

Uniwersytet Marii Curie-Skłodowskiej

Wydział Matematyki, Fizyki i Informatyki

Paweł Górny

Numeryczne rozwiązywanie
Markowskich modeli sieci
z użyciem rozkładu LU

Praca magisterska

Praca magisterska wykonana pod kierunkiem

prof. dr hab. Jerzego Kozickiego

Składam podziękowania za pomoc w napisaniu pracy

Państwu mgr Beacie i Jarosławowi Bylinom

Lublin 2005

Spis treści

1	Modele kolejkowe	4
1.1	Podstawowe pojęcia	4
1.2	Model typu $M(n)/M(n)/1$	6
1.3	Markowski model otwartej sieci stanowisk	8
2	Łańcuchy Markowa	13
2.1	Podstawowe pojęcia	13
3	Metody rozwiązywania układów równań liniowych $Ax = b$	27
3.1	Eliminacja Gaussa	29
3.2	Rozkład LU	31
3.3	Metody dokładne i łańcuchy Markowa	33
4	Zastosowanie LU do sieci optycznej	36
4.1	Sieci optyczne	36
4.2	Przełącznik optyczny	39
4.2.1	Model pakietu	42
4.3	Eksperyment numeryczny	48
5	Wnioski	54
6	Literatura	56

Wstęp

Modelowanie sieci komputerowych i komunikacyjnych jest przedmiotem zainteresowania wielu zespołów badawczych. Złożoność współczesnych sieci rodzi potrzebę stosowania przy ich projektowaniu modeli, pozwalających przewidzieć skutki podejmowanych decyzji.

W pracy opieram się na modelach kolejkowych w postaci sieci stanowisk obsługi. Stanowiska to zasoby badanego systemu, a obsługiwani klienci to wykonywane przez system zadania lub przesyłane przez sieć pakiety. Wszyscy klienci mają te same właściwości, albo dzielą się na klasy. Sieci mogą być otwarte - zasilane klientami z zewnątrz i odprowadzające ich po zakończeniu obsługi, lub zamknięte - z krążącymi w nich klientami.

Przedstawiam matematyczne metody rozwiązywania modeli w postaci sieci stanowisk oraz pokazuję przykłady ich zastosowania. Ponieważ założyłem, że czasy obsługi oraz odstępy czasów między zgłoszeniami mają rozkłady wykładnicze lub są złożone z faz o rozkładzie wykładniczym, w opisie systemu zastosowałem łańcuchy Markowa. Rozwiązanie takich modeli sprowadza się do problemów algorytmicznych i numerycznych - w swojej pracy używam metody rozkładu LU a następnie metody obliczania układów równań „zerowy wyznacznik”.

Na wstępie przybliżam podstawowe pojęcia dotyczące teorii kolejek oraz rachunku prawdopodobieństwa, których znajomość jest konieczna w dalszej części pracy. W pierwszym rozdziale zapoznaję czytelnika z wstępem do teorii kolejek.

Pokazuję klasyfikację systemów kolejkowych oraz przedstawiam ich przykłady.

Rozdział drugi to przedstawienie pojęć związanych z rachunkiem prawdopodobieństwa oraz budową łańcuchów Markowa. Na przykładzie gry w tenis i jej analizy pokazuję jak przebiega proces tworzenia łańcucha.

W trzecim rozdziale opisuję metody rozwiązywania układów liniowych $Ax = b$, przedstawiam w formie teoretycznej i praktycznej eliminację Gaussa oraz rozkład LU . Pokazuję również jak połączyć te metody numeryczne z problemem łańcuchów Markowa.

Rozdział czwarty składa się z informacji o sieciach optycznych oraz o przełączniku optycznym. Pokazuję na przykładzie w jaki sposób buduje się model kolejkowy dla prostego przypadku, a następnie przedstawiam czas i błąd obliczeń dla modelu bardziej skomplikowanego.

Korzystałem głównie z pracy Williama Stewarta „Introduction to the Numerical Solution of Markov Chains” z której zaczerpnąłem informacje o łączeniu łańcuchów Markowa z równaniami liniowymi oraz o rozwiązywaniu tych równań. Drugą istotną pozycją była książka Tadeusza Czachórskiego „Modelowanie kolejkowe w ocenie efektywności sieci i systemów komputerowych”, w oparciu o którą przedstawiam wstęp do teorii łańcuchów Markowa.

Rozdział 1

Modele kolejkowe

1.1 Podstawowe pojęcia

Każdy proces budowy modelu kolejkowego określonego systemu należy rozpocząć od jego obserwacji. Przez T określmy czas, przez jaki obserwowaliśmy interesujący nas system. Należy również zanotować liczbę zadań A , które nadeszły do systemu i liczbę zadań wykonanych C w czasie obserwacji. Pozwoli to znaleźć intensywność zgłoszeń $\lambda = \frac{A}{T}$ oraz przepustowość $X = \frac{C}{T}$ systemu. Możemy również poznać *czas aktywności* S określonego zasobu w czasie T i obliczyć *wykorzystanie zasobu* $u = \frac{S}{T}$ i *średni czas obsługi* zadania w danym zasobie $\bar{B} = \frac{S}{C}$. System jest w *stanie równowagi*, jeżeli średni czas obsługi jest mniejszy niż średni czas pomiędzy momentami nadejścia zadań. W innym przypadku system nie będzie w stanie uporać się z przychodzącymi zadaniami i kolejka zadań czekających będzie rosła. W stanie równowagi jest $\lambda = X$. Zarazem:

$$U = \frac{S}{T} = \frac{S}{C} \frac{C}{T} = \bar{B}X, \quad (1.1)$$

Wykorzystanie stanowiska jest iloczynem przepustowości i średniego czasu obsługi. Przez \bar{R} oznaczmy średni czas pomiędzy nadejściem zgłoszenia a zakończe-

niem jego obsługi (czas pobytu w systemie). Jeżeli w czasie T w systemie było \bar{N} zadań, to poprzez $\bar{N}T$ oznaczamy czas pobytu wszystkich zadań. Jeśli przez C nazwiemy ilość zadań ukończonych, z których każde spędziło w systemie \bar{R} , $\bar{N}T = C\bar{R}$, to

$$\bar{N} = \frac{C}{T}\bar{R} = X\bar{R} \quad (1.2)$$

Powyższe równanie nazywane jest prawem Little'a. Związek pomiędzy czasem pobytu, przepustowością i średnią liczbą zadań odnosi się do całego systemu lub jego podsystemu. W przypadku całego systemu, średnia liczba zadań jest sumą średnich odpowiednich podsystemów. Odnosząc prawo Little'a do kanału obsługi i pomijając kolejkę czekającą na obsługę, otrzymamy zależność:

$$\bar{N} = \frac{1 \times S + 0 \times (T - S)}{T} = \frac{S}{T} = U \quad (1.3)$$

a, czas reakcji to średni czas obsługi.

Zazwyczaj będziemy przyjmować, że klienci odwiedzają stanowiska z różną częstotliwością, a przepustowość różnych elementów systemu jest różna. Oznaczmy przez V_i średnią liczbę wizyt zadania w stacji i podczas jego pobytu w systemie. Łączny średni czas obsługi \bar{D}_i jednego zadania w stanowisku i wynosi:

$$\bar{D}_i = V_i\bar{B}_i \quad (1.4)$$

Przyjmijmy, że czas określający nadejścia klientów do stanowisk jest zmienną losową o rozkładzie prawdopodobieństwa $A(x)$, a czas obsługi jest zmienną losową o rozkładzie $B(x)$. Zazwyczaj jest tak w rzeczywistości - momenty aktywizacji zadań są przypadkowe, a czasy obsługi zależne w dużej mierze np.: od danych wejściowych. Określając typy stanowisk obsługi korzysta się z notacji Kendalla: system opisywany jest ciągiem liter $A/B/c/K/H/Z$, gdzie litery na miejscach A i B oznaczają typy rozkładów przyścia klientów i obsługi (M oznacza rozkład wykładniczy, D jednopunktowy, G - rozkład dowolny). c oznacza liczbę równoległych, pracujących jednocześnie kanałów obsługi, a zarazem maksymalną liczbę

klientów, których stanowisko może jednocześnie obsługiwać. K to ograniczenie długości kolejki - jeśli w kolejce jest K zadań, kolejne przychodzące są ignorowane. H określa wymiar źródła - liczebność populacji klientów, którzy mogą zażądać obsługi. Z jest regulaminem szeregowania zgłoszeń (*LIFO*, *FIFO*, *IS*, *PS*). Jeśli wymiar źródła jest nieskończony, brak ograniczeń długości kolejki, a regulamin jest normalny - można w zapisie pominąć K , H lub Z .

1.2 Model typu $M(n)/M(n)/1$

Założmy, że do stanowiska obsługi przychodzą klienci w odstępach czasu, które są określone przez zmienne losowe o rozkładzie wykładniczym, λ tego układu zależy od n - ilości klientów w stanowisku. Proces w którym zdarzenia następują po sobie w odstępach czasu określonych rozkładem wykładniczym nazywamy procesem Poissona. Przyjmijmy również, że rozkład czasu obsługi jest wykładniczy z parametrem μ będącym funkcją liczby klientów obecnych w stanowisku. Jeśli proces wejściowy stanowiska obsługi jest poissonowski, a czasu obsługi mają rozkład wykładniczy, to proces $N(t)$, którego wartości odpowiadają licznie klientów obecnych w stanowisku jest łańcuchem Markowa.

Niech $p(n, t; n_0) = P[N(t) = n | N(0) = n_0]$. Jeśli w chwili $t + \Delta t$ nie ma klientów w systemie, to oznacza, że od chwili t zaszła jedna z możliwości:

1. nikogo nie było i nikt w czasie Δt nie przyszedł ; prawdopodobieństwo nadejścia klienta do pustego systemu w czasie Δt to $\lambda(0)\Delta t$, prawdopodobieństwo, że nikt w tym czasie nie przyjdzie to $1 - \lambda(0)\Delta t$;
2. w chwili t był w systemie klient i jego obsługa zakończyła się, prawdopodobieństwo zakończenia obsługi w czasie Δt , jeżeli w systemie jest jeden klient, to $\mu(1)\Delta t$.

Zarazem, jeśli w chwili $t + \Delta t$ jest w systemie $n \geq 1$ klientów, to znaczy, że od chwili t zaszła jedna z sytuacji:

1. był o n klientów, nikt nie przyszedł $(1 - \lambda(n)\Delta t)$ ani nie odszedł $(1 - \mu(n)\Delta t)$; lub:
2. był o $n + 1$ klientów i jeden z nich odszedł $(\mu(n + 1)\Delta t)$; lub:
3. był o $n - 1$ klientów i przyszedł, z prawdopodobieństwem $\lambda(n - 1)\Delta t$ następny klient.

Zapiszmy zatem:

$$p(0, t + \Delta t) = p(0, t)[1 - \lambda(0)\Delta t] + p(1, t)\mu(1)\Delta t,$$

$$p(n, t + \Delta t) = p(n, t)[1 - \lambda(n)\Delta t][1 - \mu(n)\Delta t] +$$

$$p(n + 1, t)\mu(n + 1)\Delta t + p(n - 1, t)\lambda(n - 1)\Delta t, \quad n \geq 1,$$

Z powyższego wynika:

$$\frac{dp(0, t)}{dt} = -p(0, t)\lambda(0) + p(1, t)\mu(1),$$

$$\begin{aligned} \frac{dp(n, t)}{dt} = & -p(n, t)[\lambda(n) + \mu(n)] + p(n + 1, t)\mu(n + 1) + \\ & + p(n - 1, t)\lambda(n - 1), \quad n \geq 1 \end{aligned}$$

Korzystając z zapisu macierzowego, macierze \mathbf{Q}^T, \mathbf{P} są postaci:

$$\mathbf{Q}^T = \begin{bmatrix} -\lambda(0) & \mu(0) & 0 & 0 & 0 & 0 & \dots \\ \lambda(0) & -\mu(1) - \lambda(1) & \mu(2) & 0 & 0 & 0 & \dots \\ 0 & \lambda(1) & -\mu(2) - \lambda(2) & \mu(3) & 0 & 0 & \dots \\ 0 & 0 & \lambda(2) & -\mu(3) - \lambda(3) & \mu(4) & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

$$\mathbf{P}^T = \begin{bmatrix} p(0, t) & p(1, t) & p(2, t) & p(3, t) & \dots \end{bmatrix}$$

W stanie ustalonym $\lim_{t \rightarrow \infty} p(n, t) = p(n)$, $\frac{dp(n, t)}{dt} = 0$ i równania różniczkowe stają się równaniami:

$$0 = -p(0)\lambda(0) + p(1)\mu(1),$$

$$0 = -p(n)[\lambda(n) + \mu(n)] + p(n+1)\mu(n+1) + p(n-1)\lambda(n-1), n \geq 0$$

Ze względu na to, że stan systemu określamy w zależności od ilości obecnych w systemie klientów, stan w którym w systemie znajduje się n klientów będziemy nazywać stanem n . Powyższe równania przedstawiają bilans strumieni prawdopodobieństw: w ustalonym stanie strumień prawdopodobieństwa wejścia systemu w stan n : $p(n+1)\mu(n+1) + p(n-1)\lambda(n-1)$ jest równy strumieniowi prawdopodobieństwa wyjścia z tego stanu $p(n)[\lambda(n) + \mu(n)]$.

Rozwiązując ten układ równań otrzymujemy:

$$\begin{aligned} p(1) &= p(0) \frac{\lambda(0)}{\mu(1)}, \\ p(2) &= p(0) \frac{\lambda(0)\lambda(1)}{\mu(1)\mu(2)}, \\ &\dots \\ p(n) &= p(0) \frac{\lambda(0)\dots\lambda(n-1)}{\mu(1)\dots\mu(n)}, \\ &\dots \end{aligned}$$

Biorąc pod uwagę warunek normalizowania p otrzymamy:

$$p(0) = \frac{1}{1 + \sum_{n=0}^{\infty} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)}} \quad (1.5)$$

Warunkiem istnienia rozwiązania jest różne od zera prawdopodobieństwo tego, że system jest pusty - ograniczona wartość $\sum_{n=0}^{\infty} \prod_{i=0}^{n-1} \frac{\lambda(i)}{\mu(i+1)}$.

