanie próbek do drugiego rejestru.

W drugiej części cyklu pomiarowego próbka o wartości W_j^- zostaje porównana z poprzednią próbką o wartości W_{j-1}^- . Jeżeli zachodzi nierówność

$$W_j^- < W_{j-1}^- \quad (7)$$

następuje wyzerowanie rejestru drugiego i wpisanie doń nowej wartości równej W_j^- . Druga część cyklu pomiarowego kończy się na pierwszej kolejnej próbce, dla której już nie zachodzi nierówność (7). Na rys.2 jest to próbka pobrana w czasie t_n . W rejestrze drugim pozostaje zapisana wartość poprzedniej próbki, która jest najmniejszą wartością pomierzoną podczas dru-

giej części cyklu pomiarowego $W_{n-1}^- = W_{\min}^-$.

Po zakończeniu cyklu pomiarowego, zapisane w rejestrach wartości próbek W_{\max}^+ oraz W_{\min}^- zostają przepisane do układu wykonującego operację odejmowania, a wynik odejmowania

$$U_{pp} = W_{\max}^+ - W_{\min}^- \quad (8)$$

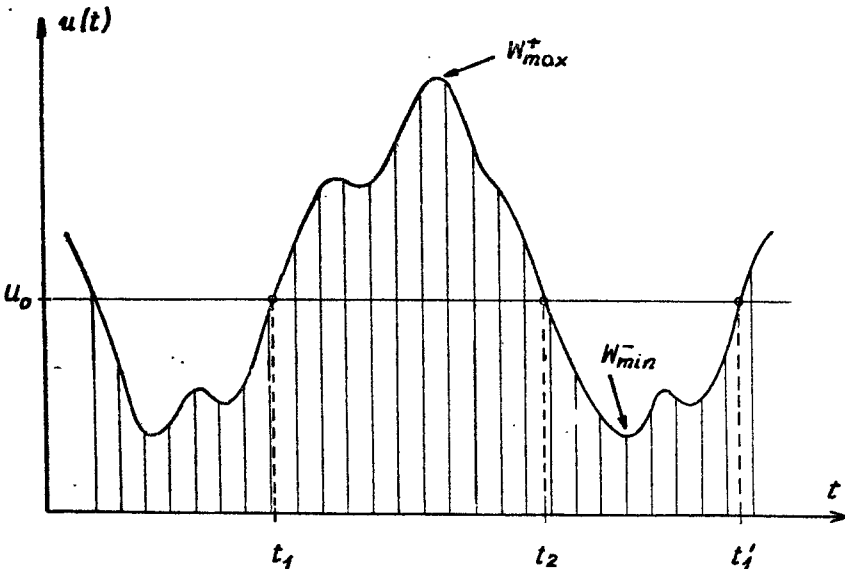
równy maksymalnej różnicy chwilowych wartości, zostaje wykazany na wskaźniku cyfrowym.

Po przepisaniu zawartości rejestrów do układu odejmującego, następuje wyzerowanie rejestru pierwszego i zapełnienie rejestru drugiego; rozpoczyna się następny cykl pomiarowy.

Opisany sposób pomiaru maksymalnej różnicy chwilowych wartości jest stosowany jedynie wtedy, gdy pomiędzy wartością minimalną a maksymalną napięcie monotonicznie narasta i monotonicznie maleje od wartości maksymalnej do występującej po niej wartości minimalnej. Nie może więc być zastosowany do napięć o przebiegu pokazanym na rys.1b.

3. POMIAR DLA PRZYPADKU DRUGIEGO

Gdy na jeden okres zmian napięcia przypadają więcej niż dwa ekstrema i spełnione są warunki charakteryzujące przypadek drugi, pomiar maksymalnej różnicy chwilowych wartości można przeprowadzić w sposób zilustrowany rys.3.



