STUDIA INFORMATICA 2012
Volume 33 Number 3A (107)

Zdzisław ONDERKA
Akademia Górniczo-Hutnicza
Katedra Geoinformatyki i Informatyki Stosowanej

STOCHASTYCZNY MODEL OBCIĄŻENIA SIECI DLA STEROWANIA OBLICZENIAMI ROZPROSZONYMI[1]

Streszczenie. Artykuł prezentuje modele: pierwszy dla aplikacji typu *manager-worker* i dla heterogenicznej sieci komputerowej oraz drugi model opisujący dynamikę zmian obciążenia sieci komunikacyjnej. Dla tych modeli przedstawiono problem stochastycznego sterowania alokacją zadań. Sformułowano również problem MDP (ang. *Markov Decision Control Problem*) dla optymalnej alokacji zadań w sieci komputerowej.

Słowa kluczowe: aplikacja *manager-worker*, stochastyczne sterowanie, obliczenia rozproszone, stochastyczna prognoza, Markowowski Proces Decyzyjny

STOCHASTIC MODEL OF THE NETWORK WORKLOAD FOR THE CONTROL OF THE DISTRIBUTED COMPUTATIONS

Summary. The paper presented the model describing the dynamics of the background load for the communication network. The problem of stochastic control for the task allocation is formulated. Than the Markov Decision Control Problem of optimal task distribution in computer network is formulated. The idea of the open-loop and closed-loop control based on the stochastic forecast are presented.

Keywords: manager-worker application, stochastic control, distributed computations, stochastic forecast, Markov Decision Process

1. Model aplikacji typu *manager-worker* i sieci heterogenicznej

Niech komputerowa sieć heterogeniczna i rozproszona aplikacja będzie układem: $< H, A, F, G_A >$ [1], gdzie: $H=\{M_0,\ldots,M_m\}$ jest zbiorem komunikujących się heterogenicznych kom-

[1] Temat realizowany w ramach prac statutowych Zakładu Geoinformatyki i Informatyki Stosowanej AGH.

puterów, dla których parametry wydajnościowe mogą zmieniać się w czasie; A jest aplikacją typu *manager-worker*, składającą się z części sekwencyjnej t_0 (*manager task*) i części rozpraszanej, składającej się z N identycznych zadań (*workers*); niech $V_D = \{t_1, \ldots t_N\}$ i $G_A = (V,E)$, ($V = \{t_0\} \cup V_D$) jest skierowanym acyklicznym grafem zadań definiującym uporządkowanie wszystkich zadań odwzorowanych na tę samą maszynę (*scheduled task graph*) [1]; bez straty ogólności można założyć, że graf G_A posiada pojedynczy węzeł wejściowy i wyjściowy identyfikowany z zadaniem t_0 oraz że dla zadania t_i istnieje ścieżka z t_0 do t_i i ścieżka od t_i do t_0; F jest odwzorowaniem $F{:}V \rightarrow H$, które definiuje graf G_A i które w przypadku dynamicznej alokacji zadań do sieci H jest zdefiniowane jako sklejenie funkcji częściowych $\{F_n\}_{n=0,1,\ldots}$, to znaczy: $\bigcup_n F_n^{-1}(M_i) = F^{-1}(M_i)\ \ \forall M_i \in M, \quad F^{-1}(M_0) = t_0$.

2. Estymacja stanu obciążenia sieci komunikacyjnej

Rozważana sieć komputerowa składa się z komputerów, których zasoby mogą być dzielone między wielu użytkowników. Dlatego na każdej maszynie różne procesy mogą pojawiać się losowo, generując tak zwane obciążenie tła maszyny. Ponadto, komunikacja użytkowników przez sieć (np. przesyłanie danych dla aplikacji, zdalne wykonywanie aplikacji, zdalna praca na odległym komputerze) generuje zwiększone obciążenie tła sieci komunikacyjnej.

Zarówno obciążenie tła maszyny, jak i obciążenie tła sieci komunikacyjnej zmienia się w czasie, co ma istotny wpływ na efektywność dużej aplikacji rozproszonej. Z tego powodu obciążenie tła maszyny oraz obciążenie tła sieci można estymować w sposób stochastyczny. Zaletą takiego modelu jest możliwość zdefiniowania funkcyjnych zależności pomiędzy systemowymi parametrami wydajnościowymi a parametrami obciążenia. Problem estymacji obciążenia tła maszyny został przedstawiony w pracy [1].

Niech $net_j \in \Re_+$ będzie dodatkowym obciążeniem sieci komunikacyjnej (obciążenie tła sieci) mierzonym na maszynie $M_j \in H$ (obciążenie niepochodzące od wykonywanej aplikacji rozproszonej A). Współczynnik ten to może być mierzona na danej maszynie jako np. ilość pakietów przechodzących przez kartę sieciową danej maszyny. Każda fizyczna wartość net_j odpowiada parametrowi spowolnienia komunikacji $\eta_j \in \Re_+$. Parametr ten jest zdefiniowany jako taki współczynnik, dla którego zachodzi: $\overline{T}_t^j = (1+\eta_j) \cdot T_t^j$, gdzie, $\overline{T}_t^j$ i T_t^j są czasami komunikacji jednego wzorcowego zadania t z procesem *manager* odpowiednio z i bez dodatkowego obciążenia sieci komunikacyjnej (obciążenia tła sieci). Dlatego dalej będzie rozważana bijekcja $\psi_j : \Re_+ \ni net_j \rightarrow \eta_j \in \Re_+$.