1.3 Markowski model otwartej sieci stanowisk

Przyjmijmy, że sieć składa się z M stanowisk, czas obsługi stanowiska $i, i = 1, \dots, M$, ma rozkład wykładniczy z parametrem $\mu_i(n_i)$, gdzie n_i jest liczbą

klientów obecnych w tym stanowisku; kolejki wszystkich stanowisk mają regulamin naturalny. Drogę klientów określa macierz prawdopodobieństw przejść

$$\mathbf{R} = [r_{ij}]_{i=1,\dots,M,j=1,\dots,M}, \quad (1.6)$$

w której r_{ij} oznacza prawdopodobieństwo, że po zakończeniu obsługi w stanowisku i klient przejdzie do stanowiska j .

Wektor $r_0 = [r_{01}, \dots, r_{0M}]$ określa prawdopodobieństwa r_{0i} , że wchodzący do sieci klient trafi najpierw do stanowiska i . Prawdopodobieństwo $r_{i0} = 1 - \sum_{j=1}^M r_{ij}$ to prawdopodobieństwo, że po zakończeniu obsługi w stanowisku i klient opuści definitywnie sieć. Strumień nadchodzących z zewnątrz zgłoszeń jest poissonowski, a jego parametr λ jest zależny od liczby $N = \sum_{i=1}^M n_i$ zgłoszeń obecnych już w sieci, $\lambda = \lambda(N)$. Stan sieci określa wektor $n(t) = [n_1(t), \dots, n_M(t)]$. Ustalmy następujące oznaczenia stanów sąsiadujących ze stanem n . Stan:

$$n^{j+} = [n_1, \dots, n_{j-1}, n_j + 1, n_{j+1}, \dots, n_M] \quad (1.7)$$

to stan, w którym, w stosunku do stanu n , jest o jednego klienta więcej w stanowisku j , a w pozostałych stanowiskach jest tyle samo co w stanie n klientów. Stan:

$$n^{j-} = [n_1, \dots, n_{j-1}, n_j - 1, n_{j+1}, \dots, n_M] \quad (1.8)$$

to stan, w którym, w stosunku do stanu n , jest o jednego klienta mniej w stanowisku j , a w pozostałych stanowiskach jest tyle samo klientów. Stan:

$$n^{j+,i-} = [n_1, \dots, n_{j-1}, n_j + 1, n_{j+1}, \dots, n_{i-1}, n_i - 1, n_{i+1}, \dots, n_M] \quad (1.9)$$

to stan, w którym, w stosunku do stanu n , jest o jednego klienta więcej w stanowisku j , o jednego klienta mniej w stanowisku i , a w pozostałych stanowiskach jest tyle samo klientów.

Tak jak w przypadku pojedynczego stanowiska, możemy napisać równanie określające prawdopodobieństwo stanu n w chwili $t + \Delta t$:

$$\begin{aligned} p(n, t + \Delta t) = & p(n, t)[1 - \lambda(N)\Delta t]\Pi_{j=1}^M\{[1 - \mu_j(n_j)\Delta t](1 - r_{jj})\} + \\ & + \sum_{j=1}^M p(n^{j+}, t)\mu_j(n_j + 1)r_{j0}\Delta t + \\ & + \sum_{j=1}^M p(n^{j-}, t)\lambda(N - 1)r_{0j}\Delta t + \\ & + \sum_{i=1}^M \end{aligned}$$

Jest to nieskończony układ równań, jako że liczba możliwych stanów n jest nieskończona. Zapisane w kolejnych liniach składniki wyrażenia po prawej stronie odpowiadają różnym możliwościom dojścia do stanu n w chwili $t + \Delta t$:

1. w chwili t sieć była już w stanie n i nic się nie zmieniło - nikt nie przyszedł i nie zakończyła się obsługa w żadnym ze stanowisk - lub obsługa zakończyła się i klient powrócił od razu do kolejki tego samego stanowiska;
2. w chwili t sieć był a w stanie n^{j+} , $j = 1, \dots, M$, w czasie Δt zakończył a się obsługa klienta w stanowisku j i klient opuścił sieć;
3. w chwili t sieć był a w stanie n^{j-} , $j = 1, \dots, M$, w czasie Δt przyszedł do sieci jeden klient i skierował się do stanowiska j ;
4. w chwili t sieć był a w stanie $n^{j+,i-}$, $j = 1, \dots, M$, $i = 1, \dots, M$, $i \neq j$, w czasie Δt zakończył a się obsługa w stanowisku j i klient przeszedł do stanowiska i .

Jeżeli istnieje stan ustalony, to z powyższych równań otrzymujemy układ:

$$\begin{aligned} 0 = & -p(n)\{\lambda(N) + \sum_{j=1}^M \mu_j(n_j)(1 - r_{jj})\} + \\ & + \sum_{j=1}^M p(n^{j+})\mu_j(n_j + 1)r_{j0} + \sum_{j=1}^M p(n^{j-})\lambda(N - 1)r_{0j} + \\ & + \sum_{i=1}^M \sum_{j=1, j \neq i}^M p(n^{j+,i-})\mu_j(n_j + 1)r_{ji} \end{aligned}$$

lub, po drobnych przekształceniach:

$$\begin{aligned}
0 &= -p(n)\{\lambda(N) + \sum_{j=1}^M \mu_j(n_j)\} + \\
&\quad \sum_{j=1}^M p(n^{j+})\mu_j(n_j + 1)r_{j0} + \sum_{j=1}^M p(n^{j-})\lambda(N - 1)r_{0j} + \\
&\quad \sum_{i=1}^M \sum_{j=1}^M p(n^{j+,i-})\mu_j(n_j + 1)r_{ji}
\end{aligned}$$

Układ ten posiada następujące rozwiązanie:

$$p(n) = \frac{1}{G}\Lambda(N)V(n), \quad (1.10)$$

gdzie

$$\begin{aligned}
\Lambda(N) &= \Pi_{m=0}^{N-1} \lambda(m) \\
V(N) &= \Pi_{i=1}^M \Pi_{m=1}^{n_i} \frac{V_i}{\mu_i(m)}
\end{aligned}$$

Elementy wektora $\mathbf{V} = [V_1, \dots, V_M]$ są rozwiązaniami równań ruchu

$$V_i = r_{0i} + \sum_{j=1}^M V_j r_{ji}, \quad i = 1, \dots, M \quad (1.11)$$

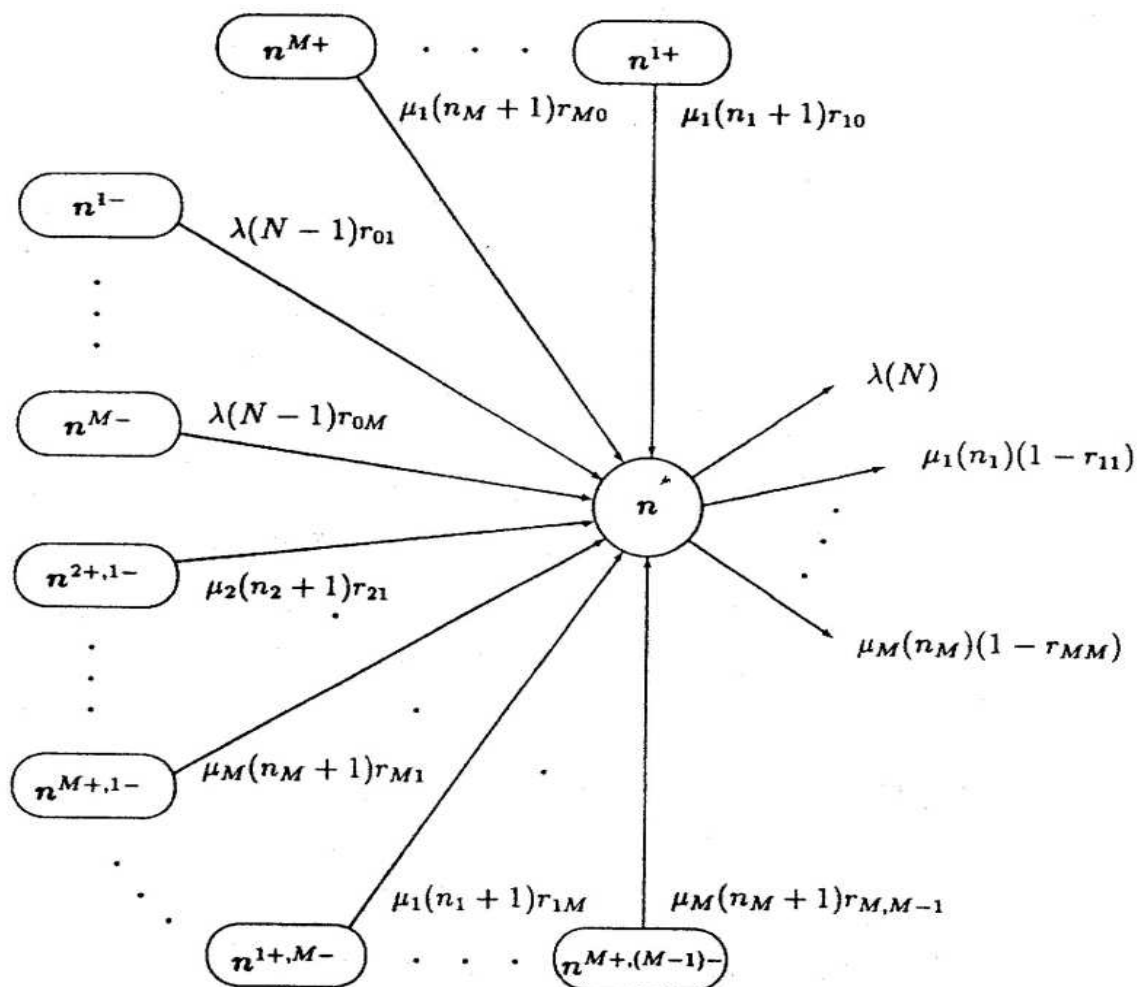
V_i jest średnią liczbą wizyt, jaką klient, raz wszedłszy do sieci, składa w stanowisku i . Stała normalizacyjna G to:

$$G = \sum_{N=0}^{\infty} H(N)\lambda(N), \quad (1.12)$$

gdzie:

$$H(N) = \sum_{n: \sum_{i=1}^M n_i = N} V(n) \quad (1.13)$$

Warunkiem istnienia stanu ustalonego jest skończona wartość stałej normalizacyjnej $G(N)$, czyli zbieżność szeregu we wzorze (1.12).



Możliwości wejścia w stan n i wyjścia ze stanu n

Rozdział 2

Łańcuchy Markowa

2.1 Podstawowe pojęcia

Zdarzeniem losowym związanym z eksperymentem losowym, a zatem i z przestrzenią próbek, jest zbiór prawdziwych stwierdzeń o wynikach eksperymentu, tzn. podzbiór przestrzeni próbek składający się z tych punktów, dla których stwierdzenia są prawdziwe.

Zdarzenia losowe oznaczane będą dużymi literami A , B ,...

Mówimy, że zdarzenie A zachodzi wtedy i tylko wtedy, gdy próbka (wynik eksperymentu) sprzyja zdarzeniu A . Całą przestrzeń próbek Ω nazywamy *zdarzeniem pewnym*, a zbiór pusty \emptyset zdarzeniem niemożliwym.

Zdarzenie A *pociąga* zdarzenie B wtedy i tylko wtedy gdy $A \subset B$, tzn. gdy A jest podzbiorem B .

Zdarzenia A i B są *równoważne* wtedy i tylko wtedy gdy $A = B$ w znaczeniu mnogościowym, tzn. gdy jedno zdarzenie pociąga drugie i na odwrót.

Sumą A i B nazywamy zdarzenie losowe składające się z próbek należących do A , albo do B , albo do obu tych zbiorów i oznaczamy $A \cup B$.

Iloczynem zdarzeń A i B nazywamy zdarzenie losowe składające się z próbek

należących do obu zbiorów A i B , i oznaczamy przez $A \cap B$. Jeśli A pociąga B , to $A \cap B = A$ i $A \cup B = B$.

Dopełnienie zdarzenia A oznaczamy jako A^c i określamy jako zbiór składający się z tych próbek, które nie należą do A .

Suma n zdarzeń losowych A_1, \dots, A_n , oznaczana przez $A_1 \cup \dots \cup A_n$ albo $\bigcup_{i=1}^n A_i$, jest zdarzeniem losowym składającym się z próbek należących do co najmniej jednego ze zdarzeń $A_i, 1 \leq i \leq n$. Analogicznie suma przeliczalnie wielu zdarzeń losowych A_1, A_2, \dots , oznaczana przez $\bigcup_{i \geq 1} A_i$ jest zdarzeniem losowym składającym się z próbek należących do co najmniej jednego ze zdarzeń $A_i, 1 \leq i$.

Podobnie: iloczyn n zdarzeń losowych A_1, \dots, A_n , oznaczany przez $A_1 \cap \dots \cap A_n$ albo $\bigcap_{i=1}^n A_i$, jest zdarzeniem losowym składającym się z próbek należących do każdego ze zdarzeń $A_i, 1 \leq i \leq n$. Analogicznie przekrój przeliczalnej liczby zdarzeń losowych A_1, A_2, \dots , oznaczany przez $\bigcap_{i \geq 1} A_i$ jest zdarzeniem losowym składającym się z próbek należących do każdego ze zdarzeń $A_i, 1 \leq i$.

