

УДК 004.6

О.М. Чаузов

Інститут СЗІ НТУ України «КПІ» імені Ігоря Сікорського, Київ

МАТЕМАТИЧНА МОДЕЛЬ РОЗПОДІЛУ ІНФОРМАЦІЙНОГО РЕСУРСУ МІЖ ТРАНЗАКЦІЯМИ ДО СХОВИЩ ДАНИХ

Розглядаються принципи побудови математичних моделей, що дозволяють провести оптимальний розподіл інформаційного ресурсу системи між транзакціями до сховищ даних (СД). Розглянута модель використовує базову модель інформаційно-телекомунікаційної мережі, що підтримує АСУ спеціального призначення. В моделі обмін здійснюється інформаційними блоками (ІБ), що надходять із сховищ даних, та у загальному випадку мають різний об'єм. Крім того, у моделі враховуються ймовірності звертання до ІБ СД.

Ключові слова: інформаційний ресурс, математична модель, інформаційно-телекомунікаційна мережа, розподілені дані.

Вступ

Особливості розподілу інформаційного ресурсу. Система планування розміщення інформаційного ресурсу у розподіленому середовищі інформаційно-телекомунікаційної мережі (ІТМ) АСУ спеціального призначення (СП) повинна забезпечувати оптимальне значення обраної цільової функції шляхом складання плану розподілу ресурсів між запитами із вхідного нестационарного потоку, враховуючи при цьому особливості завдань, котрі вирішуються системою.

Підвищення якості обслуговування користувачів ІТМ АСУ СП по обробці запитів у частині, що стосується результативності обробки запитів, потребує розробки розподіленої системи планування, що опирається на використання моделей запитів і потоку запитів, а також обчислювальної підсистеми. Розподілений характер ІТМ АСУ СП визначає необхідність використання розподіленої системи планування.

Аналіз останніх досліджень. У ряді літературних джерел [1 – 8] описані методи планування процесу обслуговування за різноманітними дисциплінами. Зокрема, у [3] показано порядок планування інформаційного ресурсу для АСУ промислового підприємства для розподілених гетерогенних систем, однак в них не враховується нестационарність потоку запитів. У [6] при розподілу інформаційного ресурсу проаналізовані нестационарні потоки, але тільки для негетерогенних систем. Також у більшості робіт [1, 2, 4, 5, 7 – 9] відсутній аналіз втрат системи при несвоєчасному виконанні деяких запитів.

В даній статті розглянемо один із варіантів вирішення даного завдання – розподілити інформаційний ресурс системи між транзакціями до сховищ даних. **Мета статті** – при розробці відповідної математичних моделей врахувати як різний об'єм інформаційних блоків, так і розподіл ймовірностей кількості звернень до різних блоків.

Результати досліджень

Розглянемо модель інформаційно-телекомунікаційної мережі (ІТМ) підтримки АСУ спеціального призначення, що характеризується інформаційними блоками (ІБ) різного об'єму, причому ймовірності звертання до деяких з них можуть бути однаковими. При створенні сховищ даних (СД) великої інформаційної ємності потрібне виділення інформаційного ресурсу, надаваного різними типами вузлів ІТМ, що відрізняються своїми характеристиками.

Розглянемо транзакцію до СД, формовану на основі інформації, обраної з M інформаційних блоків об'єму W_i ($i = \overline{1, M}$). Ймовірність звертання до i -го ІБ дорівнює p_i , де $\left(\sum_{i=1}^M p_i = 1\right)$. Для обробки та розміщення ІБ розподіленого СД може бути надано N різних типів вузлів ІПС, характеристики яких визначаються як t_j, f_j ($j = \overline{1, N}$). Визначимо показник відносної частоти звертання i -го ІБ як:

$$\omega_i = \frac{p_i}{W_i}, \quad i = \overline{1, M}. \quad (1)$$

Вочевидь, що середній час доступу до M i -го ІБ довільного об'єму, для розміщення та обробки яких інформаційний ресурс надають N типів вузлів ІПС, буде мінімальним, якщо завантажувати ІБ, починаючи з вузлів, що мають найбільшу швидкодію в порядку зменшення величини ω_i .