Niech skończony zbiór uporządkowany:

$$\Theta_j = \{\theta_j(i) \in \Re_+, i = 0,\ldots,K_j,\ K_j \in N : \forall i = 0,\ldots,K_j - 1\ \theta_j(i) < \theta_j(i+1)\ \} \tag{1}$$

tworzy tak zwany szereg rozdzielczy dla fizycznie mierzonych wartości parametru net_j. Analogicznie niech skończony zbiór uporządkowany

$$\Gamma_j = \{\eta_j(i) \in \Re_+, i = 0,\ldots,K_j,\ K_j \in N : \forall i = 0,\ldots,K_j - 1\ \eta_j(i) < \eta_j(i+1)\ \} \tag{2}$$

zawiera współczynniki spowolnienia komunikacji dla maszyny $M_j \in H$ tworzących szereg rozdzielczy dla współczynników spowolnienia komunikacji oraz

$$\forall M_j \in H\ \forall i = 0,\ldots,K_J \quad \psi_j(\theta_j(i)) = \eta_j(i) \tag{3}$$

Dla każdej maszyny $M_j \in H$ zdefiniowano zbiór $S_j=\{0,\ldots,K_j\}$, który będzie nazywany zbiorem stanów obciążenia sieci. Niech odwzorowanie $\Psi_j : \Re_+ \to S_j$ będzie zdefiniowane jako:

$$\Psi_j(\eta_j) = \begin{cases} 0 & dla \quad \eta_j(0) \le \eta_j \le \eta_j(1), \\ 1 & dla \quad \eta_j(1) \le \eta_j \le \eta_j(2), \\ \ldots & \qquad\qquad \ldots \\ K_j & dla \qquad \eta_j(1) \le \eta_j \le \infty \end{cases} \tag{4}$$

gdzie $\forall j\ \eta_j = \psi_j(net_j)$ przyporządkowuje aktualnej wartości spowolnienia obliczeń, a zatem mierzonej fizycznie wartości obciążenia sieci komunikacyjnej, stanu jego obciążenia.

Maszyna $M_j \in H$ jest w stanie $i \in \{0,\ldots K_j\text{-}1\}$ dla obciążenia tła sieci net_j, jeżeli $\Psi_j \circ \psi_j(net_j) = i$. Dodatkowo, $\forall M_j \in H$ zdefiniujmy bijekcję $\varphi_1^j : S_j \ni i \to \eta_j(i) \in \Gamma_j$.

3. Stochastyczny model opisujący dynamikę zmian obciążenia tła sieci komunikacyjnej

Dla dalszego opisu modelu założono, że istnieją przedziały czasu $\Delta_n = [\tau_n, \tau_{n+1}),\ n = 0,1\ldots$, w których wartości średnie parametrów obciążenia tła maszyny, a w tym również obciążenia tła sieci reprezentują w zadowalający sposób średnie wartości parametrów charakterystyki pracy maszyny w Δ_n. Jako model obciążenia tła sieci komunikacyjnej zdefiniowano układ $\langle \wp, \{X_n\}_{n=0,1,\ldots}, \mathcal{P}, \Gamma, \Theta, S, \varphi_1 \rangle$ [1], gdzie:

- $\wp$ jest układem $<\wp_1,\ldots, \wp_m>$, gdzie $\forall j = 1,\ldots,m\ \wp_j = (\Omega, \Im, P^j)$ jest przestrzenią prawdopodobieństwa, dla której, Ω jest zbiorem zdarzeń elementarnych polegających na tym, że maszynie M_j odpowiada pewna średnia wartość współczynnika spowolnienia komunikacji w przedziale czasu Δ_n,

- $\{X_n\}_{n=0,1,\ldots}$ jest układem $<\{X_n^1\}_{n=0,1,\ldots}, \ldots, \{X_n^m\}_{n=0,1,\ldots}>$ niestacjonarnych dyskretnych stochastycznych łańcuchów Markowa [2,3], które opisują dynamikę zmian stanu obciążenia tła sieci komunikacyjnej dla każdej maszyny $Mj \in H$ w czasie. Dynamika ta jest opisana wzorem:

$$P^k(X_{n+1}^k = j \,|\, X_n^k = i) = P^k(X_{n+1}^k = j \,|\, X_0^k = i_0, \ldots, X_{n-1}^k = i_{n-1}, X_n^k = i) \tag{5}$$