Dwa zdarzenia losowe A i B *wykluczają się wzajemnie*, albo są *rozłączne* jeśli nie istnieją próbki należące równocześnie do obu zbiorów, tzn. wtedy i tylko wtedy, gdy zajście obu zdarzeń A i B w tej samej próbce eksperymentu jest niemożliwe. Zapisujemy to jako $A \cap B = \emptyset$. Ogólniej, o zdarzeniach A_1, \dots, A_n , mówimy że wykluczają się wzajemnie albo że są rozłączne, wtedy i tylko wtedy, gdy $A_i \cap A_j = \emptyset, i \neq j, 1 \leq i, j \leq n$. Oznacza to, że w danej próbce eksperymentu może zajść nie więcej niż jedno z tych zdarzeń. Podobnie, o przeliczalnej liczbie zdarzeń A_1, A_2, \dots mówimy, że wykluczają się wzajemnie, wtedy i tylko wtedy gdy $A_i \cap A_j = \emptyset$, dla $i \neq j, 1 \leq i, j$

Skończona lub przeliczalna rodzina wzajemnie wykluczających się zdarzeń losowych jest *wyczerpująca* wtedy i tylko wtedy, gdy suma zdarzeń równa jest Ω .

Miarą prawdopodobieństwa lub *prawdopodobieństwem* nazywać będziemy funkcję P , która każdemu zdarzeniu losowemu $A \in \mathcal{P}(\Omega)$ przyporządkuje liczbę nie-

ujemną $P(A)$ taką, że $P(\Omega)=1$ i

$$P\left(\bigcup_{i \geq 1} A_i\right) = \sum_{i \geq 1} P(A_i), \quad (2.1)$$

dla rozłącznych zdarzeń A_1, A_2, \dots , tzn. P jest *przeliczalnie addytywna*.

Przyjmując $A_i = \emptyset$, $i \geq 1$, wynika z powyższej nierówności, że $P(\emptyset)=0$, a następnie przyjmując $A_i = \emptyset$, dla $i \geq n$ otrzymujemy nierówność:

$$P\left(\bigcup_{i \geq 1}^n A_i\right) = \sum_{i \geq 1}^n P(A_i), \quad (2.2)$$

dla dowolnych rozłącznych zdarzeń losowych A_1, \dots, A_n . Z ostatniej równości wynika również, że prawdopodobieństwo P jest również funkcją *skończenie addytywną*. Zauważmy, że ze skończonej addytywności nie wynika przeliczalna addytywność funkcji.

Skończona addytywność pociąga zależność

$$P(A^c) = 1 - P(A), \quad (2.3)$$

jako, że $A \cup A^c = \Omega$, $A \cap A^c = \emptyset$, a z $P(A^c) \geq 0$ otrzymujemy $P(A) \leq 1$ dla każdego $A \in \mathcal{P}(\Omega)$. Ogólniej, jeżeli $A \subset B$, to $P(A) \leq P(B)$. Wynika to z rozłączności zdarzeń A i $B \cap A^c$, i z równości $B = A \cup (B \cap A^c)$.

Konsekwencją przeliczalnej addytywności jest równość

$$P\left(\bigcup_{i \geq 1} A_i\right) = \lim_{x \rightarrow \infty} P(A) \quad (2.4)$$

dla dowolnego zstępującego ciągu zdarzeń losowych $A_1 \supset A_2 \supset \dots$. Z kolei z powyższej równości wynika, że dla dowolnego wstępującego ciągu zdarzeń losowych $A_1 \subset A_2 \subset \dots$

$$P\left(\bigcap_{i \geq 1} A_i\right) = \lim_{x \rightarrow \infty} P(A) \quad (2.5)$$

Przez \mathcal{K} oznaczmy pewną klasę podzbiorów Ω , z $\mathcal{P}(\Omega)$ zmienionym na \mathcal{K} . Taką klasę podzbiorów nazwiemy *σ -algebrą* podzbiorów Ω .

Miarą probabilistyczną na \mathcal{K} nazywamy funkcję, która każdemu zdarzeniu losowemu $A \in \mathcal{K}$ przyporządkuje nieujemną liczbę $P(A)$ taką, że $P(\Omega)=1$ i która jest przeliczalnie addytywna, tzn.

$$P\left(\bigcup_{i \geq 1} A_i\right) = \sum_{i \geq 1} P(A_i), \quad (2.6)$$

dla dowolnych rozłącznych zdarzeń losowych A_1, A_2, \dots

Trójkę (Ω, \mathcal{K}, P) nazywamy *przestrzenią probabilistyczną*.

Mówimy, że dwa zdarzenia losowe $A, B \in \mathcal{K}$ są stochastycznie niezależne wtedy i tylko wtedy, gdy $P(A \cap B) = P(A)P(B)$.

Funkcję X , która przyporządkowuje każdemu punktowi ω liczbę rzeczywistą $X(\omega)$ nazywamy *zmienną losową o wartościach rzeczywistych* wtedy i tylko wtedy, gdy zbiór $\{\omega: X(\omega) < x\}$ tych punktów ω dla których $X(\omega) < x$, jest zdarzeniem losowym $\in \mathcal{K}$ dla dowolnego rzeczywistego x .

Zmienną losową X nazywamy (*absolutnie*) *ciągłą* wtedy i tylko wtedy, gdy istnieje taka nieujemna funkcja p_X określona na prostej rzeczywistej, że

$$\int_{-\infty}^{\infty} p_X(x) dx = 1 \quad (2.7)$$

oraz

$$P(a < X < b) = \int_a^b p_X(x) dx \quad (2.8)$$

dla dowolnych liczb rzeczywistych $a < b$. Funkcja p_X zwana jest *gęstością (prawdopodobieństwa)* zmiennej losowej X .

Rozpatrzmy przykład gęstości:

$$p(X) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2.9)$$

Mamy $p(x) \geq 0$ i

$$\int_{-\infty}^{\infty} p(x) dx = \lambda \int_0^{\infty} e^{-\lambda x} dx = -e^{-\lambda x} \Big|_0^{\infty} = 1$$

zatem p jest gęstością. Nazywana jest ona *gęstością wykładniczą* o parametrze λ . Ciągłą zmienną losową o wykładniczej gęstości nazywamy *wykładniczą zmienną losową*.

Stwierdzenie Wykładnicza zmienna losowa X ma własność braku pamięci, tzn.

$$P(X > x + y | X > y) = P(X > y) \quad (2.10)$$

dla dowolnych liczb nieujemnych x i y .

Dowód: Mamy

$$\begin{aligned} P(X > x + y | X > x) &= \frac{P(\{X > x + y | X > y\} \cap \{X > x\})}{P(X > x)} = \\ &= \frac{P(X > x + y)}{P(X > x)} = \frac{\lambda \int_{x+y}^{\infty} e^{-\lambda u} du}{\lambda \int_x^{\infty} e^{-\lambda u} du} = \\ &= \frac{-e^{-\lambda u} \big|_{x+y}^{\infty}}{-e^{-\lambda u} \big|_x^{\infty}} = \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}} = e^{-\lambda y} = \lambda \int_y^{\infty} e^{-\lambda u} du = P(X > y) \end{aligned}$$

Brak pamięci gęstości wykładniczej jest podstawowym faktem teorii procesów Markowa. Gęstość wykładnicza jest również wygodnym modelem różnego typu problemów *czasu oczekiwania*, dotyczących czasów obsługi, defektów oprzyrządowania przemysłowego, wezwań telefonicznych itp. W wielu takich sytuacjach uzasadnione jest założenie, że czas upływający do zajścia pewnego zdarzenia losowego jest zmienną wykładniczą.

Założmy, że dla dowolnego punktu t rzeczywistej półosi nieujemnej $[0, \infty)$ mamy daną dyskretną zmienną losową, przyjmującą wartości w skończonym lub przeliczalnie nieskończonym zbiorze S .

Rodzinę zmiennych losowych $(X(t))_{t \geq 0}$ nazywamy procesem stochastycznym (albo losowym). Skończenie wymiarowe rozkłady łączne procesu definiujemy jako rozkłady łączne zmiennych losowych $X(t_1), \dots, X(t_n)$, dla dowolnych wartości $t_1 < \dots < t_n$, $n \geq 1$, tzn.:

$$p_{t_1, \dots, t_n}(x_1, \dots, x_n) = P(X(t_1) = x_1, \dots, X(t_n) = x_n) \quad (2.11)$$

Nieujemną macierz A nazywamy stochastyczną wtedy i tylko wtedy, gdy suma wyrazów w każdym wierszu A jest równa 1. Mówimy, że macierz A jest podwójnie stochastyczna wtedy i tylko wtedy, gdy suma wyrazów w każdej kolumnie również jest równa jeden. Każdy wiersz macierzy stochastycznej zwany jest wektorem probabilistycznym. W macierzy podwójnie stochastycznej zarówno wiersza jak i kolumny są wektorami probabilistycznymi.

Mówimy, że ciąg zmiennych losowych $(X(t))_{t \geq 0}$ o wartościach w S jest (jednorodnym, skończonym) łańcuchem Markowa z przestrzenią stanów S , rozkładem początkowym p i macierzą przejść P wtedy i tylko wtedy, gdy $P(X(0) = i) = p(i), i \in S$, oraz

$$\begin{aligned} P(X(n+1) = i_{n+1} | X(n) = i_n, \dots, X(0) = i_0) = \\ = P(X(n+1) = i_{n+1} | X(n) = i_n) = p(i_n, i_{n+1}) \end{aligned}$$

Pierwsza z powyższych równości znana jest jako *własność Markowa*. Jednorodność wprowadzona jest drugą równością, która orzeka, że prawdopodobieństwo przebywania w stanie i_{n+1} w chwili $n+1$ pod warunkiem, że w chwili n proces przebywał w stanie i_n , ma określoną wartość niezależną od n .

Mówimy, że zdarzenie A *poprzedza moment* $n \geq 0$ wtedy i tylko wtedy, gdy można je określić za pomocą zmiennych losowych $X(0), \dots, X(n)$, tzn. gdy można je przedstawić w postaci $A = \{(X(0), \dots, X(n)) \in A^{n+1}\}$, gdzie A^{n+1} jest zbiorem ciągów długości $n+1$ o wyrazach z przestrzeni stanów łańcucha. Podobnie mówimy, że zdarzenie B jest *poprzedzane przez moment* $n \geq 0$ wtedy i tylko wtedy, gdy można je określić za pomocą zmiennych losowych $X(n), X(n+1), \dots$, tzn. gdy można je przedstawić w postaci $B = \{(X(n), X(n+1), \dots) \in B^{(\infty)}\}$, gdzie $B^{(\infty)}$ jest zbiorem ciągów nieskończonych o wyrazach z przestrzeni stanów łańcucha.

Dla dowolnych zdarzeń A i B , odpowiednio, poprzedzającego i poprzedzanego przez moment n , mamy

$$\begin{aligned}
P(B|X(n) = i_n, A) &= P(B|X(n) = i_n, A)P(A|X(n) = i_n) = \\
&= P(B|X(n) = i_n)P(A|X(n) = i_n)
\end{aligned}$$

Wobec tego dla dowolnego łańcucha Markowa przeszłość (A) i przyszłość (B) są warunkowo niezależne pod warunkiem teraźniejszości ($\{X(n) = i_n\}$).

Prawdopodobieństwa przejścia $p(i, j); i, j \in S$, odpowiadają za zmiany stanu łańcucha Markowa z kroku na krok. Prawdopodobieństwa opisujące zmiany stanu po $n > 1$ krokach określone są w prosty sposób poprzez $p(i, j)$. Oznaczmy symbolem $p(n, i, j)$ prawdopodobieństwo przejścia ze stanu i do stanu j w dokładnie n krokach,

$$p(n, i, j) = P(X(m+n) = j | X(m) = i) \quad (2.12)$$

dla dowolnej liczmy naturalnej $m \geq 0$. Prawa strona równości nie zależy od m i ma wartość

$$\begin{aligned}
\sum_{i_1, \dots, i_{n-1} \in S} P(X(m+n) = j, X(m+n-1) = i_{n-1}, \dots, X(m+1) = i_1 | X(m) = i) = \\
= \sum_{i_1, \dots, i_{n-1} \in S} p(i, i_1)p(i_1, i_2) \dots p(i_{n-1}, j), n > 2.
\end{aligned}$$

Oczywiście $p(1, i, j) = p(i, j)$. Zdefiniujmy też $p(0, i, j) = \delta(i, j)$, $i, j \in S$. Z powyższych wynika, że

$$p(2, i, j) = \sum_{k \in S} p(i, k)p(k, j) \quad (2.13)$$

czyli macierz $[p(2, i, j)]_{i, j \in S}$ jest równa kwadratowi macierzy przejść

$\mathbf{P} = [p(i, j)]_{i, j \in S}$ - prawdopodobieństwa $p(n, i, j)$ (*prawdopodobieństwa przejścia w n krokach*) są elementami n -tej potęgi macierzy \mathbf{P} .

Macierz \mathbf{P}^n jest nazywana *macierzą przejść w n krokach*, $n \geq 1$. Dla każdego n jest to macierz stochastyczna, czyli

$$p(n, i, j) \geq 0; i, j \in S, \sum_{j \in S} p(n, i, j) = 1; i, j \in S \quad (2.14)$$

Przykład modelu matematycznego - gra sportowa

Przyjmijmy istnienie gry prowadzonej przez dwóch graczy A i B , która kończy się w momencie uzyskania pewnego rezultatu. Gracz odnosi zwycięstwo, jeżeli wygra pewną liczbę jednostek podstawowych, które mogą być połączone w jednostki złożone. Podstawowa jednostka polega na wywalczeniu pewnej liczby punktów i jest wygrywana przez pierwszego gracza, który uzyska r punktów przy różnicy co najmniej dwóch punktów pomiędzy graczami. Załóżmy, że prawdopodobieństwa uzyskania punktu przez A i B są odpowiednio równe p i q , $p + q = 1$.

Podobną regułę stosujemy do jednostek złożonych.