Пронумеруємо типи вузлів ІПС в порядку зменшення швидкодії, тобто

$$t_j \leq t_{j+1}, \quad j = \overline{1, N-1},$$

а ІБ – за убаванням відносної частоти об'їгу, тобто

$$\omega_i \geq \omega_{i+1}, \quad i = \overline{1, M-1}.$$

Позначимо k_j – кількість ІБ, розміщених у типах вузлів ІПС із номерами $1, \dots, j$. Кількість ІБ, ро-

зміщених в j -му типі вузла, визначається як $k(j) = k_j - k_{j-1}$, при цьому $k_0 = 0, k_N = M$.

Розглянемо дискретну випадкову величину Θ – задіяний ресурс для обробки та розміщення ІБ із можливими значеннями $\theta \in \{y_1, \dots, y_N\}$, де

$$y_j = \sum_{i=1}^{k_j} W_i, \quad j = \overline{1, N}. \quad (2)$$

Тоді сумарний обсяг задіяних вузлів j -го типу визначається як:

$$y_j^0 = \sum_{i=k_{j-1}+1}^{k_j} W_i, \quad j = \overline{1, N}. \quad (3)$$

Для всіх типів вузлів можна розглядати N -вимірний вектор $\bar{y}^0 = \{y_1^0, \dots, y_N^0\}$.

Можна відмітити, що змінні y_j та y_j^0 пов'язані відношенням:

$$y_j = y_1^0 + \dots + y_j^0, \quad j = \overline{1, N}.$$

З (1) та (3) витікає, що сумарний обсяг задіяних вузлів j -го типу однозначно визначається за допомогою змінних y_j таким чином:

$$y_j^0 = y_j - y_{j-1}, \quad j = \overline{1, N}, \quad (4)$$

якщо покласти $y_0 = 0$.

Функція розподілу випадкової величини Θ має такий вигляд:

$$F(\theta) = \sum_{y_r < \theta} P(\Theta = y_r). \quad (5)$$

З (1) та (5) витікає, що в точках розриву функція $F(\theta)$ приймає такі значення:

$$\begin{cases} F(y_1) = 0, \\ F(y_j) = \sum_{i=1}^{k_{j-1}} p_i, \quad j = \overline{2, N+1}, \end{cases} \quad (6)$$

де $y_{N+1} = y_N + c, c = \text{const} \neq 0$, тобто

$$F(\theta) = \begin{cases} 0, & \theta \leq y_1; \\ \sum_{i=1}^{k_{j-1}} p_i, & \theta \in (y_{j-1}, y_j], \quad j = \overline{2, N+1}. \end{cases} \quad (7)$$

Для вираження залежності між імовірністю звертання до ІБ і сумарним об'ємом пам'яті для їх розміщення введемо в розгляд функцію ψ , визначену в точках розриву дискретної випадкової величини Θ таким чином:

$$\psi(y_j) = \sum_{i=1}^{k_j} p_i, \quad j = \overline{1, N}. \quad (8)$$

Можна побачити, що

$$\psi(y_j) = F(y_{j+1}), \quad j = \overline{1, N},$$

тобто

$$\psi(\theta) = \begin{cases} 0, & \theta < y_0; \\ \sum_{i=1}^{k_j} p_i, & \theta \in (y_{j-1}, y_j], \quad j = \overline{1, N}. \end{cases} \quad (9)$$

Апроксимація функції за результатами аналітичної обробки даних дає такий результат:

$$F(\theta) = 1 - e^{-(\alpha + \beta\theta)}, \quad (10)$$

де α, β – коефіцієнти, що залежать від об'єму ІБ і ймовірностей звертання до них і що розраховуються таким чином:

$$\alpha = \frac{C \cdot y_1}{y_{N+1} - y_1}; \quad \beta = -\frac{C}{y_{N+1} - y_1}; \quad C = 20.$$