- $\mathcal{P}$ jest układem $< \mathcal{P}^1, \ldots, \mathcal{P}^m >$, gdzie $\mathcal{P}^k = \{P_n^k\}_{n=0,1,\ldots}$ jest ciągiem Markowowskich macierzy przejść. $p_{ij}^k(n) = P_n^k(X_{n+1}^k = j \,|\, X_n^k = i)$ jest prawdopodobieństwem, że proces będzie mieć wartość stanu j obciążenia tła sieci w kroku $n+1$, zakładając, że ma wartość stanu i w kroku n dla maszyny M_k. Jeżeli oznaczymy przez $\Pi(n) = (\Pi^1(n), \ldots, \Pi^m(n))$ wektor rozkładów prawdopodobieństwa stanu dla komputerów sieci w n-tym kroku dla sieci H oraz $\Pi^j(\mu) = \beta_\mu = (\beta_\mu^1, \ldots, \beta_\mu^m)$ oznaczymy początkowy rozkład prawdopodobieństw, to kolejne rozkłady mogą zostać obliczone dla każdej maszyny M_j według wzoru: $\Pi^j(n) = \Pi^j(n-1) \cdot P_n^j$, $n > \mu$ [2,3],
- Γ jest układem $<\Gamma_1, \ldots, \Gamma_m>$, gdzie Γ_j jest szeregiem rozdzielczym dla parametrów spowolnienia komunikacji (zdefiniowanym wzorem 2),
- Θ jest układem $<\Theta_1, \ldots, \Theta_m>$, gdzie Θ_j jest szeregiem rozdzielczym dla fizycznie mierzonych parametrów wydajnościowych dla maszyny M_j (zdefiniowanym wzorem 1),
- S jest układem $<S_1, \ldots, S_m>$, gdzie S_j jest zbiorem stanów obciążenia sieci komunikacyjnej dla maszyny $Mj \in H$ i $card(S_j) = card(\Gamma_j)$ $\forall j$,
- φ_1 jest układem $<\varphi_1^1, \ldots \varphi_1^m>$ oraz $\forall j = 1, \ldots, m$ φ_1^j jest odwzorowaniem $\varphi_1^j : S_j \to \Gamma_j$.

Mając proces stochastyczny $\{X_n^j\}_{n=0,1,\ldots}$ oraz dwie bijekcję φ_1 oraz φ_2 ($\varphi_2 = <\varphi_2^1, \ldots \varphi_2^m>$, gdzie $\varphi_2^j : \Gamma_j \to \Re_+$, $\varphi_2^j(\eta_j) = (1 + \eta_j) \cdot T_t^j$), możemy zdefiniować nowy proces stochastyczny $\{\overline{T_t^j}(n)\}_{n=0,1,\ldots}$ opisujący dynamikę zmian czasów komunikacji pojedynczego zadania $t \in V_D$ z zadaniem t_0 na maszynie Mj $\forall j = 1, \ldots, m$ $n=0,1,\ldots$

$$\overline{T_t^j}(n) = \varphi_2^j \circ \varphi_1^j(X_n^j) = (1 + \eta_j(X_n^j)) \cdot T_t^j \tag{6}$$

4. Stochastyczne sterowanie oparte na modelu obciążenia sieci

Polityką sterowania nazywamy ciąg funkcji $\{F_n\}_{n=\mu,\mu+1,\ldots}$, gdzie indeks $\mu \geq 0$ oznacza chwilę czasu τ_μ (krok czasowy niestacjonarnego procesu stochastycznego, opisującego dynamikę zmian obciążenia sieci), w której rozpoczyna się stosowanie polityki sterowania wykonaniem aplikacji rozproszonej.

Niech $\overline{\mathrm{U}}_\mu$ oznacza zbiór wszystkich polityk sterowania. Podzbiór $U_\mu \subset \overline{U}_\mu$ nazywa się *zbiorem polityk dopuszczalnych*, jeśli każda polityka $u \in U_\mu$. spełnia zdefiniowane ograniczenia. Przy założeniu że kolejne decyzje realizowanej polityki są podejmowane w chwilach czasowych τ_μ , $n = \mu, \mu+1, \ldots$, to dla opisywanego modelu sterowania zdefiniowano politykę jako dopuszczalną, jeżeli zachodzi następujący warunek:

$$\exists q \in (0,1],\ \forall j = 1,\cdots,m\ \forall n = \mu,\ldots\ \ q \le P_n^j\left(\left(\ \overline{T}_t^j(n) \cdot card(F_n^{-1}(M_j))\ \right) < \Delta_n\ \right) \le 1 \qquad (7)$$

gdzie $P(\cdot)$ oznacza prawdopodobieństwo.

Niech $c^j(n, s^j, u)$ oznacza tak zwaną *funkcję kosztu natychmiastowego* [2], która jest związana z pojedynczym krokiem $n=\mu,\mu+1,\ldots$ realizowanej polityki $u \in U_\mu$ dla maszyny Mj będącej w stanie $s^j \in S_j$. Załóżmy, że dla $\forall u \in U_\mu$ oraz $\forall \mu$ i $\tau_n \ge \tau_\mu$ oraz dla dowolnego rozkładu początkowego β_μ^j (w chwili τ_μ) stanu obciążenia tła sieci dla maszyny M_j i dowolnego stanu obciążenia sieci w chwili czasowej τ_n istnieje wartość oczekiwana $E_{\beta_\mu^j}(c^j(n, s^j, u))$ [2,3].

Niech $C_n(\beta_\mu, s, u)$ oznacza *funkcję kosztu* dla n-tego kroku $n=\mu,\mu+1,\ldots,\mu+Z-1$, $Z \in N, Z < \infty$ realizowanej polityki $u \in U_\mu$, jednocześnie jest to koszt dla funkcji F_n stosowanej w przedziale czasu $[\tau_n, \tau_{n+1})$, $n = \mu, \mu+1, \ldots \mu+Z-1$, (ostatnia decyzja jest podejmowana w chwili czasowej $\tau_{\mu+Z-1}$). Założenie $Z < \infty$ oznacza, że jest rozpatrywany *problem sterowania z horyzontem skończonym*.