Rozpatrzmy przypadek tenisa, gdzie jednostką podstawową jest gem, $r=4$, jednostką złożoną set, a mecz wygrywany przez gracza, który zwycięży w 3 setach. Każda z następujących $r^2 + 1$ par liczb może być wynikiem jednostki podstawowej: $u - v, 0 \leq u, v \leq r - 1, (u, v) \neq (r - 2, r - 2)(r - 2, r - 1)(r - 1, r - 1)(r - 1, r - 2)$, przewaga A (A prowadzi $(u + 1) - u, u \geq r - 2$), przewaga B , równowaga ($u - u, u \geq r - 2$), jednostka dla A , jednostka dla B . Wyniki te będziemy uważali za stany łańcucha Markowa o prawdopodobieństwach przedstawionych na rysunku, w przypadku $r=4$

$$p(4) = p^r(1 + \binom{r}{1}q + \binom{r+1}{2}q^2 + \dots + \binom{2r-4}{r-3}q^{r-3})$$

Stany łańcucha Markowa można pogrupować w klasy w zależności od tego czy mogą one być osiągnięte z dowolnego danego stanu. Mówimy, że „stan j jest osiągalny ze stanu i ” i piszemy wówczas $i \rightarrow j$ wtedy i tylko wtedy, gdy istnieje takie $n \geq 1$, że $p(n, i, j) > 0$. Z kolei mówimy, że „stan i komunikuje się ze stanem j ” i piszemy $i \leftrightarrow j$ wtedy i tylko wtedy, gdy zarówno $i \rightarrow j$, jak i $j \rightarrow i$. Relacja „ \rightarrow ” jest przechodnia, tzn. jeżeli $i \rightarrow j$ i $j \rightarrow k$ to $i \rightarrow k$. Wynika to z nierówności

$$p(m + n, i, k) \geq p(m, n, i)p(n, j, k).$$

Relacja \leftrightarrow jest również przechodnia, a także jest relacją symetryczną, tzn. $i \leftrightarrow j$ pociąga $j \leftrightarrow i$.

Klasą stanów nazywamy największy podzbiór przestrzeni stanów taki, że dowolne dwa jego stany komunikują się. Dwie różne klasy muszą być rozłączne, ponieważ istnienie elementu wspólnego powodowałoby komunikowanie się pomiędzy ich stanami, co implikowałoby połączenie się tych klas w jedną. Stan *niepowracalny* to stan nie komunikujący się ze sobą, a zarazem z żadnym innym stanem. Stan taki można uważać za osobną klasę. Pomijając stany niepowracalne, relacja \leftrightarrow jest relacją równoważności i jej klasy abstrakcji są identyczne ze zdefiniowanymi powyżej klasami stanów.

Podzbiór M przestrzeni stanów nazywamy *zamkniętym* wtedy i tylko wtedy, gdy $\sum_{j \in M} p(i, j) = 1$ dla każdego $i \in M$. Oznacza to, że podmacierz $[p(i, j)]_{i, j \in M}$ macierzy przejść P jest macierzą stochastyczną. M jest zbiorem zamkniętym wtedy i tylko wtedy, gdy (przy założeniu, że łańcuch startuje ze stanu należącego do zbioru M) niemożliwe jest wyjście z tego zbioru. Najprostszym przykładem zbioru zamkniętego jest klasa składająca się z jednego stanu i , przy czym $p(i, i) = 1$. Stan taki nazywamy *pochłaniającym*.

Jeżeli i jest stanem *powracalnym* ($p(n, i, i) > 0$ dla pewnego $n \geq 1$) to możemy zdefiniować jego okres d_i jako największy wspólny dzielnik wszystkich liczb naturalnych m , dla których $p(m, i, i) > 0$. O stanie i mówimy, że jest *okresowy* albo *nieokresowy*, w zależności czy $d_i > 1$, czy $d_i = 1$.

Rozważmy proces stochastyczny $(X(t))_{t \geq 0}$ o skończonym zbiorze stanów S . Załóżmy, że dla dowolnego ciągu rosnącego $0 = t_0 < \dots < t_n < t_{n+1} < \dots$ wartości parametru t ciąg zmiennych losowych $(X(t))_{t \geq 0}$ ma własność Markowa, czyli

$$\begin{aligned} P(X(t_{n+1})) = i_{n+1} \mid X(t_n) = i_n, \dots, X(t_0) = i_0) = \\ = P(X(t_{n+1}) = i_{n+1} \mid X(t_n) = i_n) \end{aligned}$$

dla dowolnych $i_0, \dots, i_{n+1} \in S; n \geq 0$. Załóżmy dalej, że ostatnie prawdopodobieństwo warunkowe jest funkcją jedynie i_n, i_{n+1} oraz różnicy $t_{n+1} - t_n$, tzn.

$$P(X(t_{n+1}) = i_{n+1} \mid X(t_n) = i_n) = p(t_{n+1} - t_n, i_n, i_{n+1}).$$

Oczywiście $P(t) = [p(t, i, j)]_{i, j \in S}$ jest macierzą stochastyczną dla dowolnego $t > 0$.

Założmy istnienie procesu stochastycznego, takiego że

$$P(X(t_1) = i_1, \dots, X(t_n) = i_n) = p_{t_1 \dots t_n}(i_1, \dots, i_n)$$

dla dowolnych $0 < t_1 < \dots < t_n, i_1, \dots, i_n \in S$ i $n \geq 1$. Nazwijmy go (*jednorodnym, skończonym*) *łańcuchem Markowa o przestrzeni stanów S , rozkładzie początkowym $\mathbf{p} = [p(i)]_{i \in S}$ i macierzowej funkcji przejść $P(\bullet)$* .

Przyjmijmy istnienie

$$\mathbf{P}'(0) = \mathbf{Q} = [q(i, j)]_{i, j \in S} \quad (2.15)$$

którą nazywamy *macierzą intensywności* funkcji macierzowej $P(\bullet)$. Zauważmy, że dla każdego $i, j \in S, i \neq j$

$$q(i, j) = \lim_{h \rightarrow 0^+} \frac{p(h, i, j)}{h} \geq 0$$

i dla każdego $i \in S$

$$q(i, i) = \lim_{h \rightarrow 0^+} \frac{p(h, i, i) - 1}{h} \leq 0.$$

Zamiast $q(i, i)$ będziemy pisali $-q(i)$, przy czym $q(i) \geq 0; i \in S$. Z równości $\sum_{j \in S} p(t, i, j) = 1; i \in S$ mamy:

$$\sum_{j \in S} q(i, j) = 0; i \in S$$

albo w zapisie macierzowym

$$\mathbf{Qe} = \mathbf{0}$$

Macierz intensywności \mathbf{Q} odgrywa podobną rolę do tej, jaką w przypadku łańcucha Markowa odgrywa macierz prawdopodobieństw przejścia w jednym kroku. Wobec tego skończony proces Markowa jest zazwyczaj opisywany podaniem macierzy intensywności macierzowej funkcji przejść.

Badając probabilistyczne znaczenie macierzy intensywności \mathbf{Q} mamy:

$$\begin{aligned} q(i) = -qi, i &= \lim_{h \rightarrow 0^+} \frac{1 - p(h, i, i)}{h} = \\ &= \lim_{h \rightarrow 0^+} \frac{P(X(t+h) = i \mid X(t) = i) - 1}{h}, t \geq 0 \end{aligned}$$

skąd wynika, że

$$P(X(t+h) = i \mid X(t) = i) = 1 - q(i)h + o(h), \quad (2.16)$$

gdzie $o(h)$ oznacza pewną wielkość taką, że $\lim_{h \rightarrow 0^+} \frac{o(h)}{h} = 0$. Zatem, dla dowolnego $t \geq 0$, $q(i)h + o(h)$ jest prawdopodobieństwem warunkowym zdarzenia polegającego na tym, że proces w chwili $t + h$ nie znajduje się w stanie i pod warunkiem, że w chwili t znajdował się właśnie w tym stanie. W związku z tym wielkość $q(t)$ nazywamy *intensywnością ucieczki* ze stanu i .

Zaś dla $i \neq j$ mamy:

$$q(i, j) = \lim_{h \rightarrow 0^+} \frac{p(h, i, j)}{h} = \lim_{h \rightarrow 0^+} \frac{P(X(t+h) = j \mid X(t) = i)}{h}; \quad t \geq 0$$

stąd

$$P(X(t+h) = j \mid X(t) = i) = q(i, j)h + o(h).$$

Wobec tego wartość $q(i, j)$ jest związana z przejściem ze stanu i do stanu j i dlatego jest nazywana intensywnością przejścia z i do j .

Rozpatrzmy przypadek czasu ciągłego. Z twierdzenia Doeblina wynika, że istnieją granice

$$\lim_{t \rightarrow \infty} p(t, i, j) = \pi(i, j); \quad i, j \in S$$

i zbieżność jest typu geometrycznego. Macierz $\Pi = [\pi(i, j)]_{i,j} \in S$ spełnia równania

$$\Pi P(t) = \mathbf{P}(t) \Pi = \Pi = \Pi^2, \quad t \geq 0$$

Podobnie jak dla czasu dyskretnego, mówimy, że rozkład prawdopodobieństwa \mathbf{p} na przestrzeni stanów S jest *stacjonarny*, wtedy i tylko wtedy, gdy $\mathbf{p}'\mathbf{P}(t) = \mathbf{p}'$ dla dowolnego $t \geq 0$. Wiersze macierzy Π tworzą rozkłady stacjonarne.

Zarazem \mathbf{p} jest rozkładem stacjonarnym wtedy i tylko wtedy, gdy $\mathbf{p}'\mathbf{Q} = \mathbf{0}$. Jeżeli \mathbf{p} jest rozkładem stacjonarnym, to $\mathbf{p}'\mathbf{P}(t) = \mathbf{p}'$ dla dowolnego $t \geq 0$, zatem

różniczkując obie strony względem t i biorąc pochodną w punkcie $t = 0$, otrzymamy $\mathbf{p}'\mathbf{Q}=\mathbf{0}$. W drugą stronę: jeżeli $\mathbf{p}'\mathbf{Q}=\mathbf{0}$ to $\mathbf{p}'\mathbf{P}(t)=\mathbf{p}'$ dla dowolnego $t \geq 0$, czyli \mathbf{p} jest rozkładem stacjonarnym. A więc $\Pi Q = 0$. Mamy też $Q\Pi = 0$. Równość tą otrzymujemy różniczkując względem t równanie $P(t)\Pi = \Pi$ i wstawiając $t = 0$.

Rozdział 3

Metody rozwiązywania układów równań liniowych $Ax = b$

Metody numeryczne, które obliczają rozwiązania problemów matematycznych w ustalonej liczbie operacji nazywane są ogólnie metodami bezpośrednimi lub dokładnymi. Dla przykładu eliminacja Gaussa niejednorodnego układu $(n \times n)$ równań liniowych z nieosobliwą macierzą współczynników znajduje rozwiązanie w dokładnie $\frac{n^3}{3} + \frac{n^2}{2} - \frac{5n}{6}$ mnożeniach i $\frac{n^2}{2} + \frac{n}{2}$ dzieleniach. Cechy metod dokładnych:

- * metody bezpośrednie dają rozwiązanie w skończonej liczbie kroków wykorzystując dekompozycję;

- * w bezpośrednich metodach rozwiązywania równań eliminacja jednego niezerowego elementu macierzy w fazie redukcji często pociąga za sobą stworzenie kilku niezerowych elementów tam, gdzie wcześniej były zera. Efekt ten (nazywany wypełnianiem, ang. *fill-in*) nie tylko sprawia, że organizacja przechowywania macierzy staje się bardziej skomplikowana (bo trzeba przewidzieć wstawianie i usuwanie elementów), ale także wielkość wypełnienia może być tak wielka, że dostępna pamięć szybko się skończy. Mająca przynieść sukces implementacja metody bezpośredniej musi przezwyciężyć te trudności.

* metody dokładne umożliwiają określenie liczby wymaganych operacji przed zakończeniem obliczeń;

* dla pewnych klas problemów metody bezpośrednie dają dokładniejszą odpowiedź w krótszym czasie.

Macierz $\mathbf{U} = (u_{ij})$ nazywamy macierzą *górną trójkątną*, jeśli wszystkie jej elementy poniżej głównej przekątnej są zerami, tj. $u_{ij} = 0$ dla $j < i, i = 2, 3, \dots, n$. Analogicznie macierz $\mathbf{L} = (l_{ij})$ nazywamy macierzą *dolną trójkątną*, gdy $l_{ij} = 0$ dla $j > i$.