Уведемо x_{ij} – булеві змінні розподілу ІБ по типах апаратно-програмних засобів (АПЗ), котрі приймають ненульові значення тільки тоді, коли i -й ІБ розміщено у j -му вузлі АПС ІПС. :

Сумарні наведені витрати на необхідну кількість обчислювальних вузлів не повинні перевищувати максимально допустимого розміру наведених витрат f на модернізацію АПС ІПС при розподілі ресурсів, виділених для забезпечення функціонування ІБ, що визначається нерівністю:

$$\Phi(\bar{y}^0) = \sum_{j=1}^N f_j y_j^0 \leq F_{\max}. \quad (10)$$

Враховуючи (3), обмеження відносно сумарних затрат (10) в термінах змінних y_j записується як:

$$\begin{aligned} \Phi(\bar{y}) &= \sum_{j=1}^N f_j (y_j - y_{j-1}) = \\ &= f_1(y_1 - y_0) + f_2(y_2 - y_1) + \dots + f_N(y_N - y_{N-1}) = \end{aligned} \quad (11)$$

$$= \sum_{j=1}^N (f_j - f_{j+1}) y_j \leq F_{\max}, \quad f_{N+1} = 0$$

Середній час $T(\bar{y})$ доступу к M ІБ, які зберігаються у N типах вузлів ІПС визначається такою формулою:

$$\begin{aligned} T(\bar{y}) &= \sum_{j=1}^N \sum_{i=1}^M t_j p_i x_{ij} = \\ &= \sum_{j=1}^N t_j (p_1 x_{1j} + p_2 x_{2j} + \dots + p_M x_{Mj}) = \\ &= t_1(p_1 + \dots + p_{k_1}) + t_2(p_{k_1+1} + \dots + p_{k_2}) + \\ &+ \dots + t_N(p_{k_{N-1}+1} + \dots + p_{k_N}) = t_1(y_1) + \\ &+ t_2((y_2) - (y_1)) + \dots + t_N((y_N) - (y_{N-1})) = \\ &= (t_1 - t_2)(y_1) + (t_2 - t_3)(y_2) + \dots + \\ &+ t_N(y_N) = \sum_{j=1}^N (t_j - t_{j+1}) \psi(y_j) \end{aligned} \quad (12)$$

за умови, що $t_{N+1} = 0$.

Вищевикладені міркування приводять до постановки задачі математичного програмування:

потрібно знайти вектор $\bar{y} \in \mathbb{R}^N$, при якому досягається мінімум функції:

$$T(\bar{y}) = \sum_{j=1}^N (t_j - t_{j+1}) \psi(y_j) \quad (13)$$

при обмеженні на середній час доступу:

$$\sum_{j=1}^N (f_j - f_{j+1}) y_j \leq F_{\max}, \quad (14)$$

$$f_j, y_j \in \mathbb{R}^N \quad \forall j = \overline{1, N}.$$

Сформульована задача класифікується як задача нелінійного програмування в силу нелінійності функції $\psi(\bar{y})$ [4].

Розв'язання даної задачі дозволяє в самому загальному випадку на будь-яких рівнях ієрархії обробити та розмістити раціональним методом ІБ.

ВИСНОВКИ

Розглянуто принципи побудови математичних моделей, що дозволяють провести оптимальний розподіл інформаційного ресурсу системи між транзакціями до сховищ даних. Розглянута модель використовує базову модель інформаційно-телекомунікаційної мережі, що підтримує АСУ спеціального призначення. В моделі обмін здійснюється інформаційними блоками, що надходять із сховищ даних, та у загальному випадку мають різний об'єм. Крім того, у моделі враховуються ймовірності звертання до інформаційних блоків сховищ даних.