Koszt $C_n(\beta_\mu, s, u)$ jest zdefiniowany dla dowolnego rozkładu początkowego $\beta_\mu = (\beta_\mu^1, \ldots, \beta_\mu^m)$ w początkowej chwili czasowej τ_μ oraz osiągniętego stanu $s = (s^1, \ldots, s^m)$, $s^j \in S_j$ w τ_n jako: $C_n(\beta_\mu, s, u) = \max\limits_{M_j \in H}(c^j(n, s^j, u))$.

Niech $C^Z(\beta_\mu, u)$ oznacza *całkowity koszt* dla sieci H dla realizowanej polityki $u \in U_\mu$ w przedziale czasowym $[\mu,Z]$ (skończonym horyzoncie, $Z<\infty$) zdefiniowanym jako:

$$C^Z(\beta_\mu, u) = E_{\beta_\mu}\left(\sum_{n=\mu}^{\mu+Z+1} C_n(\beta_\mu, s, u)\right) \qquad (8)$$

Wtedy można zdefiniować problem sterowania jako: *znaleźć taką politykę dopuszczalną* $u \in U_\mu$, *która minimalizuje funkcję kosztu* $C^Z(\beta_\mu, u)$ *w zbiorze* U_μ.

5. Problem sterowania w języku teorii MDP

Opisany problem sterowania został oparty na teorii Markowowskich Procesów Decyzyjnych (*Markov Decision Process* – MDP) [2]. Niech $\mathcal{T}= \{\tau_{\mu,} \ldots, \tau_{\mu+Z-1}\}$ będzie zbiorem kolejnych chwil czasowych dla horyzontu $Z\in\mathcal{N}$, $Z<\infty$, w których są podejmowane kolejne decyzje realizowanej polityki sterowania (np. przez jakiś proces lub agenta decyzyjnego), odpowiadające początkom przedziałów czasowych Δ_n, $n=\mu,\mu+1\ldots$ Zakładamy, że ostatnia decyzja jest podejmowana w chwili $\tau_{\mu+Z-1}$.

W każdej chwili czasowej proces decyzyjny otrzymuje informacje o stanie $s=\{s^1,\ldots,s^m)\in S$ dla sieci H, w którym może wykonać akcję $a=(a^1,\ldots,a^m)$ ze wszystkich możliwych akcji w tym stanie $A_s = A^1_{s^1}\times\ldots\times A^m_{s^m}$, gdzie $A^k_{s^k}$ jest dyskretnym i skończonym zbiorem akcji, które mogą być podejmowane, jeżeli stan maszyny M_k jest s^k. Ponadto, zakładamy, że akcje są podejmowane w sposób deterministyczny oraz zakładamy, że $\mathcal{A}=\bigcup_s A_s$ oraz $\mathcal{A}=\mathcal{A}^1\times\ldots\times\mathcal{A}^m$, gdzie $\mathcal{A}^k=\bigcup_{s^k\in S_k} A^k_{s_k}$.

Rezultatem wyboru akcji $a_k\in A^k_s$ w stanie s^k w chwili czasowej τ_i dla maszyny M_k jest *funkcja kosztu natychmiastowego* $c^k(i,s^k,u)$ oraz stan maszyny M_k w chwili czasowej τ_{i+1} jest określony rozkładem prawdopodobieństwa $p^k_{\tau_i}(\cdot\mid s^k,a^k)$ i $\sum_{j\in S} p^k_{\tau_i}(j\mid s^k,a^k)=1$.

Procedurę wyboru akcji w każdej chwili czasowej τ_i nazywamy *prawem decyzyjnym.* Prawa decyzyjne mogą być deterministyczne markowowskie albo losowe zależne od historii, w zależności od tego, jak wykorzystują poprzednie informacje oraz jak dokonują wyspecyfikowania akcji. W opisywanym modelu będą stosowane deterministyczne markowowskie prawa decyzyjne, które są zdefiniowane jako funkcje $d_\tau : S\ni s\to d_\tau(s)\in A_s$, gdzie jeżeli $s=(s^1,\ldots,s^m)$, to $d_\tau(s)=(d^1_\tau(s^1),\ldots,d^m_\tau(s^m))$. Przez całkowite obciążenie sieci będzie rozumiane obciążenie tła sieci oraz obciążenie pochodzące od komunikującego się z innymi procesami (na innych maszynach lub innej maszynie) zadania aplikacji. Polityką lub strategią sterowania nazywamy ciąg praw decyzyjnych $u=(d_0,d_1,\ldots,d_{Z-1})$.

W dalszym ciągu będziemy rozważać w celu sterowania obliczeniami rozproszonymi markowowskie deterministyczne polityki dla skończonego horyzontu, w których podejmowana decyzja, w każdej chwili czasowej jest podejmowana na podstawie poprzedniego stanu obciążenia całkowitego sieci oraz wybranej decyzji w tym stanie i każda akcja jest podejmowana z prawdopodobieństwem równym 1.