Rozwiązanie układów równań liniowych o macierzach trójkątnych jest znacznie prostsze niż w przypadku pełnych macierzy. Dla przykładu rozpatrzmy układ o macierzy górnej trójkątnej \mathbf{U} :

$$\begin{array}{ccccccc} u_{11}x_1 & + & u_{12}x_2 & + & \dots & + & u_{1n}x_n & = & b_1 \\ & & u_{22}x_2 & + & \dots & + & u_{2n}x_n & = & b_2 \\ & & & & \ddots & & \vdots & & \vdots \\ & & & & & & u_{nn}x_n & = & b_n \end{array}$$

Zakładamy, że \mathbf{U} jest nieosobliwa, a więc że wszystkie elementy diagonalne u_{kk} są różne od zera. Niewiadomą x_n wyznaczamy z ostatniego równania, $x_n = \frac{b_n}{u_{nn}}$. Dalej, znając $x_n, x_{n-1}, \dots, x_{k+1}$ z k -tego równania obliczamy

$$x_k = \frac{-\sum_{j=k+1}^n u_{kj}x_j + b_k}{u_{kk}} \quad (3.1)$$

Algorytm ten wymaga około $\frac{1}{2}n^2$ mnożeń i tyleż dodawań. Sprawdzimy, że jest numerycznie poprawny. Niech $x = [x_1, \dots, x_n]^T$ będzie rozwiązaniem obliczonym w arytmetyce fl . Mamy więc

$$x_n = fl\left(\frac{b_n}{u_{nn}}\right) = \frac{b_n}{u_{nn}(1 + \varepsilon_{nn})}$$

$$|\varepsilon_{nn}| \leq \frac{2^{-t};(R)}{(4+\sqrt{2})2^{-t}(C)}$$

gdzie (R) to przypadek rzeczywisty, a (C) to przypadek zespolony.

$$x_k = fl\left(\frac{-\sum_{j=k+1}^n u_{kj}x_j + b_k}{u_{kk}}\right) = \frac{-\sum_{j=k+1}^n u_{kj}(1 + \varepsilon_{kj})x_j + b_k}{u_{kk}(1 + \varepsilon_{kk})} \quad (3.2)$$

gdzie ε_{kj} są błędami wytworzonymi przy obliczaniu iloczynu skalarnego o oszacowaniach:

$$|\varepsilon_{kj}| \leq \left\{ \begin{array}{l} (n-j+2)2^{-t}; (R) \\ (n-j+2+\sqrt{2})2^{-t}; (C) \end{array} \right.$$

a ε_{kk} jest błędem wytworzonym przy dodawaniu b_k i dzieleniu przez u_{kk}

$$|\varepsilon_{kk}| \leq \left\{ \begin{array}{l} 2*2^{-t} (R) \\ (5+\sqrt{2})2^{-t} (C) \end{array} \right.$$

Obliczone rozwiązanie \mathbf{x} jest zatem dokładnym rozwiązaniem układu $(\mathbf{U} + \delta\mathbf{U})\mathbf{x} = \mathbf{b}$, przy czym macierz $\delta\mathbf{U} = (u_{kj}\varepsilon_{kj})$ pozornych zaburzeń równoważnych. Jej elementy mają oszacowania:

$$|\delta u_{kk}| \leq \left\{ \begin{array}{l} 2*2^{-t}|u_{kk}|; (R) \\ (5+\sqrt{2})2^{-t}|u_{kk}|; (C) \end{array} \right.$$

$$|\delta u_{kj}| \leq \left\{ \begin{array}{l} (n-j+2)2^{-t}|u_{kj}|; (R) \\ (n-j+2+\sqrt{2})2^{-t}|u_{kj}|; (C) \end{array} \right.$$

3.1 Eliminacja Gaussa

Metoda eliminacji Gaussa polega na przekształceniu rozwiązywanego układu równań $Ax = b$ do równoważnego układu o macierzy górnej trójkątnej. Pokażemy, że jest to równoważne rozkładowi macierzy \mathbf{A} na iloczyn macierzy dolnej trójkątnej i górnej trójkątnej. Proces eliminacji przebiega w $n-1$ etapach. Oznaczmy $A^{(1)} = \mathbf{A}$, $b^{(1)} = \mathbf{b}$, a \mathbf{A}^k i \mathbf{b}^k niech będą macierzą układu i wektorem prawej strony konstruowanymi w k -tym kroku metody ($k=2,3,\dots,n$). Załóżmy tymczasem, że element $a_{11}^{(1)}$ i kolejno obliczane $a_{kk}^{(k)}$ są $\neq 0$. W pierwszym kroku metody, za pomocą pierwszego równania układu wyjściowego

$$a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \dots + a_{1n}^{(1)}x_n = b_1^{(1)}$$

$$a_{21}^{(1)}x_1 + a_{22}^{(1)}x_n + \dots + a_{2n}^{(1)}x_n = b_2^{(1)}$$

...

$$a_{n1}^{(1)}x_1 + a_{n2}^{(1)}x_n + \dots + a_{nn}^{(1)}x_n = b_n^{(1)}$$

eliminuje się niewiadomą x_1 z pozostałych równań. W tym celu pierwsze równanie pomnożone przez $l_{i1} = a_{i1}^{(1)}/a_{11}^{(1)}$ należy odjąć stronami od dalszych równań ($i = 2, 3, \dots, n$). Przechodzimy w ten sposób od układu $\mathbf{A}^{(1)}\mathbf{x} = \mathbf{b}^{(1)}$ do równoważnego układu $\mathbf{A}^2\mathbf{x} = \mathbf{b}^2$ postaci:

$$a_{11}^{(1)}x_1 + a_{12}^{(1)}x_n + \dots + a_{1n}^{(1)}x_n = b_1^{(1)}$$

$$a_{22}^{(2)}x_n + \dots + a_{2n}^{(2)}x_n = b_2^{(2)}$$

...

$$a_{n2}^{(2)}x_n + \dots + a_{nn}^{(2)}x_n = b_n^{(2)}$$

Wykonane przekształcenia możemy zapisać przy użyciu macierzy elementarnego przekształcenia nieunitarnego $\mathbf{L}^{(1)}$ której elementy - l_{i1} zostały zdefiniowane powyżej. Mamy zatem $\mathbf{A}^{(2)} = \mathbf{L}^{(1)}\mathbf{A}^{(1)}$, $\mathbf{b}^{(2)} = \mathbf{L}^{(1)}\mathbf{b}^{(1)}$.

Analogicznie eliminujemy kolejne niewiadome - w k -tym kroku metody eliminujemy niewiadomą x_k z równań $k+1, k+2, \dots, n$ odejmując od nich stronami równanie k -te pomnożone przez l_{ik} ,

$$l_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$$

Tak więc dla $i = k+1, k+2, \dots, n$ mamy

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - l_{ik}a_{kj}^{(k)}$$

$$b_i^{(k+1)} = b_i^{(k)} - l_{ik}b_k^{(k)}$$

lub krócej

$$\mathbf{A}^{(k+1)} = \mathbf{L}^{(k)}\mathbf{A}^{(k)}, \mathbf{b}^{(k+1)} = \mathbf{L}^{(k)}\mathbf{b}^{(k)}$$

Po wykonaniu $n - 1$ kroków eliminacyjnych otrzymujemy układ o macierzy górnej trójkątnej $\mathbf{A}_{(n)}$

$$\begin{array}{cccccc} a_{11}^{(1)}x_1 & + & a_{12}^{(1)}x_2 & + & \dots & + & a_{1n}^{(1)}x_n & = & b_1^{(1)} \\ & & a_{22}^{(2)}x_2 & + & \dots & + & a_{2n}^{(2)}x_n & = & b_2^{(2)} \\ & & & & \ddots & & \vdots & & \vdots \\ & & & & & & a_{nn}^{(n)}x_n & = & b_n^{(n)} \end{array}$$

Jego rozwiązanie jest już proste. Zarazem jest:

$$A^{(n)} = L^{(n-1)}L^{(n-2)}\dots L^{(1)}A^{(1)}$$

a stąd

$$A = A^{(1)} = (L^{(1)})^{-1}(L^{(1)})^{-1}\dots(L^{(n-1)})^{-1}A^{(n)}$$

Iloczyn $(L^{(1)})^{-1}(L^{(1)})^{-1}\dots(L^{(n-1)})^{-1}A^{(n)}$ jest macierzą dolną trójkątną równą

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ \dots & & & \\ l_{n1} & \dots & l_{n,n-1} & 1 \end{pmatrix}$$

Widzimy, że eliminacja Gaussa jest równoważna rozkładowi macierzy \mathbf{A} na iloczyn \mathbf{LU} macierzy dolnej i górnej trójkątnej.

3.2 Rozkład LU

Niech \mathbf{A} będzie macierzą nieosobliwą, która może być zapisana jako iloczyn dwóch trójkątnych i nieredukowalnych macierzy \mathbf{L} i \mathbf{U} , takich, że

$$\sum_{k=1}^n l_{ik}u_{kj} = a_{ij}$$

czyli w zapisie macierzowym $\mathbf{A}=\mathbf{L}\mathbf{U}$. Ponieważ znalezienie rozkładu \mathbf{A} jest równoznaczne z rozwiązaniem układu n^2 równań, a mamy $n^2 + n$ niewiadomych

$$l_{ik}; \quad i = 1, 2, \dots, n; \quad k = 1, 2, \dots, i$$

$$u_{kj}; \quad j = 1, 2, \dots, n; \quad k = 1, 2, \dots, i$$

daje nam to możliwość przypisania określonych wartości n niewiadomym. Wartości te są dobierane tak, aby odnalezienie pozostałych n^2 wartości było jak najprostsze. Zazwyczaj są one równe:

$$l_{ii} = 1; \quad 1 \leq i \leq n \quad (\text{dekompozycja Doolittle'a})$$

lub

$$u_{ii} = 1; \quad 1 \leq i \leq n \quad (\text{dekompozycja Crout'a}).$$

Zasadniczo, powinno się przeprowadzać sprawdzanie wyznacznika, aby zbadać stabilność dekompozycji. Jednak w przypadku łańcuchów Markowa nie jest to wcale potrzebne. Sprawdźmy przypadek, gdy $i > j$. Aby policzyć i -ty wiersz \mathbf{L} oraz j -tą kolumnę \mathbf{U} nie trzeba poruszać się poza u_{jj} . Zatem sumowanie $\sum_{k=1}^n$ może być zastąpione przez $\sum_{k=1}^j$. Podobnie, dla $i \leq j$ dostaniemy $\sum_{k=1}^i$.

Zatem ostatecznie dostajemy:

$$i \leq j : \sum_{k=1}^i l_{ik} u_{kj} = a_{ij}$$

$$i > j : \sum_{k=1}^j l_{ik} u_{kj} = a_{ij}$$

Zakładając, że wyszliśmy z dekompozycji Doolittle'a, wtedy:

$$\sum_{k=1}^{i-1} l_{ik} u_{kj} + u_{ij} = a_{ij}; \quad i \leq j \quad (l_{ii} = 1)$$

i

$$\sum_{k=1}^{j-1} l_{ik} u_{kj} + l_{ij} u_{jj} = a_{ij}; \quad i > j$$

Z pierwszego równania otrzymujemy:

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}; \quad i \leq j, \quad (3.3)$$

a z drugiego:

$$l_{ij} = [a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj}] / u_{jj}; \quad i > j \quad (3.4)$$

Dekompozycja Doolittle'a polega na powtarzaniu tych dwóch równań do uzyskania pierwszego wiersza \mathbf{U} , a następnie odpowiadającej kolumny \mathbf{L} , aż wszystkie wiersze i kolumny \mathbf{U} i \mathbf{L} będą policzone. Eliminacja Gauss'a i dekompozycja Doolittle'a są ze sobą powiązane: górna trójkątna macierz uzyskana z dekompozycji Doolittle'a to nic innego jak zredukowana macierz trójkątna po eliminacji Gauss'a.

3.3 Metody dokładne i łańcuchy Markowa

Naszym problemem jest znalezienie wektora prawdopodobieństwa π z równań

$$\pi Q = 0; \pi \geq 0; \pi e = 1$$

gdzie Q jest macierzą częstości przejść powiązaną z określonym łańcuchem Markowa, a $e = [1, 1, \dots, 1]^T$

Zerowy wyznacznik

W podejściu 'zerowy wyznacznik' robimy kilka specjalnych założeń. Wymagamy, by istniał rozkład \mathbf{LU} macierzy $A = Q^T$. Ponieważ macierz Q jest nieredukowalna, istnieją macierze dolna i górna takie, że $Q^T = LU$. Gdyby Q była redukowalna, problem sprowadziłby się do kilku mniejszych macierzy. Dla przykładu weźmy

$$\text{macierz: } Q = \begin{pmatrix} -2.0 & 1.0 & 1.0 \\ 1.0 & -3.0 & 2.0 \\ 2.0 & 0.0 & -2.0 \end{pmatrix}$$

Zgodnie z regułami rozkładu LU Doolittle'a macierzy Q^T , otrzymujemy co następuje:

$$\begin{aligned}
u_{1j} &= a_{ij}; j = 1, 2, 3 & \implies & u_{11} = -2.0; u_{12} = 1.0; u_{13} = 2.0 \\
l_{i1} &= a_{i1}/u_{11}; i = 2, 3 & \implies & l_{21} = -0.5; l_{31} = -0.5 \\
u_{2j} &= a_{2j} - l_{21}u_{1j}; j = 2, 3 & \implies & u_{22} = -2.5; u_{23} = 1.0 \\
l_{i2} &= (a_{i2} - l_{i1}u_{12})/u_{22}; i = 3 & \implies & l_{32} = -1.0 \\
u_{j3} &= a_{j3} - l_{31}u_{1j} - l_{32}u_{2j}; j = 3 & \implies & u_{33} = 0.0
\end{aligned}$$

Daje to rozkład:

$$\begin{pmatrix} 1.0 & 0.0 & 0.0 \\ -0.5 & 1.0 & 0.0 \\ -0.5 & -1.0 & 1.0 \end{pmatrix} \begin{pmatrix} -2.0 & 1.0 & 2.0 \\ 0.0 & -2.5 & 1.0 \\ 0.0 & 0.0 & 0.0 \end{pmatrix} = \begin{pmatrix} -2.0 & 1.0 & 2.0 \\ 1.0 & -3.5 & 0.0 \\ 1.0 & 2.0 & -2.0 \end{pmatrix}$$

gdzie ostatni wiersz macierzy U jest wyzerowany - jako wynik tego, iż macierz jest osobliwa.