Rys.3. Ilustracja drugiego przypadku pomiaru napięcia międzyszczytowego

W tym przypadku okres jest również dzielony na dwie części, lecz nie tak jak to miało miejsce dla przypadku pierwszego. O podziale decyduje wartość napięcia odniesienia U_0 . W znacznej części przypadków może być ono równe średniej wartości mierzonego napięcia.

W pierwszej części cyklu pomiarowego, tj. w przedziale czasowym od t_1 do t_2 , próbki o wartościach W_k^+ są kierowane do rejestru pierwszego, a w drugiej części cyklu, tj. w przedziale od t_2 do t_1' , do rejestru drugiego. Podobnie jak dla pierwszego sposobu, każda próbka zostaje porównana z poprzednią. Jeśli w pierwszej części cyklu wartość bieżącej próbki W_k^+ jest większa od wartości zapisanej w rejestrze pierwszym W_{rej}^+

$$W_k^+ > W_{rej}^+ \quad (9)$$

następuje wyzerowanie rejestru pierwszego i wpisanie doń nowej wartości równej W_k^+ . Jeżeli nierówność (9) nie zachodzi, stan rejestru nie ulega zmianie. W ten sposób w chwili t_2 , kończącej pierwszą część cyklu pomiarowego, w rejestrze pierwszym zapisana jest największa wartość próbki W_{max}^+ . Jeżeli w drugiej części cyklu pomiarowego wartość bieżącej próbki W_{min}^- jest mniejsza od wartości zapisanej w rejestrze drugim W_{rej}^-

$$W_m^- < W_{rej}^- \quad (10)$$

następuje wyzerowanie rejestru drugiego i wpisanie doń nowej wartości W_m^- . W przeciwnym przypadku stan rejestru drugiego nie ulega zmianie. Dzięki temu w chwili t_1' jest w rejestrze drugim zapisana najmniejsza wartość W_{min}^- , jaka wystąpiła podczas drugiej części cyklu pomiarowego.

Po zakończeniu drugiej części cyklu pomiarowego następuje przepisanie zawartości obu rejestrów do układu wykonującej operację odejmowania i wyświetlenie wyniku odejmowania równego maksymalnej różnicy chwilowych wartości napięcia. Po przepisaniu zawartości rejestrów następuje wyzerowanie rejestru pierwszego i zapełnienie rejestru drugiego, po czym rozpoczyna się następny cykl pomiarowy.

4. UWAGI KOŃCOWE

Chociaż przy opisie obu sposobów pomiaru maksymalnej różnicy chwilowych wartości napięcia przebiegi napięcia traktowano jako przebiegi okresowe, to jednak sposoby te można stosować również i dla przebiegów prawie okresowych, a nawet dla pojedynczych, nie powtarzalnych przebiegów.

W obu przypadkach rozpatrywano przebiegi $u(t)$ zawierające składową stałą. Postąpiono tak, by uzyskać sposób stosowalny również dla przebiegów wolnozmiennych, w przypadku których mogą wystąpić trudności przy oddzieleniu składowej stałej z jednoczesnym wymaganiem zachowania kształtu składowej zmiennej. Gdy przebieg mierzony nie zawiera składowej stałej, można

albo dodać składową stałą, albo osobno w opisany sposób pomierzyć wartości szczytowe: dodatnią oraz ujemną i do wartości szczytowej dodatniej dodać moduł wartości szczytowej ujemnej.

Jeżeli sposób pomiaru, przewidziany dla przypadku pierwszego, zostanie zastosowany dla przebiegów przedstawionych na rys.1c, uzyskany wynik pomiaru będzie równy różnicy dwu kolejno po sobie występujących ekstremalnych wartości przebiegu.

Aparatura działająca w oparciu o wyżej opisany sposób pomiaru pozwala na wyeliminowanie pracochłonnych czynności związanych z wyznaczaniem wartości wielkości mierzonej metodami dotychczas stosowanymi.