Polityka dopuszczalna została zdefiniowana w rozdziale 4, przy czym:

$$\{d_n(s_n)\}_{n=0,1,\ldots Z-1} = \{a_n\}_{n=\mu,\mu+1,\ldots} = \{F_n\}_{n=\mu,\mu+1,\ldots} \text{ i } s_n = (s_n^1,\ldots,s_n^m),\ a_n = (a_n^1,\ldots,a_n^m) \tag{9}$$

Przestrzeń zdarzeń elementarnych Ω^H dla sieci H ma w tym przypadku następującą postać: $\Omega^H = \{S\times\mathcal{A}\}^{Z\text{-}1}\times S$ oraz zdarzenie elementarne $\omega^H \in \Omega^H$ jest ciągiem:

$$(s_0, a_0, s_1, a_1, \ldots a_{Z-1}, s_Z)\ \forall s_n = (s_n^1,\ldots,s_n^m),\ a_n = (a_n^1,\ldots,a_n^m) \tag{10}$$

Zmienne losowe $\overline{X}_n(\omega^H),\ n = 0,1,\ldots$ zdefiniowane jako odwzorowania $\overline{X}_n : \Omega^H \to S$, $\overline{X}_n(\omega^H) = s_n$, gdzie n odpowiada chwili czasowej τ_n (i przedziałowi czasu Δ_n), opisują dynamikę zmian obciążenia całkowitego sieci komunikacyjnej dla sieci komputerowej H. Dodatkowo, zdefiniowano ciąg wektorów losowych $Y_n(\omega^H)$, które dla każdego n przyjmują wartości: $Y_n(\omega^H) = a_n,\ a_n \in A_{s_n},\ \overline{X}_n(\omega^H) = s_n$.

Analogicznie jak dla sieci komputerowej można zdefiniować dla maszyny $M_k\in H$ zbiór zdarzeń elementarnych jako $\Omega = \{S_k\times\mathcal{A}^k\}^{Z\text{-}1}\times S_k$. Ciąg zmiennych losowych $\overline{X}_n^k(\omega)$, $\omega\in\Omega,\ n = 0,1\ldots,\ k = 1,\ldots m$ zdefiniowanych jako $\overline{X}_n^k : \Omega \to S^k$ i $\overline{X}_n^k(\omega), = s_n^k$, gdzie s_n^k jest stanem całkowitego obciążenia sieci komunikacyjnej. Natomiast zmienne losowe $Y_n^k(\omega)$ dla każdego n przyjmują wartości: $Y_n^k(\omega) = a_n^k,\ a_n^k \in A_{s_n^k}^k,\ \overline{X}_n^k(\omega) = s_n^k$.

Jeżeli przez h_n^k oznaczymy historię dla maszyny $M_k\in H$, tj.

$$h_n^k = (\overline{X}_0^k = s_0^k, Y_0^k = a_0^k, \ldots, Y_{n-1}^k = a_{n-1}^k, \overline{X}_n^k = s_n^k),\quad \mu \le \tau_n \le \mu + Z - 1 \tag{11}$$

to dynamika zmian całkowitego obciążenia sieci komunikacyjnej pochodzącego od obciążenia tła sieci oraz obciążenia pochodzącego od komunikującego się zadania aplikacji (sterowanej zgodnie z polityką $u \in U_\mu$) wykonywanego na maszynie M_k wyraża się wzorem:

$$\begin{aligned} &P_u^k\left(\overline{X}_{n+1}^k = j \,\middle|\, h_n^k = (h_{n-1}^k, a_{n-1}^k, i), Y_n^k = a_n^k\right) = \\ &P_u^k\left(\overline{X}_{n+1}^k = j \,\middle|\, \overline{X}_n^k = i, Y_n^k = a_n^k\right) = p_{\tau_n}^k(j \mid i, a_n^k) \end{aligned} \tag{12}$$

Jednocześnie zachodzi następująca równość: $P_u^k\left(Y_n^k = a_n^k \,\middle|\, h_n^k\right) = 1$.

Rozkład prawdopodobieństw $\overline{\Pi}^j(\mu+i+1)$ dla zmiennych losowych $\overline{X}_{\mu+i+1}^k$ w chwili czasowej $\tau_{\mu+i+1}$ można obliczyć według wzoru:

$$\overline{\Pi}^j(\mu+i+1) = \sum_{j\in S_k} p_{\tau_{\mu+i}}^k(j \mid s_j^k, a_j^k)\cdot\overline{\Pi}^k(\mu+i) \tag{13}$$

gdzie $\overline{\Pi}^k(\mu+i)$ jest rozkładem prawdopodobieństw w chwili czasowej $\tau_{\mu+i}$.

Rozkłady prawdopodobieństw $\overline{\Pi}^k(\mu+i)$ mogą być obliczane na podstawie rozkładów $\Pi^k(\mu+i)$ stanu obciążenia tła sieci komunikacyjnej. Oznaczmy przez $r_i^k = \eta_k(i+1) - \eta_k(i)$, $\eta_k(i) \in \Gamma_k$ oraz δ^k niech będzie średnim obciążeniem sieci pochodzącym od komunikującego się zadania aplikacji. Załóżmy dodatkowo, że $\forall i\ \forall k\ \ \delta^k < r_i^k$ oraz $\alpha_i^k = \frac{\delta^k}{r_i^k}$.