Jeśli ustalimy teraz, że $Ux = z$ i rozwiążemy $Lz = 0$, otrzymamy musi być $z = 0$, ponieważ L jest nieosobliwa. Wróćmy teraz do równania $Ux = z = 0$, w którym ostatni wiersz U jest równy 0. Zatem możemy przypisać dowolną niezerową wartość (nazwijmy ją η) do x_n , jako że będzie zawsze mnożony przez zero z ostatniego wiersza U . W kolejnych krokach możemy teraz wyznaczyć następne składowe wektora x w oparciu o η . Wydawać by się mogło, że wynik zależy od obranego η , jednakże wektor prawdopodobieństw musi spełniać jeszcze jeden warunek-suma jego składowych powinna wynosić 1. Gdy znormalizujemy otrzymane wyniki, uzyskamy szukany wektor prawdopodobieństw.

$$\begin{pmatrix} 1.0 & 0.0 & 0.0 \\ -0.5 & 1.0 & 0.0 \\ -0.5 & -1.0 & 1.0 \end{pmatrix} \begin{pmatrix} -2.0 & 1.0 & 2.0 \\ 0.0 & -2.5 & 1.0 \\ 0.0 & 0.0 & 0.0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$\begin{pmatrix} 1.0 & 0.0 & 0.0 \\ -0.5 & 1.0 & 0.0 \\ -0.5 & -1.0 & 1.0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \begin{pmatrix} 0.0 \\ 0.0 \\ 0.0 \end{pmatrix} = 0$$

$$\begin{pmatrix} -2.0 & 1.0 & 2.0 \\ 0.0 & -2.5 & 1.0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \eta \end{pmatrix} \begin{pmatrix} 0.0 \\ 0.0 \end{pmatrix} = 0$$

$$x = \eta \begin{pmatrix} 1.2 \\ 0.4 \\ 1.0 \end{pmatrix}$$

$$\pi = (\frac{6}{13}, \frac{2}{13}, \frac{5}{13})$$

Rozdział 4

Zastosowanie LU do sieci optycznej

4.1 Sieci optyczne

Internet i aplikacje oparte na protokole internetowym IP spowodowały rewolucję w dziedzinie telekomunikacji. Cztery komputery połączone ze sobą w sieć w 1969 roku w ramach projektu wojskowego DARPA stały się początkiem współczesnego internetu, wykorzystywanego szeroko w cywilnych dziedzinach życia, umożliwiająccego dostęp do informacji różnego rodzaju, także multimedialnej, wymianę poczty elektronicznej, wspólną pracę, wideokonferencje, a w dziedzinie rozrywki przesył muzyki i filmów oraz wieloosobowe gry komputerowe. Pojawiające się nowe aplikacje wymagają z reguły dużych przepustowości i mogą funkcjonować tylko przy wykorzystaniu szybkich połączeń. Zróżnicowanie charakteru aplikacji internetowych rodzi potrzebę zróżnicowania jakości usług w sensie szybkości i niezawodności transmisji. Zmusza to operatorów telekomunikacyjnych do wprowadzenia nowych technik zapewniających gradację jakości usług w miejsce jednakowej dla wszystkich usługi bez gwarancji (*best effort*). Jednym z możliwych rozwiązań są zaproponowane przez IEFT (Internet Engineering Task Force) mechanizmy *Intserv* (*Integrated Services*) i *Diffserv* (*Differentiated Services*), które realizować można

w nowej wersji protokołu IP v.6.

W przeciwieństwie do szybko rozwijających się aplikacji, struktura Internetu zmienia się powoli. Była zaprojektowana głównie do transportu głosu, istotne jej części wciąż bazują na tradycyjnych sieciach SONET/SDH, i wymaga adaptacji do obecnie przesyłanych strumieni danych. Szansę realizacji tych potrzeb stwarzają sieci całkowicie optyczne w których strumień informacji niesiony promieniem nie jest nigdzie między nadawcą a odbiorcą konwertowany do postaci elektronicznej. W obecnej chwili powszechne jest wykorzystanie światłowodów jako nośnika transmisyjnego, jednakże w każdym węźle sieci trzeba dokonać konwersji sygnału świetlnego na sygnał elektroniczny, by komutować pakiety. Pakiety przychodzą z pewnej liczby kierunków i są kierowane dalej, zgodnie ze swoim przeznaczeniem. Z każdego kierunku przychodzą pakiety należące do różnych połączeń, kierowane do różnych portów wyjściowych. Czas przyjęcia pakietu i analizy nagłówka jest zwykle znacznie mniejszy od czasu wysłania pakietu, wynikającego z jego rozmiaru i przepustowości łącza (lub przypustowości przyznanej połączeniu do którego należy pakiet), dlatego w modelu uwzględnione są tylko kolejki przy portach wyjściowych przełącznika. Sposób zarządzania kolejką może uprzywilejować pakiety należące do wybranych połączeń i skrócić ich czas czekania w kolejce, kosztem wydłużenia czasu czekania innych pakietów. Zarządzanie kolejkami pakietów w formie elektronicznej nie przedstawia żadnej trudności - pakiety są buforowane w pamięci i pobierane z niej w miarę potrzeby.

Rozwój technologii WDM (Wavelength Division Multiplexing) i DWDM (Dense WDM) dla jednoczesnego przesyłu wielu wiązek światła o różnej długości fali z wielokrotnia możliwości łącz: dostępne są handlowo systemy przesyłu w jednym łączu 96 wiązek o przepustowości 10 Gb/s każda, a więc łącznie 960 Gb/s; systemu 200-kanalowe pracują już na platformie, możliwy jest transport 40 Gb/s na jednej długości fali, ocenia się, że przy obecnej technologii światło-

wodowej przepustowości jednego włókna sięgną 50 Tb/s. Wąskim gardłem sieci stał się proces konwersji sygnału na formę elektroniczną i na odwrót oraz proces komutacji pakietów w przełącznikach elektronicznych. Trzeba więc stworzyć przełączniki optyczne, komutujące pakiety w formie optycznej. Jedną z trudności przy ich stosowaniu jest brak pamięci optycznej: nie można zapamiętać pakietów zakodowanych w wiązce światła, można jedynie opóźnić ich przesył o jeden lub więcej taktów pracy sieci, wpuszczając niosące pakiety światło do odpowiednio opóźniającej pętli światłowodu. Przy projektowaniu sieci całkowicie optycznej wymagane są więc szczególne rozwiązania. Po pierwsze, pakiety powinny przychodzić do węzła możliwie regularnie, by nie powstawało zatłoczenie. Trzeba to zorganizować, bo strumień transmitowanego przez użytkowników ruchu charakteryzują się dużą zmiennością natężenia, a prócz tego wykazują długoterminową autokorelację, która bardzo wyraźnie zwiększa długość kolejek obserwowanych w konwencjonalnych sieciach. Po drugie, trzeba tak zaprojektować algorytm wyboru trasy pakietów, by wykluczyć ich kolejkowanie, np. jeśli nie można wysłać pakietu w pożądanym kierunku, trzeba wysłać go w innym kierunku, tylko po to by go nie stracić. Ponieważ sieci całkowicie optyczne są dopiero w fazie projektowania, trwa dyskusja nad różnymi możliwymi rozwiązaniami, a użyteczność poszczególnych opcji może być badana poprzez modelowanie.

Modele kolejkowe są stosowane w ocenie efektywności pracy systemów komputerowych i sieci komputerowych od połowy lat 60-tych. Początkowo korzystano z istniejących już wcześniej modeli, opracowanych przez Erlanga i Molinę dla analizy pracy central telefonicznych. Wśród wybitnych matematyków rozwijających teorię kolejek byli Feller, Kołmogorow, Chinczyn, Crommelin, Palm. Później zaczęły powstawać modele opracowane już specjalnie z myślą o systemach informatycznych. Modele kolejkowe przedstawiają badany system jako *stanowisko obsługi*, do którego dochodzą w losowych odstępach czasu *klienci* i są obsługiwani przez

czas, który również jest zmienną losową. Czas obsługi to czas, przez który zdanie wykorzystuje dany zasób. Gdy stanowisko nie jest dostępne, klienci są ustawiani w kolejkę. Stanowiska mogą być łączone w sieć; klienci krążą w niej odwiedzając poszczególne stanowiska z różną częstotliwością. Stanowiska to zasoby badanego systemu, a obsługiwani klienci to wykonywanie przez system zadania. Wszyscy klienci albo mają te same własności albo dzielą się na klasy różniące się między sobą czasem obsługi w stanowiskach i drogą w sieci. Sieć może być zamknięta - krąży w niej wtedy ustalona liczba klientów - lub otwarta, zasilana strumieniem klientów przychodzących z zewnątrz i odprowadzająca ich na zewnątrz po zakończeniu obsługi. Modele kolejkowe powinny, przy założeniu znajomości opisu strumienia wejściowego klientów (najczęściej zakłada się, że znany jest rozkład odstępów czasu pomiędzy nadejściem kolejnych klientów), rozkładu czasu obsługi i regulaminu szeregowania zgłoszeń w kolejce podać rozkład długości kolejki, rozkład czasu czekania w kolejce (lub tylko wartości średnie), a w przypadku kolejki o ograniczonej długości także prawdopodobieństwo, że klient zostanie odrzucony z powodu braku miejsca w kolejce. Modele te, mimo licznych uproszczeń, potrafią dobrze uchwycić zależności pomiędzy podstawowymi parametrami systemu, wskazać jego wąskie gardła i przewidzieć czas odpowiedzi w funkcji obciążenia.

4.2 Przełącznik optyczny

Ze względu na wysoką przepustowość jaką oferują sieci oparte na pakietach optycznych, w ostatnich latach wzrosło zainteresowanie wykorzystaniem i optymalizacją ruchu w tego typu sieciach. Jedną z cech takich sieci jest niemożność buforowania pakietów wewnątrz sieci. Zatem nie można użyć algorytmu transportu "przechowaj i przekaz". W zamian tego opracowano transport oparty na Odbiciu oraz na Skupianiu. Powyższe strategie transportu nie pozwalają na tracenie pakietów.

Jednak, ze względu na przechowywanie pakietów wewnątrz sieci, wzrastają opóźnienia i zarazem zmniejsza się przepustowość. W transporcie Odbicia opartym na najkrótszej ścieżce, przełączniki przesyłają pakiety zgodnie z najkrótszą drogą do celu. Każdy przełącznik może przesłać skończoną ilość pakietów w jednostce czasu. Pakiety przychodzące są natychmiast przesyłane dalej - do następnego przełącznika. Jeśli ilość pakietów, które chcą wejść do przełącznika jest większa niż jego przepustowość, tylko część pakietów będzie przyjęta - pozostałe zostaną skierowane do innego i droga którą będą musiały przebyć będzie dłuższa. Jest to główna wada tej techniki. W skrajnym przypadku pakiety nigdy nie osiągną swojego celu.

Symulacje wykazały, że ilość odbić może być znaczna, zwłaszcza gdy ruch nie jest zrównoważony. Bezwzględna ilość odbić może być niewielka, jednak może być duża w stosunku do aktualnej ilości pakietów w sieci. Zaobserwowano np. ponad 1000 odbić dla kilku pakietów podczas symulacji na dwuwymiarowym modelu 10×10 . Pakiety z bardzo dużą ilością odbić stanowią problem, ponieważ niejednokrotnie ich opóźnienie jest większe niż dopuszczalny czas całkowitego transportu. Pakiety nigdy nie gina fizycznie - po prostu nie docierają do celu gdyż opóźnienie jest zbyt wielkie. Zagadnienia opóźnień w transporcie i przeciętnej przepustowości są głównymi przedmiotami badań. Poprzednie teorie proponowały modele dla sieci przełączników 2×2 bez kolejkowania nowych pakietów. Bardziej współcześnie, Fabrega i Munoz przedstawili model sieci z odbiciami oparty na łańcuchach Markowa. Jednakże topologia zakładała, że istnieje tylko jedna najkrótsza droga pomiędzy źródłem i celem. Ponadto nie udało im się przewidzieć wydajności sieci zbudowanych z większej ilości przełączników. S. Yao przedstawił model dla bardziej ogólnych topologii. Metoda polega na dekompozycji sieci i aproksymacji natężenia ruchu wzdłuż ścieżki. Inny sposób, bazujący na macierzy wielomianów został zaproponowany przez Szcześniaka, by modelować bardziej realistyczny algo-

rytm, w którym pakiet musi być usuwany po S odbiciach. Ponownie analizowane jest przeciętne opóźnienie, a nie jego rozkład. Jak widać, wszystkie poprzednio proponowane modele używają uproszczonych i nierealistycznych topologii sieci.

Rozpatrzmy zestawy przełączników 4×4 i dwuwymiarową topologię torusa. Projekt ROM dowiódł, że dwuwymiarowa siatka może dobrze odzwierciedlać rzeczywistą topologię i łatwo ją przekształcić w torus. Zakładamy ponadto, że przepustowość łącza wynosi tylko 1.

Modelujemy sieć poprzez ogólną reprezentację pakietów. Najpierw tworzymy wektor odległości do celu i dzielimy pakiety na dwie grupy w zależności od ilości możliwych punktów docelowych w kolejnym kroku. Zakładamy że pakiety chcą się poruszać najkrótszą ścieżką. Tylko niektóre z kierunków (z czterech istniejących w torusie) mogą prowadzić najkrótszą drogą. W torusie nieparzystym i układzie współrzędnych będziemy rozróżniać tylko dwa typy: pakiety, które mają tylko jeden dozwolony kierunek zgodny z zasadą najkrótszej ścieżki, i te które mają dwa kierunki. Odpowiednio będziemy je nazywać pakietami 1. typu i pakietami 2. typu. W torusie parzystym istnieją również pakiety typów trzeciego i czwartego. Podczas rozpatrywania odbić, ilość możliwych kierunków ma duże znaczenie dla opóźnień w transporcie. Nasza analiza będzie się opierać o konstrukcję łańcucha Markowa, który będzie reprezentował ewolucję pakietu. Stany będą zależeć od typu pakietu oraz wektora odległości do celu.

Najpierw modelujemy ścieżkę pakietu wewnątrz sieci używając łańcucha Markowa. Rozkład czasu transportu będzie policzony numerycznie gdy zostaną poznane prawdopodobieństwa odbić pakietów obu typów. Prawdopodobieństwa będą określone na podstawie systemu przekazywania pakietów typu 1 i typu 2 - niektóre z reguł pochodzą z pracy Qin Wei'a o algorytmach odbić, a niektóre są rezultatem badań z ostatnich czterech lat. W mojej pracy, przeciwnie do założeń Qin Wei'a, uznajemy, że przed siecią istnieje kolejka i bierzemy pod uwagę

własności algorytmów odbić przy obliczaniu prawdopodobieństw.