Przedstawiony sposób pomiaru maksymalnej różnicy chwilowych wartości napięcia, opracowany w ramach prac własnych prowadzonych w Zakładzie Podstaw Elektroniki Instytutu Telekomunikacji i Elektrotechniki ATR w Bydgoszczy, stanowi przedmiot zgłoszeń patentowych.

L i t e r a t u r a

- [1] Budkowski St., Papliński A., Sosnowski J.: Zespoły i urządzenia cyfrowe. WNT, Warszawa 1979

DIGITAL PEAK-TO-VALLEY VALUE MEASUREMENT METHOD FOR NON - SINUSOIDAL VOLTAGE

Summary

The subject of this paper pertains to the problem of digital measurement of the peak-to-valley value of non-sinusoidal voltage. There are two cases discussed-instantaneous value of the measured voltage reaches 1/two extrema over the whole period, 2/ more than two extrema. The peak-to-valley value is measured by a sample-by-sample process.

МЕТОД ЦИФРОВОГО ИЗМЕРЕНИЯ МАКСИМАЛЬНОЙ РАЗНОСТИ МГНОВЕННЫХ ЗНАЧЕНИЙ НЕ- СИНУСОИДАЛЬНЫХ НАПРЯЖЕНИЙ

Резюме

Проблема работы касается цифрового измерения максимальной разности мгновенных значений несинусоидальных напряжений. Рассмотрены два случая: кривая напряжения имеет: 1/ два экстремальные значения во время одного периода, 2/ больше экстремальных значений. Измерение проводится с использованием квантования по времени.

Andrzej Pawlak
Jan Piątkowski
Roman Wiatr

METODYKA DOBORU SILNIKA STEMPLOWNICZY STATYCZNO - DYNAMICZNEJ

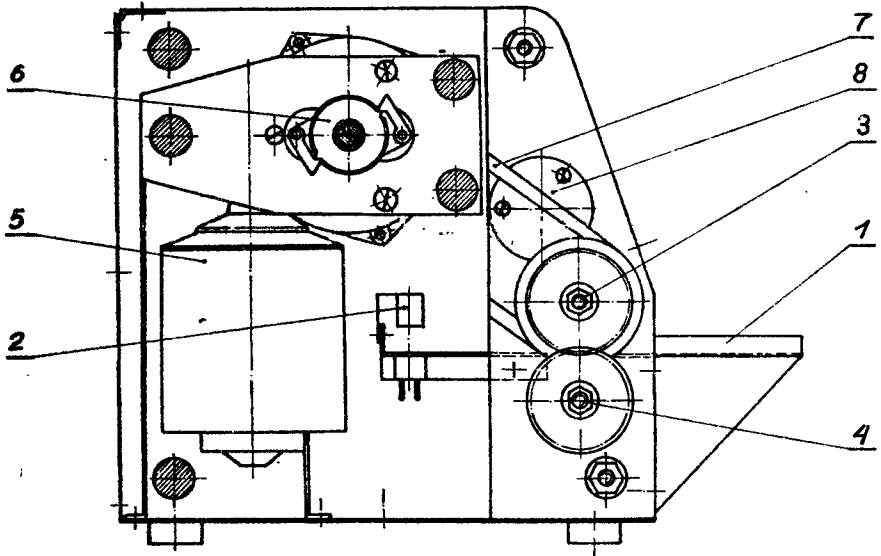
W artykule przedstawiono podstawy teoretyczne i metodykę doboru silnika stemplownicy druków manipulacyjnych na podstawie momentu obrotowego, niezbędnego do wykonania odcisku datowników.

1. WSTĘP

Wzrost zapotrzebowania na usługi pocztowe stworzył realne przesłanki dla masowego wypierania pracy ręcznej, tak bardzo jeszcze rozpowszechnionej w technologicznych procesach obróbki poczty. Jednym z kierunków mechanizacji poczty jest tzw. "mała mechanizacja", czyli mechanizacja obsługi klienta. Specyfika usług pocztowych zmusza do poszukiwania nowych, oryginalnych rozwiązań konstrukcyjnych. Uwzględniając istniejące tendencje rozwojowe, w Zakładzie Mechanizacji Poczty Instytutu Telekomunikacji i Elektrotechniki Akademii Techniczno-Rolniczej w Bydgoszczy opracowano stemplownicę druków manipulacyjnych SDM-1.