Jeżeli net_k^c oznacza fizycznie mierzony parametr całkowitego obciążenia sieci komunikacyjnej, a net_k oznacza fizycznie mierzoną wartość obciążenia tła sieci komunikacyjnej oraz $\eta'_k(i) = \eta_k(i+1) - \delta^k$ dla i=0,1,…,K-1, to:

$$\overline{X}_n^k = \Psi_k(\psi_k(net_k^c)) = \begin{cases} 0 & jezeli \quad net_k \in [\eta_k(0), \eta'_k(0)) \\ i & jezeli \quad net_k \in [\eta'_k(i-1), \eta_k(i)) \cup [\eta_k(i), \eta'_k(i)) \\ K & jezeli \quad net_k \in [\eta'_k(K-1), \eta_k(K)) \cup [\eta_k(K), \eta'_k(K)) \end{cases} \tag{14}$$

Wtedy przy założeniu równomiernych rozkładów wewnątrz szeregu rozdzielczego i korzystając z prawdopodobieństwa geometrycznego, rozkłady prawdopodobieństwa $\overline{\Pi}^k(\mu+i)$ można obliczyć według wzoru w postaci macierzowej:

$$(\overline{\Pi}^k(n))^T = B \cdot (\Pi^k(n))^T = B \cdot (P^k(n) \cdot \Pi^k(n-1))^T \tag{15}$$

gdzie operacja macierzowa $(\cdot)^T$ oznacza transponowanie macierzy oraz macierz B ma następującą postać:

$$\begin{bmatrix} 1-\alpha_1 & 0 & 0 & \cdots \\ \alpha_1 & 1-\alpha_2 & 0 & \cdots \\ 0 & \alpha_2 & \ddots & \\ \vdots & \cdots & \alpha_{K-1} & 1 \end{bmatrix} \tag{16}$$

Łatwo zauważyć, że rozkład prawdopodobieństwa $\overline{\Pi}^k(n)$ obliczany za pomocą powyższego układu równań spełnia równość $\sum_{i=0}^{K} \overline{\Pi}_i^k(n) = 1$.

Funkcja kosztu dla skończonego horyzontu Z dla sieci H dla deterministycznych Markowowskich polityk $u \in U_\mu$ jest zdefiniowany wzorem:

$$C^Z(\beta_\mu, u) = E_{\beta_\mu}^u \left(\sum_{n=\mu}^{\mu+Z+1} C_n(n, \overline{X}_n, d_n(h_n)) \right) \tag{17}$$

gdzie $d_n(h_n) = Y_n$, $h_n = (h_{n-1}, a_{n-1}, s)$ oraz $P_n^k(Y_n^k = a_n^k \mid s_n^k) = 1$ i $\beta_\mu = (\beta_\mu^1, \ldots, \beta_\mu^m)$.

Teraz problem sterowania (zdefiniowany w rozdziale 4) polega na znalezieniu takiej polityki $u \in U_\mu$, która minimalizuje funkcję kosztu zdefiniowaną wzorem (17).

Dla problemu sterowania w teorii Markowowskich Procesów Decyzyjnych można też zastosować inne funkcje kosztu:

- Przewidywany koszt dyskontowy dla współczynnika 0<α<1:

$$C_{\alpha}^{Z}(\beta_{\mu},u)=(1-\alpha)\cdot\sum_{n=\mu}^{\mu+Z+1}\alpha^{n-\mu}C_{n}(n,\overline{X}_{n},d_{n}(h_{n}))$$

- Przewidywany koszt średni: $C_{ave}{}^{Z}(\beta_{\mu},u)=C^{Z}(\beta_{\mu},u)/Z$

- Koszt średni (jako zmienna losowa): $C_{\lim}^{Z}(\beta_{\mu},u)=\limsup\limits_{Z\to\infty}\sum\limits_{n=\mu}^{\mu+Z+1}\alpha^{n-\mu}C_{n}(n,\overline{X}_{n},d_{n}(h_{n}))$

6. Polityki sterowania aplikacją

6.1. Sterowanie w układzie otwartym

Polityka sterowania w układzie otwartym (*Open Loop Control*) jest oparta na modelu przedstawionym w rozdziale 4, tzn. jest oparta na prognozie średniego stanu obciążenia tła sieci komunikacyjnej generowanej tylko raz w początkowej chwili czasowej τ_{μ} dla całego horyzontu czasowego. Dokładniej, polityka sterowania polega na obliczeniu rozkładów prawdopodobieństwa $\Pi_{i}^{k}(n)$, mając dany rozkład początkowy β_{μ}^{k} dla każdej maszyny $M_{k}\in H$ rekurencyjnie według wzoru:

$$\Pi_{i}^{k}(n)=\sum_{j\in S}p_{ij}^{k}(n-1)\cdot\Pi_{j}^{k}(n-1),\ \tau_{n}>\tau_{\mu} \qquad (18)$$