4.2.1 Model pakietu

Ostatnio dowiedziono kilka istotnych własności wyboru pakietów, które mają być odbijane. Po pierwsze, pakiety są dzielone na dwa zbiory: pakietów które mają tylko jeden dopuszczony kierunek ruchu oraz pakietów które mają dwa takie kierunki. Dalej będziemy je nazywać odpowiednio pakietami typu 1 i pakietami typu 2. Pakiet typu 1 osiągnęły jedną z koordynat celu, natomiast pakiety typu 2 muszą poruszać się w dwóch kierunkach by osiągnąć cel. Dopuszczalna jest zmiana typu pakietu podczas jego poruszania się, zależnie od obranego algorytmu odbić.

LEMAT: *Jeśli pakiet ma tylko jedno możliwe odbicie, musimy spełnić jego żądanie przed szukaniem pakietów drugiego typu.*

Wróćmy do modelu transportu. Założyliśmy, że wybór pakietów jest oparty na ich typie. Zatem prawdopodobieństwo odbicia nie zależy od dystansu do celu, ale tylko od typu. Niech $p1$ i $p2$ będą odpowiednio prawdopodobieństwami tego, że pakiet typu 1 i typu 2 będzie odbity w jednym kroku. Wartości te zostaną wyprowadzone w następnym rozdziale. Stwórzmy model ewolucji jednego pakietu wewnątrz sieci. Najpierw budujemy macierz przejścia R typowego pakietu i przystosowujemy ją do torusa z 7 wierszami i kolumnami. Model bazuje na następującym zestawie stanów: stan inicjujący (stan 0) zanim pakiet wejdzie do sieci i stan wyjściowy (stan 1) gdzie pakiet opuszcza sieć. Stan 1 jest stanem absorbującym. Określamy typ pakietu oraz wektor odległości do celu. Pamiętajmy, że typ pakietu zależy od ilości dopuszczalnych kierunków w danym kroku. Oczywiście pakiet może zmieniać swój typ w zależności od odległości do celu oraz założeń algorytmu odbić. Na przykład, dla torusa 7×7 takimi stanami są: $(t1,1)$, $(t1,2)$, $(t1,3)$, $(t2,1,1)$, $(t2,1,2)$, $(t2,1,3)$, $(t2,2,2)$, $(t2,2,3)$, $(t2,3,3)$ i łańcuch ma 11 stanów. Czas transportu jest czasem przebycia ścieżki zaczynającej się w stanie 0 i

kończącej się w stanie 1.

Spójrzmy na wpływ algorytmu transportu i odbić na stany. Na początek określimy dwie najprostsze zmiany:

- Pakiet typu 1 który nie jest odbity i który jest w odległości k pozostaje w typie 1 jeśli porusza się do celu w jednym kierunku - wtedy odległość zmienia się na $k-1$.

- Pakiet typu 2 który jest odbity pozostaje pakietem typu 2. Zasadniczo, odbicie zwiększa odległość do celu, z wyjątkiem krawędzi torusa. Każda ze składowych odległości może wzrosnąć z prawdopodobieństwem $\frac{1}{2}$.

Teraz przypatrzmy się dwóm innym przypadkom:

- Pakiet typu 1 który jest odbity ma trzy możliwe kierunki. Dwa kierunki prowadzą do zmiany w pakiet typu 2 i jeden kierunek pozostanie w typie 1. Jeśli założymy równe prawdopodobieństwo wyboru kierunku, pakiet utrzyma typ 1 z prawdopodobieństwem $\frac{1}{3}$, a zmieni typ na 2 z prawdopodobieństwem $\frac{2}{3}$. Odległość do celu to $(1,k)$ poza określonymi przypadkami: kiedy pakiet jest na skraju torusa i utrzymuje typ 1; kiedy pakiet jest wciąż w odległości k ale w odwrotnym kierunku (pamiętajmy, że zajmujemy się torusem nieparzystym).

- Pakiet typu 2 znajdujący się w odległości (m,k) który nie jest odbity zmniejsza odległość do $(m,k-1)$ lub $(m-1,k)$. I w zależności od jego pozycji i obranego kierunku może stać się pakietem typu 1 (jeśli $m-1=0$ lub $k-1=0$) lub pozostać pakietem typu 2.

Macierz przejść R może być określona opierając się na wcześniej pokazanych zasadach i prawdopodobieństwach odbić $p1$ i $p2$. Każdy z punktów w sieci ma to samo prawdopodobieństwo ($\frac{1}{N^2-1}$) stania się celem. Zatem musimy policzyć ilość punktów typu 1 i 2 oraz dystanse k i (m,k) .

$$R(0, (t1, k)) = \frac{4}{N^2-1} \text{ i } R(0, (t2, m, k)) = \frac{x_2(k+m)}{N^2-1}$$

gdzie $x_2(m+k)$ jest liczbą punktów w torusie w odległości k , które mogą przyj-

mować pakiety typu 2 wysyłane do punktu centralnego. Inaczej: $x_2(m+k) = 4k+4m-4$ jeśli $k+m \leq Z$ i $x_2(m+k) = 8Z+4-4k-4m$, gdzie $Z \leq m+k \leq 2Z$.

Przedstawimy dwa modele do obliczania prawdopodobieństw odbić. Ponieważ wszystkie pakiety są z probabilistycznego punktu widzenia identyczne, rozpatrujemy pakiety typu 1 i 2 oraz przełączniki - zauważmy, że ze względu na założenia topologii, przełączniki też są statystycznie jednakowe. Prawdopodobieństwo odbicia pakietu jest obliczone poprzez uwarunkowanie wszystkich możliwych konfiguracji przyjść pakietów typu 1 i typu 2.

Pamiętajmy, że najlepszy algorytm musi przekazywać pakiety typu 1 z najwyższym priorytetem. Wystarczy policzyć prawdopodobieństwo śledzonego pakietu typu 1 i znać ilość wszystkich pakietów typu 1 w przełączniku. Zauważmy, że górna granica sumy to 3, ponieważ śledzony klient zajmuje jeden z linków switcha. Zarazem prawdopodobieństwo to nie zależy od odległości do celu k .

$$p1 = \sum_{i=0}^3 Pr(i \text{ typu 1 przybywa}) d1(i)$$

gdzie $d1$ jest prawdopodobieństwem tego, że pakiet typu 1 będzie odbity jeśli inny pakiet typu 1 przybędzie. Niektóre z prawdopodobieństw $d1(i)$ są liczone w następnym podrozdziale i będą później wytłumaczone. Obliczenia wykorzystują niektóre z własności i charakterystycznych cech nieparzystego torusa.

Lemat *Wszystkie możliwe konfiguracje żądań nie są możliwe ze względu na algorytm przekazywania oraz topologię. Dla odbić wzdłuż najkrótszej ścieżki w nieparzystym torusie, pakiety typu 2 nie mogą żądać dwóch przeciwnych kierunków (np. PÓŁNOCY i POŁUDNIA).*

Prawdopodobieństwa przybyć są określane przez zdarzenia niezależne. Określmy przez u_1 użycie pewnego linku przez pakiety typu 1.

$$Pr(i \text{ typu 1 przybywa}) = C(3, i) (u_1)^i (1 - u_1)^{3-i}$$

Ponieważ klienci typu 2 mają mniejszy priorytet w algorytmie, prawdopodobieństwo ich odbić wymaga wzięcia pod uwagę większej ilości zmiennych.

$$p2 = \sum_{i=0, j=0}^{i+j=3} Pr(i \text{ typu 1 oraz } j \text{ typu 2 przybywa}) d2(i, j)$$

gdzie $d2(i, j)$ określa prawdopodobieństwo z którym śledzony pakiet typu 2 będzie odbity w związku z przybyciami i pakietów typu 1 i j pakietów typu 2.

$$Pr(i \text{ typu 1 i } j \text{ typu 2 przybywa}) = M(i, j, 3 - i - j)(u_1)^i (u_2)^j (1 - u_1 u_2)^{3-i-j},$$

gdzie $M(i, j, 3 - i - j)$ jest mnożnikiem wielomianowym. Pamiętajmy, że $p1$ i $p2$ są sumami po $d1(i)$ i $d2(i, j)$. Własność ta okaże się bardzo użyteczna przy udowadnianiu istnienia rozwiązania.

Teraz musimy określić elementarne prawdopodobieństwa $d1(i)$ i $d2(i, j)$. Najpierw ustalmy śledzony pakiet typu 1. Pakiet ten wchodzi do switcha i konkuruje z i przybyłymi pakietami typu 1. Jeśli nie ma współzawodnictwa, $d1(0)=0$. Podobnie $d2(0,0)=0$, $d2(1,0)=0$ oraz $d2(0,1)=0$. Przyjmijmy teraz, że dwa pakiety (w tym śledzony) typu 1 przybywają do switcha. Wtedy $d1(1)$ jest prawdopodobieństwem tego, że oba pakiety żądają tego samego wyjścia pomnożone przez prawdopodobieństwo odbicia śledzonego pakietu. Zakładamy równe prawdopodobieństwo wyboru gdy kilka pakietów żąda tego samego celu oraz równe prawdopodobieństwo wyboru kierunku, co daje nam: $d1(1) = 1/4 \times 1/2 = 1/8$.

Policzmy teraz prawdopodobieństwo odbicia śledzonego pakietu 1 typu kiedy dwaj inni klienci przybywają do switcha. Najpierw ustalmy cel dla śledzonego klienta i określmy wszystkie możliwe konfiguracje dla pozostałych pakietów. Istnieje 16 przypadków: 9 bez koniecznych odbić, 6 powodujących konflikt ze śledzonym pakietem i ostatni przypadek, gdy mamy konflikt pomiędzy trzema pakietami. Ponieważ odbijany pakiet jest wybierany w czasie konfliktu, otrzymujemy: $d1(2) = 1/16 * (6 * 1/2 + 1 * 2/3) = 11/48$.

Ustalmy teraz relację pomiędzy użyciem linku u i prawdopodobieństwem odbicia. Musimy się skupić na ilości pakietów $n1(k)$ i $n2(m,k)$, a nie na stanie określonego pakietu. Dla przekształceń wewnątrz sieci macierz $M(i,j)$ jest podobna do $R(i,j)$. Zarazem cała grupa pakietów podlega podobnym przekształceniom co jeden, śledzony pakiet. Jedyne różnice są w transmisjach wewnątrz sieci i poza nią.

Na wstępie zapisujemy równanie przepływu dla przeciętnej ilości $n1(k)$ i $n2(m,k)$. Musimy policzyć ilość nowych pakietów 1 i 2 typu przychodzących do sieci w odległości k lub (m,k) . Ustalmy te wielkości jako $l2(k)$ oraz $a2(m,k)$. Przeciętna ilość pakietów wchodzących do sieci jest zarazem przeciętną ilością pakietów docierających do bufora - o ile system jest stabilny. Wartość ta wynosi zazwyczaj λN^2 .

Ale przyływ do sieci musi być równy odpływowi ze switcha z udanym wyjściem z punktu przy odległości 1. Zatem $\lambda N^2 = n1(1) * (1 - p1)$. Usuwanie dwa pierwsze stany z macierzy R i zmieniamy pierwszy wiersz by wziąć pod uwagę przychodzenia pakietów do sieci. Średni wektor rozkładu przychodzących pakietów jest rozwiązaniem układu równań:

$$\begin{aligned}(n1, n2) &= (n1, n2)M \\ n1(1) &= \frac{\lambda N^2}{(1 - p1)}\end{aligned}$$

Używając formuły Little'a, otrzymujemy:

$$n = \lambda N^2 T,$$

gdzie T jest średnią ilością skoków (lub potrzebnym czasem w optycznej części sieci), a n jest przeciętną ilością klientów. Wróćmy do obciążenia linku u . Jest to powiązane z u_1 i u_2 : $u = u_1 + u_2$. Ponieważ wszystkie linki uznajemy za jednakowe, mamy: $u = \frac{n}{4N^2}$. W dalszym ciągu otrzymujemy:

$$n = \sum_{k=1}^Z n_1(k) + \sum_{k=2}^{2Z} n_2(k) \text{ i } u_1 = \frac{\sum_{k=1}^Z n_1(k)}{4N^2} \text{ i } u_2 = \frac{\sum_{k=2}^{2Z} n_2(k)}{4N^2}$$

W ten sposób określiliśmy nasz model.

Musimy przeprowadzić dowód istnienia rozwiązania i algorytmu numerycznego. Najpierw pokażemy że istnieje co najmniej jedno rozwiązanie takiego modelu używając twierdzenia Brouwer'a:

Twierdzenie Brouwer'a: *Niech S będzie zwartym i wypukłym zbiorem z niepustym wnętrzem; jeśli F jest ciągła i jeśli $F(S) \subseteq S$ to F jest punktem stałym.*

Przypomnijmy kilka cech funkcji F i zbioru S .

Lemat p_1 jest wypukłą sumą prawdopodobieństw $d_1(i)$. Dla wszystkich wartości u_1 (które mogą się zmieniać podczas przekształceń) p_1 jest zawsze pomiędzy 0 i maksymalną wartością d_1 , tzn.:

$$p_1 \in [0, \frac{81}{256}]$$

Jako że p_2 jest skończoną sumą prawdopodobieństw $d_2(i,j)$, otrzymujemy:

$$p_2 \in [0, \frac{9}{32}]$$

Niech S będzie podzbiorem R^2 zdefiniowanym przez przez dwie powyższe wartości. $S = [0, \frac{81}{256}] \times [0, \frac{9}{32}]$. S jest ograniczonym i domkniętym zbiorem R^2 i S ma niepuste wnętrze. Ponieważ domknięte i ograniczone podzbiory R^2 są zwarte, S spełnia wymogi założenia. Wróćmy do funkcji F .

Otrzymaliśmy określone wartości p_1 i p_2 i te wartości są w S . Rozwiążemy układ dla przeciętnej ilości pakietów.

Lemat Ponieważ $p_1 < 1$ i $p_2 < 1$, macierz M jest nieredukowalna.