Działanie urządzenia SDM-1 przedstawionego na rys.1 polega na tym, że wałek datownika 3 z osadzonymi na nim trzema datownikami wykonując jeden obrót nanosi odcisk stempla datownika na druk manipulacyjny nadając mu jednocześnie ruch posuwisty powodujący usunięcie go ze strefy działania datownika na prowadnicę 1. Ruch obrotowy zostaje przekazany na wałek datujący 3 za pośrednictwem układu napędowego składającego się z przekładni pasowej 7, sprzęgła zapadkowego 6 i silnika elektrycznego 5. Układ napędowy zostaje uruchomiony sygnałem z fotokomórki 2. Synchronizację ruchu druku manipulacyjnego osiągnięto dzięki zastosowaniu przekładni zębatej przenoszącej ruch obrotowy na wałek stemplujący 4. Barwnik zostaje naniesiony na datowniki wałka datującego 3 każdorazowo przed dokonaniem operacji stemplowania dzięki zastosowaniu swobodnie zawieszzonego bezpośrednio nad datownikami wałka barwiącego 8.

W dotychczas stosowanych stemplownicach wykorzystywano bądź stemplowanie statyczne, jak na przykład w maszynach do frankowania korespondencji, bądź stemplowanie dynamiczne jak w stemplownicach kolumnowych i stolowych.



Rys.1. Stemplownica druków manipulacyjnych SDM-1, widok z boku

Niska jakość papieru stosowanego na blankiety druków manipulacyjnych zmusiła konstruktorów do zastosowania stemplowania statyczno-dynamicznego. Metoda ta stawia określone wymagania układowi napędowemu. W związku z tym zrodziła się potrzeba opracowania metodyki doboru silnika stemplownicy statyczno-dynamicznej. Parametrem określanym za pomocą tej metody jest moment obrotowy silnika konieczny do rozpędzenia wirujących części mechanizmu stemplownicy druków manipulacyjnych.

2. DYNAMIKA MECHANIZMU STEPLUJĄCEGO

W ogólnym przypadku, gdy mamy do czynienia z nieswobodnym ciałem sztywnym mogącym się obracać wokół nieruchomej osi i gdy jako współrzędną uogólnioną obierzymy kąt obrotu φ , wówczas praca przygotowana sił zewnętrznych wynosi [1]:

$$\delta L = \delta \varphi \sum_{i=1}^{i=n} M_{iz} \quad (1)$$

przy czym M_{iz} oznacza sumę momentów względem osi obrotu sił zewnętrznych działających na rozpatrywane ciało. Stąd otrzymujemy:

$$Q \varphi = \sum_{i=1}^{i=n} M_{iz} \quad (2)$$

Z drugiej strony dla ruchu układu materialnego o więzach idealnych i holonomicznych można zapisać równania różniczkowe Lagrange'a:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} = Q_j \quad (3)$$

gdzie:

T - energia kinetyczna rozpatrywanego układu

Równań tego typu można zapisać tyle, ile jest współrzędnych uogólnionych, a więc tyle ile rozpatrywany układ materialny ma stopni swobody.