Następnie mając obliczony wektor prawdopodobieństw dla średnich stanów obciążenia tła sieci komunikacyjnej dla każdego kroku czasowego, można obliczyć spodziewany czas komunikacji zadania (z procesem *manager*) wykonywanego na każdej maszynie, co prowadzi do obliczenia dla każdego kroku *współczynnika mocy* λ_j dla każdej maszyny M_j jako $\lambda_{j}=1/\overline{T}_{t}^{j}$. Kolejne decyzje polityki sterownia alokują podzbiór zadań $F^{-1}(M_{k})$, których liczności zbiorów są proporcjonalne do spodziewanego współczynnika mocy maszyny λ_j dla $n=\mu,\mu+1,\ldots,\mu+Z-1$. Wartość kosztu $C_{n}(\beta_{\mu},s_{n},u)$ dla jednego kroku realizowanej polityki $u=\{F_{n}\}_{n=\mu,\mu+1,\ldots}$ jest następująca:

$$C_{n}(\beta_{\mu},s_{n},u)=\begin{cases}\max_{F_{n}(M_{k})}\{v^{k}(n)\cdot(1+\eta_{k}(j))\cdot T_{t}^{k}\} & dla\ n=\mu\\ \max_{F_{n}(M_{k})}\{v^{k}(n)\cdot\sum_{j=0}^{K}(\Pi_{j}^{k}(n)\cdot(1+\eta_{k}(j))\cdot T_{t}^{k})\} & dla\ n>\mu\end{cases} \qquad (19)$$

gdzie $s_{n}=(X_{n}^{1},\ldots,X_{n}^{m})$, $v^{k}(n)=card(F_{n}^{-1}(M_{k}))$ oraz ponieważ u jest polityką dopuszczalną, $\forall k=1,\ldots m$ funkcje $F_{n}^{-1}(M_{k})$ spełniają ograniczenia (7).

6.2. Sterowanie w układzie zamkniętym

W przypadku sterowania w układzie zamkniętym (*Closed Loop Control*) wybór pierwszej decyzji (tzn. w chwili czasowej τ_n) jest identyczny jak w przypadku sterowania w układzie otwartym. Różnica polega na tym, że decyzje (dotyczące alokacji zadań do maszyn sieci) w kolejnych krokach czasowych τ_n, $n = \mu, \mu + 1, \ldots, \mu + Z + 1$ są oparte na prognozie średnich stanów obciążenia całkowitego sieci komunikacyjnej dla każdej maszyny w sieci H i są obliczane w każdym kroku czasowym. Czyli obliczenia prognozowanego czasu komunikacji pojedynczego zadania z procesem *manager* są wykonywane kolejno za każdym razem przed wykonaniem kolejnego kroku czasowego (czyli na jeden kolejny krok czasowy).

Wartość oczekiwana czasu komunikacji pojedynczego zadania na maszynie M_k w każdym kolejnym kroku czasowym $n = \mu + 1, \mu + 2, \ldots$ jest następująca:

$$E_{\beta_\mu^k}(\overline{T}_t^k(n)) = E_{\beta_\mu^k}(\varphi_2^k \circ \varphi_1^k(\overline{X}_n^k)) = \sum_{j=0}^{K}(\overline{\Pi}_j^k(n) \cdot (1+\eta_k(j)) \cdot T_t^k) \tag{20}$$

oraz oczekiwany czas komunikacji dla $v(a_n^k) = card(F_n^{-1}(M_k))$ zadań na maszynie M_k w kroku n jest równy $v(a_n^k) \cdot E_{\beta_\mu^k}(\overline{T}_t^k(n))$, a więc wartość kosztu $C_n(\beta_\mu, s_n, u)$ dla każdego kroku realizowanej polityki $u = \{a_n\}_{n=\mu+1,\mu+2,\ldots} = \{F_n\}_{n=\mu+1,\mu+2,\ldots}$ jest następujący:

$$C_n(\beta_\mu, s_n, u) = \max_{a_n^k \in A_{s_n^k}^k} \left\{ v(a_n^k) \cdot \sum_{j=0}^{K}(\overline{\Pi}_j^k(n) \cdot (1+\eta_k(j)) \cdot T_t^k) \right\} \tag{21}$$

gdzie $s_n = (s_n^1, \ldots, s_n^m)$, $a_n = (a_n^1, \ldots, a_n^m)$ oraz ponieważ u jest polityką dopuszczalną, $\forall n \ \forall k \ a_n^k$ musi spełniać ograniczenia (7). Dla pierwszego kroku czasowego (tj. $n=\mu$) koszt wyraża się wzorem (20).

7. Podsumowanie i wnioski

Na podstawie przyjętych założeń dotyczących komputerowej sieci heterogenicznej i aplikacji rozproszonej zdefiniowano stochastyczny model opisujący dynamikę zmian obciążenia tła sieci komunikacyjnej. Sformułowano dwa problemy sterowania alokacją zadań. Pierwszy w układzie otwartym oparty na jednokrotnym przygotowaniu prognozy dla całego horyzontu czasowego i drugi oparty na teorii Markowoskich Procesów Decyzyjnych, w którym oblicza się prognozę w każdym kolejnym kroku czasowym, biorąc pod uwagę stan obciążenia sieci i decyzję podjętą w tym stanie.