Напрямок подальших досліджень пов'язаний із подальшого розвитку методів та моделей планування розподілу інформаційного ресурсу із врахуванням особливостей АСУ спеціального призначення з метою зниження складності алгоритму, що реалізує математичну модель.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ РАСПРЕДЕЛЕНИЯ ИНФОРМАЦИОННЫХ РЕСУРСОВ МЕЖДУ ТРАНЗАКЦИИ К ХРАНИЛИЩУ ДАННЫХ

А.Н. Чаузов

Рассматриваются принципы построения математических моделей, позволяющих провести оптимальное распределение информационного ресурса системы между транзакциями в хранилища данных (ХД). Рассматриваемая модель использует базовую модель информационно-телекоммуникационной сети, поддерживает АСУ специального назначения. В модели обмен осуществляется информационными блоками (ИБ), поступающих из хранилищ данных, и в общем случае имеют разный объем. Кроме того, в модели учитываются вероятности обращения к ИБ ХД.

Ключевые слова: информационный ресурс, математическая модель, информационно-телекоммуникационная сеть, распределенные данные.

DISTRIBUTION MATHEMATICAL MODEL OF INFORMATION RESOURCES BETWEEN TRANSACTIONS TO THE DATA WAREHOUSE

O.N. Chauzov

In article are considered the principles of creation of mathematical model allowing of resource information system between transactions in the data warehouse (DW). The model uses a basic model of information-telecommunication network, supports automation of special purpose. The model exchanged information blocks (IB), coming from the data warehouse, and generally have a different amount. In addition, the model takes into account the probability of recourse to the IB DW.

Keywords: information resource, mathematical model, information and telecommunications network, distributed data.

Список літератури

1. Agarwal, S. Traffic engineering in software defined networks [Text] / S. Agarwal, M. Kodialam, T.V. Lakshman [Text] // INFOCOM, 2013 Proceedings IEEE. – IEEE, 2013. – P. 22.11-22.19.
2. Baki, A.K.M. Continuous monitoring of smart grid devices through multi protocol label switching [Text] / A.K.M. Baki // IEEE Transactions on Smart Grid. – 2014. – Vol. 5, № 3. – P. 12.10-12.15.
3. Логинов, И.В. Оптимизация модели распределенной гетерогенной вычислительной системы, используемой для планирования обработки запросов [Текст] / И.В. Логинов, Е.В. Лебедев // Информатика и системы управления. – 2009. – № 3(21). – С. 118-124.
4. Кучук, Г.А. Інформаційні технології управління інтегральними потоками даних в інформаційно-телекомунікаційних мережах систем критичного призначення [Текст] / Г.А. Кучук. – Х.: ХУ ПС, 2013. – 264 с.
5. Кучук, Г.А. Синтез стратифікованої інформаційної структури інтеграційної компоненти гетерогенної складової Єдиної АСУ Збройними Силами України [Текст] / Г.А. Кучук, О.П. Давікоза // Наука і техніка Повітряних Сил Збройних Сил України. – 2013. – № 3. – С. 154-158.
6. Schramm, C. Application-oriented network modeling with mobile agents [Text] / C. Schramm, A. Bieszczad, B. Pagurek // Network Operations and Management Symposium, NOMS 98., IEEE. – IEEE, 1998. – Vol. 2. – P. 696-700.
7. Gentsch, W. Sun Grid Engine: towards creating a compute power grid [Text] / W. Gentsch // Cluster Computing and the Grid, 2011. Proceedings. First IEEE/ACM International Symposium on. USA, 2011. – P. 35-36.
8. Subramanyan, R. A scalable SNMP-based distributed monitoring system for heterogeneous network computing [Text] / R. Subramanyan, Jose Miguel-Alonso, Jose A.B. Fortes // Conference on High Performance Networking and Computing. USA. No. 14. 2000. – P. 116-119.
9. Азаренко Е.В. Проектирование автоматизированных систем управления на компьютерных сетях : монография [Текст] / Азаренко Е.В., Герасимов Б.М., Шохин Б.П. – Севастополь: Гос. океанариум, 2007. – 272 с.

Надійшла до редколегії 29.08.2015

Рецензент: д-р техн. наук, проф. С.В. Козелков, Державний університет телекомунікацій, Київ.