Rozpatrywany konkretny układ materialny /węzeł datownika maszyny stemplującej/ jest układem o jednym stopniu swobody. Parametrem uogólnionym określającym ten stopień swobody jest kąt obrotu φ układu względem osi obrotu. Dlatego też dla opisu ruchu mechanizmu stemplującego zastosujemy równania Lagrange'a w postaci (3). Biorąc pod uwagę, że współrzędną uogólnioną jest kąt obrotu φ mamy:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{\varphi}} \right) - \frac{\partial T}{\partial \varphi} = Q_\varphi \quad (4)$$

3. OKREŚLENIE MOMENTU OBROTOWEGO SILNIKA

Mechanizm stemplujący można przedstawić schematycznie w następujący sposób: dla mechanizmu stemplującego, przedstawionego na rys.2, zachodzą następujące zależności kinematyczne:

$$\begin{aligned} \omega_1 R_1 &= \omega_2 R_2 = \omega_3 R_3 \\ \omega_2 &= \omega_1 \frac{R_1}{R_2} ; \quad \omega_3 = \omega_1 \frac{R_1}{R_3} \end{aligned} \quad (5)$$

Natomiast równanie energii kinetycznej rozpatrywanego układu możemy zapisać w postaci

$$T = \frac{J_{01}}{2} \omega_1^2 + \frac{J_{02}}{2} \omega_2^2 + \frac{J_{03}}{2} \omega_3^2 \quad (6)$$

gdzie:

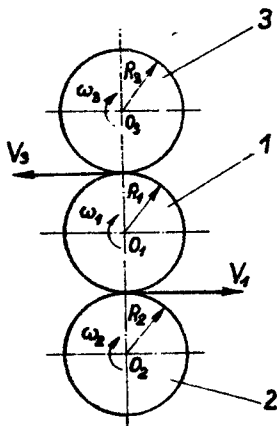
- J_{01} - moment bezwładności wałka datownika
- J_{02} - moment bezwładności rolki stemplującej
- J_{03} - moment bezwładności rolki barwiącej

Wprowadzając zależności (5) do równania (6) otrzymamy postać:

$$T = \frac{1}{2} \left[J_{O1} + J_{O2} \left(\frac{R_1}{R_2} \right)^2 + J_{O3} \left(\frac{R_1}{R_3} \right)^2 \right] \omega_1^2 \quad (7)$$

którą możemy zapisać

$$T = \frac{1}{2} \left[J_{O1} + J_{O2} \left(\frac{R_1}{R_2} \right)^2 + J_{O3} \left(\frac{R_1}{R_3} \right)^2 \right] \left(\frac{d\varphi}{dt} \right)^2 \quad (8)$$



Rys.2. Schemat mechanizmu stemplującego

- 1 - wzeł wałka datownika
- 2 - wzeł rolki stemplującej
- 3 - wzeł rolki barwiącej

Przechodząc do wielkości uogólnionych; tzn.

$$J_0 = J_{O1} + J_{O2} \left(\frac{R_1}{R_2} \right)^2 + J_{O3} \left(\frac{R_1}{R_3} \right)^2 \quad (9)$$

gdzie:

J_0 - uogólniony moment bezwładności mechanizmu stemplującego względem uogólnionej osi obrotu, za którą przyjmujemy oś obrotu wałka datownika,

równanie energii kinetycznej mechanizmu stemplującego możemy zapisać w postaci uogólnionej, a mianowicie:

$$T = \frac{J_0}{2} \left(\frac{d\varphi}{dt} \right)^2 \quad (10)$$

Określimy teraz składowe równania Lagrange'a (4) pamiętając, że przy wyznaczaniu pochodnych cząstkowych $\frac{\partial T}{\partial \dot{\varphi}}$ i $\frac{\partial T}{\partial \varphi}$ zmienne φ i $\dot{\varphi}$ należy traktować jako niezależne od siebie.

$$\frac{\partial T}{\partial \varphi} = 0 \quad - \text{ponieważ w równaniu (10) występuje zmienna } \dot{\varphi} \text{ a nie } \varphi$$

$$\frac{\partial T}{\partial \dot{\varphi}} = J_0 \dot{\varphi} = J_0 \frac{d\varphi}{dt}$$

i dalej

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{\varphi}} \right) = J_0 \frac{d^2 \varphi}{dt^2}$$