Sterowanie w układzie zamkniętym zwykle będzie lepsze niż w układzie otwartym, dla horyzontu dłuższego niż jeden krok czasowy, gdyż dostarcza lepszej estymacji obciążenia dla każdego kolejnego kroku czasowego. Sterowanie w układzie otwartym generuje prognozę tylko raz w początkowej chwili czasowej dla całego horyzontu czasowego. Sterowanie w układzie zamkniętym dynamicznie sprawdza stan obciążenia sieci i sterowanie oparte jest nie tylko na podstawie stanu obciążenia aktualnego sieci (jak w zwykłych procesach markowowskich [3]), ale również na podstawie podjętej decyzji.

W opisanych politykach sterowania użyto statystykę λ_j zdefiniowaną jako $\lambda_j = 1/\overline{T}_t^j$. Można również zastosować inne statystyki, jak np. $1/x_p^j$, gdzie x_p^j jest kwantylem $\overline{T}_t^j$ obliczanym przy danym poziomie ryzyka 1-*p* oraz w ograniczeniu dla zbioru polityk dopuszczalnych *q*=1-*p* [2] lub statystykę $1/(\alpha \cdot \overline{T}_t^j(n) + (1-\alpha) \cdot x_j^p)$, gdzie $\overline{T}_t^j(n)$ jest aktualną wartością procesu stochastycznego oraz $\alpha \in [0,1]$ i w ograniczeniu polityk dopuszczalnych $q = (1-\alpha)(1-p) + \alpha * 0.5$ (jest to wtedy tak zwana polityka *mieszana inercyjno-markowowska*).

Oczywiście, dla opisanego sterowania niezbędne jest zaprojektowanie i zaimplementowanie odpowiedniego systemu agentowego dla śledzenia parametrów obciążenia sieci na każdym komputerze, gdzie narzut komunikacyjny pomiędzy agentami jest znikomy. Podobny system dla śledzenia obciążenia komputerów został opisany w pracy [4].

Niestety, opisane modele teoretyczne mogą mieć sensowne zastosowanie tylko w przypadku, gdy czas komunikacji procesów wykonujących obliczenia (*workers*) z procesem głównym (*manager*) istotnie wpływa na czas całkowity obliczeń aplikacji rozproszonej. Nie mają zastosowania, gdy narzut komunikacyjny związany z rozpraszaniem obliczeń jest niewielki (niewielki wpływ na wartość przyśpieszenia obliczeń).

Jednakże opisane modele można zastosować nie tylko do sterowania obliczeniami rozproszonymi wykonywanymi w kilku krokach czasowych przy dużym narzucie komunikacyjnym, w których następuje alokacja wielu procesów do każdej maszyny w sieci, ale również dla przypadku komunikacji w sieci dwóch procesów, np. klient-serwer (jeden klient i wiele serwerów), gdzie proces klienta na podstawie prognozy obciążenia sieci komunikacyjnej będzie wybierać któremu serwerowi zlecić obsługę, lub w przypadku obiektów rozproszonych [5], z którego obiektu zdalnego wywołać metodę, lub w przypadku rozproszonych aplikacji bazodanowych [6].

Dlatego, kolejnym krokiem rozwoju opisanych modeli będzie połączenie modelu opisanego w pracy [1] z modelem opisanym w tym artykule oraz sformułowanie nowych algorytmów tak w układzie otwartym, jak i zamkniętym oraz analiza optymalności zdefiniowanego sterowania.

BIBLIOGRAFIA

1. Onderka Z.: Stochastic Control of the Scalable High Performance Distributed Computations [in:] Wyrzykowski R. et al. (eds.): PPAM 2011, Part II, LNCS 7204, Springer, Heidelberg 2012, s. 181÷190.
2. Puterman M. L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming, Wiley Series in Probability and Statistics, 2005.
3. Astrom K. J.: Introduction to Stochastic Control Theory, Dover Publications, 2006.
4. Lepiarz M., Onderka Z.: Agents System for Load Monitoring of the Heterogeneous Computer Network, PPAM'2001, LNCS 2328, 2001, s. 364÷368.
5. Cichy M., Onderka Z.: The comparision of the communication performance for the CORBA and DCOM standards in the client-server systems, (in polish) Conference Computer Networks 2011 (SK'11), Studia Informatica, 2011.
6. Piórkowski A.: Metody tworzenia rozproszonych aplikacji bazodanowych w technologii .NET [w:] Sieci komputerowe: Aplikacje i zastosowania, WKŁ, Warszawa 2007, t. 2, s. 195÷202.

Wpłynęło do Redakcji 18 marca 2012 r.

Abstract

The paper presented two models: first one for the master-slave application and for the heterogeneous computer network and the second one the model describing the dynamics of the background load for the communication network. For these models, two control problems of the task allocation were formulated. First one, the problem of stochastic control based on the forecast of network background load generated for the whole horizon only at the starting time epoch. The second one, the Markov Decision Control Problem of optimal task distribution, where the forecast of the network whole workload is generated in each consecutive time epoch of the realized policy. The idea of the open-loop and closed-loop control based on the stochastic forecast are presented.

Adres

Zdzisław ONDERKA: Akademia Górniczo-Hutnicza, Katedra Geoinformatyki i Informatyki Stosowanej, WGGiOŚ, Al. Mickiewicza 30, 30-059, Kraków, Polska, zonderka@agh.edu.pl, onderka@ii.uj.edu.pl