Dowód:

1. Istnieje ścieżka od stanu 1 do dowolnego punktu i . Przejście od 1 do i zawsze ma dodatnie prawdopodobieństwo.

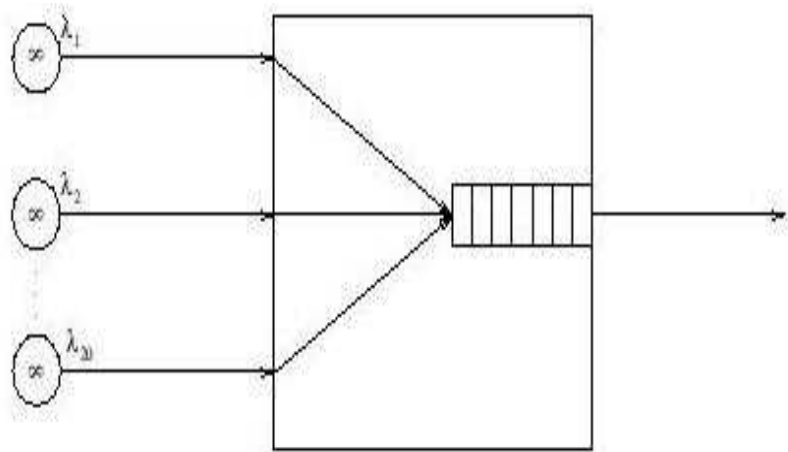
2. istnieje ścieżka od stanu i do stanu 1. Z każdego stanu można przejść do innego z dodatnim prawdopodobieństwem ($p_1 < 1$ i $p_2 < 1$). Poprzez indukcję zawsze istnieje ścieżka do stanu 1.

M jest zawsze ergodyczna gdy (p_1, p_2) jest zawarta w S . Zatem możemy zastosować obliczenia Malisev'a, które zakładają ciągłość rozkładów dla nieredukowalnych, skończonych i dyskretnych łańcuchów Markowa. Zatem wartości $n_1(i)$ i $n_2(i)$ są ciągłe w p_1 i p_2 . Wobec otrzymujemy, że n , n_1 i n_2 są ciągłe. Ponieważ u_1 to $\frac{n_1}{4\pi N^2}$, jest to także ciągłe. podobnie u_2 . Ostatecznie p_1 i p_2 są wielomianowymi funkcjami u_1 i u_2 , więc również są ciągłe. F jest złożeniem funkcji ciągłych, więc jest ciągła. Zmiana wartości u_1 i u_2 zmieni sumy które określają p_1 i p_2 , ale wciąż będą one w S .

Zatem S jest zwarty, wypukły i ma niepuste wnętrze, F jest ciągła i $F(S) \subseteq S$. F spełnia założenia twierdzenia Brouwer'a i F jest punktem stałym.

4.3 Eksperyment numeryczny

Metodę **LU** zastosowano do znajdowania charakterystyk liczbowych analitycznego modelu formowania pakietu optycznego z pakietów elektronicznych. W celu analitycznego zmodelowania formowania z bloków pakietów optycznych stosuje się zazwyczaj model kolejkowy oparty na kolejce FIFO i pewnej ilości źródeł Poissonowskich. W przykładzie ograniczono się do 250 bloków i 20 źródeł. Odpowiada to długości pakietu równej 250 i 20 źródłom. Ponieważ pakiety mogą mieć różną długość, od 1 do 20 bloków, w modelu założono pewien zbiór źródeł, w którym każde źródło generuje pakiety o określonej długości; np.: źródło 1 pakiety składające się z jednego bloku, źródło 2 pakiety składające się z 2 bloków itd.



Schemat przełącznika

Proces tworzenia kolejki jest następujący: Pakiet optyczny jest gotowy do wysłania kiedy 250 bloków przybyło, lub mniej niż 250 bloków jest w kolejce ale pakiet który właśnie przybył spowodowałby wydłużenie kolejki ponad 250 bloków. Ponadto zakłada się, że proces ten nie może przekroczyć pewnego założonego czasu (ang. timeout). Powyższe założenie definiują określony łańcuch Markowa (jak na rysunku 2). Węzły grafu są Markowskimi stanami systemu. Poziomo, krawędzie między węzłami oznaczają przybycie kolejnych pakietów. Ponieważ przybywające pakiety mogą mieć różną długość, może istnieć kilka krawędzi z tego samego węzła - przejść z jednego stanu do kolejnych - w zależności od długości przybyłego pakietu. Dla stanów z ilością bloków w kolejce bliską 250 przejścia mogą spowodować powrót do stanów z małą ilością bloków - próba wstawienia przychodzącego pakietu może powodować przepełnienie kolejki; wtedy tworzenie kolejki pakietów jest kończone, pakiet jest wysyłany i zaczyna się tworzenie nowej kolejki. Pionowe rzędy węzłów oznaczają okresy gdy żaden pakiet nie przybył i bloki czekają określony czas (który nie może być przekroczony) na zakończenie formowania kolejki. Ponieważ model ma charakter łańcucha Markowa z ciągłym czasem, T jest przybliżone rozkładem Erlanga 10-tego rzędu, tzn. jest 10 faz o rozkładzie wykładniczym (średni czas

trwania każdej fazy to w przykładzie $\mu_r = 10$). Kiedy zostaje osiągnięty limit czasu, system wraca do stanu początkowego - pakiet optyczny utworzony do tej pory zostaje wysłany i zaczyna się tworzenie nowej kolejki. Schematycznie działanie przełącznika ukazane jest na diagramie.

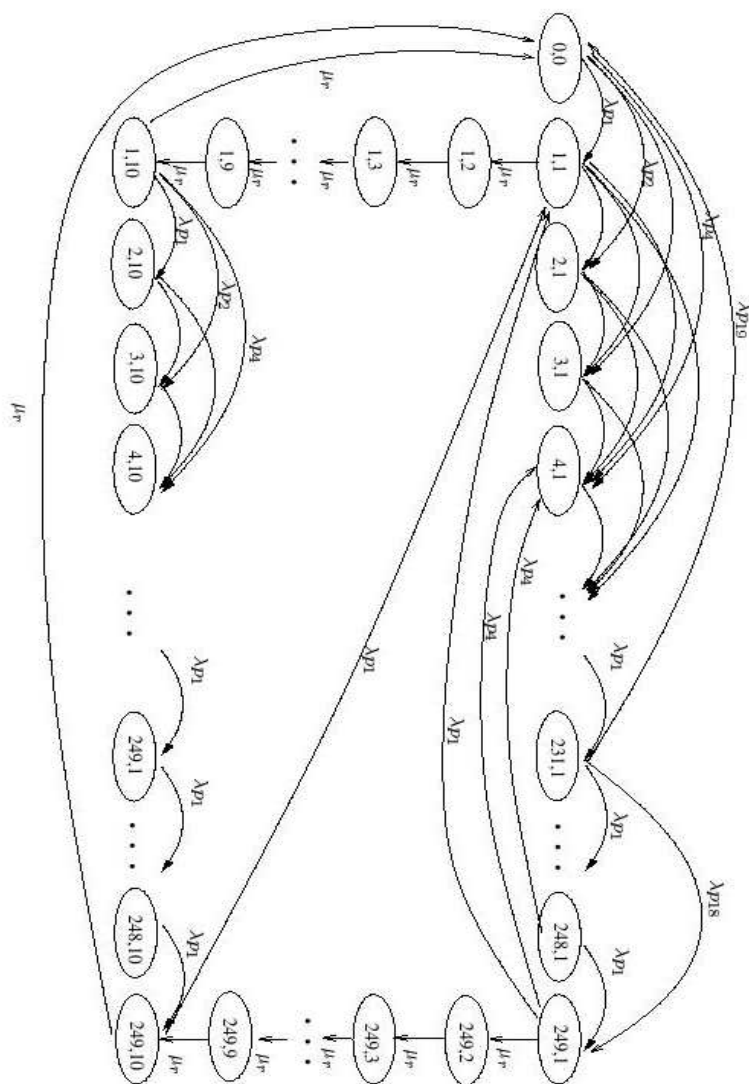


Diagram przejść stanów

Macierz przejścia dla modelu formowania węzła ma 2491 elementów (jest względnie mała) dlatego przechowywano ją w dwuwymiarowej tablicy bez kom-

presji. Algorytmy zaimplementowano w języku C, a program skompilowano przy użyciu kompilatora *gcc* z opcją **-O3**.

```
void RozkladLU(double ** wiersze, double ** gora, int rozmiar)
{
    int i;
    int j;
    unsigned int k;
    double wart;
    double tmp;
    double ** dol;
    dol=new double * [rozmiar+1];
    for(i=rozmiar; i>=0; i--)
        dol[i] = new double [i+1];

    for (i=rozmiar; i>=0; i--)
        for (j=i; j>=0; j--)
            dol[i][j] = 0.0;
    for (i=rozmiar; i>=0; i--)
        dol[i][i]=1.0;

    for (i=0; i<=rozmiar; i++)
    {
        for (j=i; j<=rozmiar; j++)
        {
            wart=wiersze[i][j];
            for (k=0; k<=i; k++)
                wart-=dol[i][k] * gora[k][j];
            gora[i][j]=wart;
        }
        tmp=1.0/gora[i][i];
        for (j=i+1; j<=rozmiar; j++)
        {
            wart=wiersze[j][i];
            for (k=0; k<=i; k++)
                wart-=dol[j][k] * gora[k][i];
            dol[j][i]=wart*tmp;
        }
    }
    //przetwarzanie
} //rozklad LU
```

```

double PrzetwarzanieUkladu(int rozmiar, double ** gora, double * wektor)
{
double suma,sumatmp,tmp,wart;
int i,j;
for (i=0; i<=rozmiar-1; i++)
    wektor[i]=0.0;

suma=wektor[rozmiar]=1;

for (i=rozmiar-1; i>=0; i--)
{
    wart=gora[i][i];
    if (wart!=0.0)
    {
        sumatmp=0.0;
        for (j=i+1; j<=rozmiar; j++)
            sumatmp+=(gora[i][j] * wektor[j]);
        if (sumatmp!=0.0)
        {
            tmp=(sumatmp/wart)*(-1.0);
            wektor[i]=tmp;
            suma+=tmp;
        }
    }
}
suma=1.0/suma;
for (i=rozmiar; i>=0; i--)
    wektor[i]*=suma;
return suma;
} //przetwarzanie ukkladu

```

W poniższej tabeli znajduje się czas rozwiązania problemu wraz z błędem obliczeń.

Środowisko	Czas obliczeń	Błąd ($\ Q^T x\ _2$)
<i>Pentium III 1GHz</i>		
<i>gcc v. 3.3.2</i>	136,92s	$2.205377e^{-3}$
<i>Pentium M 735 (1.7 GHz)</i>		
<i>gcc v. 3.4.2</i>	50,49s	$4,048849e^{-3}$
<i>Pentium IV 733 MHz</i>	98,89s	$3,69610e^{-3}$
<i>Athlon XP 3.2GHz</i>		
<i>gcc v. 3.3.4</i>	89.63s	$2.643332e^{-3}$

Rozdział 5

Wnioski

W swojej pracy skupiłem się na jednej z metod rozwiązywania układów równań (metodzie LU) oraz pokazałem jak można połączyć ją z zagadnieniami Markowskich modeli sieci.

Metoda ta należy do grupy metod dokładnych - jesteśmy w stanie z góry przewidzieć czas wykonania zadania. Jednakże w przypadku metod numerycznych o zastosowaniu tej a nie innej metody często decydują dwa aspekty - dokładność otrzymanych wyników oraz szybkość ich uzyskania. Metoda LU jest metodą dość dobrze opisaną i nie najnowszą, jednak w dalszym ciągu konkurencyjną i umożliwiającą uzyskanie wyników w dość dobrym czasie.

W naszym przykładzie mieliśmy do czynienia z macierzą o rozmiarze $n = 2491$ - mogliśmy sobie pozwolić na przechowywanie jej bez oszczędności pamięci. Przy większych macierzach konieczne wydaje się zastosowanie specjalnej formy przechowywania macierzy - bez pamiętania miejsc na których występują zera.

Różnice w czasie wykonania obliczeń zależą w dużej mierze od sprzętu - uwagę zwraca dość słaby wynik na procesorze Athlon. Na najlepszy wynik - na komputerze z procesorem Pentium M - wpływ mógł mieć 2mb cache procesora.

Błąd powstały podczas obliczeń metodami LU oraz *zerowy wyznacznik* może

być w niektórych przypadkach (gdy np. chodzi o obliczenia związane ze stratami pakietów itp.) zbyt duży. Istotnym wydaje się używanie metod, przy których błąd ten byłby o kilka rzędów mniejszy.

Rozdział 6

Literatura

Stewart W. „Introduction to the Numerical Solution of Markov Chains”;
Princeton University Press; Chichester, West Sussex 1994

**Czachórski T. „Modelowanie kolejkowe w ocenie efektywności sieci
i systemów komputerowych”;**
Pracownia Komputerowa Jacka Skalmierskiego, Gliwice, 1999

Iosifescu M. „Skończone procesy Markowa i ich zastosowania”;
PWN, Warszawa, 1988.

**Bylina B., Górny P. „Metody dokładne w zastosowaniu do roz-
wiązywania łańcuchów Markowa”;**
Varia Informatica, Polskie Towarzystwo Informatyczne; Lublin 2005

**Czachórski T., Fourenau J.M. „Performance Evaluation of an Optimal
Deflection Routing Algorithm on an Odd Torus”;**
Polska Akademia Nauk, 2004

Czachórski T. „Internet a sieci optyczne”;

Instytut Informatyki Teoretycznej i Stosowanej PAN

Szczerbiński Z. „A Markov Chain Model of Optical Packet Formation: Solution Process Acceleration With Parallel Computing”;

Polska Akademia Nauk

Qin Wei „Performance du routage par deflexion pour les rseaux tout optique”;

PhD, University of Versailles, 1999