W związku z tym równanie (4) przyjmuje postać:

$$J_0 \frac{d^2 \varphi}{dt^2} = Q \omega \quad (11)$$

Podstawiając 11 do 2 otrzymamy równanie:

$$J_0 \frac{d^2 \varphi}{dt^2} = \sum_{i=1}^{i=n} M_{iz} \quad (12)$$

w którym

$$\sum_{i=1}^{i=n} M_{iz} = M_{O1} + M_{O2} \quad (13)$$

gdzie:

M_{O1} - moment obrotowy, niezbędny do pokonania bezwładności mechanizmu stemplującego

M_{O2} - moment oporu obrotowego, określany stratami na tarcie w łożyskach oraz stratami energii na naniesienie barwnika na datowniki i odcisnięcie stempla [2].

W związku z tym

$$J_0 \frac{d^2 \varphi}{dt^2} = M_{O1} + M_{O2}$$

lub

$$J_0 \frac{d\omega}{dt} = M_{O1} + M_{O2} \quad (14)$$

Rozdzielając zmienne i całkując lewą stronę równania od 0 do , a prawą od 0 do t_r (t_r - czas rozruchu silnika), otrzymamy:

$$\int_0^{\omega} J_0 d\omega = \int_0^{t_r} (M_{O1} + M_{O2}) dt$$

i rozwiązując względem $(M_{O1} + M_{O2})$ mamy:

$$M_{O1} + M_{O2} = \frac{J_0 \omega}{t_r} \quad (15)$$

Przyjmując, że moment obrotowy podawany do mechanizmu stemplującego musi pokonać zarówno bezwładność mechanizmu stemplującego jak i opór obrotowy, możemy zapisać, że

$$M_0 = M_{O1} + M_{O2}$$

gdzie:

M_0 - moment obrotowy niezbędny do rozpędzenia mechanizmu stemplującego

i dalej

$$M_0 = \frac{J_0 \omega}{t_r} \quad (16)$$

Uwzględniając fakt, że ruch obrotowy do mechanizmu stemplującego zostaje podany z pomocą układu napędowego, przy doborze silnika należy uwzględnić również straty na rozpędzenie tego układu. W pocztowych maszynach stemplujących straty te uwzględnia się wprowadzając współczynnik zapasu w granicach 10 + 15 % [2]. W związku z tym moment obrotowy silnika określa się wzorem:

$$M_{sil} = k M_0 \quad (17)$$

lub w postaci rozwiniętej:

$$M_{sil} = k \frac{J_0 \omega}{t_r} \text{ kGm} \quad (18)$$

gdzie:

$k = 1,1 + 1,15$ - bezwymiarowy współczynnik zapasu

J_0 - kGm s²

ω - $\frac{1}{s}$

t_r - s

4. WNIOSKI

Zaproponowana metodyka doboru silnika stemplownicy statyczno-dynamicznej, z punktu widzenia jego momentu obrotowego, stwarza możliwość wstępnego określenia typu silnika już we wczesnej fazie konstruowania.

Literatura

- [1] Leyko J.: Mechanika ogólna. T.2. PWN, Warszawa 1974
- [2] Sokołow W.P.: Pocztoobrabatywajuszczije maszyny i awtomaty. Swiaź, Moskwa 1977

METHODS OF ENGINE SELECTION FOR STATIC-DYNAMIC STAMPING MACHINE

Summary

On the basis of the turning moment necessary for performing a date stamp on a form, there are presented in the paper theoretical bases and methods of engine selection for a stamping machine used for manipulation prints.

МЕТОДИКА РАСЧЁТА ЭЛЕКТРОДВИГАТЕЛЯ СТАТИЧЕСКИ-ДИНАМИЧЕСКОЙ ШТЕМПЕЛЕВАЛЬНОЙ МАШИНЫ

Резюме

В статье представлены теоретические основы и методика расчёта крутящего момента электродвигателя штемпелевальной машины применяемой для штемпелевания бланков денежных переводов.

Cena zł 84,-