

УДК 681.324 : 621.325

І.В. Ільїна, Г.А. Кучук, Ю.В. Мишуніна

Харківський університет Повітряних Сил ім. Івана Кожедуба, Харків

МЕТОД МІНІМІЗАЦІЇ СУМАРНОГО ОБСЯГУ ІНТЕГРАЛЬНИХ ПОТОКІВ РОЗПОДІЛЕНОЇ ІНФОРМАЦІЙНО-ДОВІДКОВОЇ СИСТЕМИ

Проаналізовано засоби організації графічного представлення статистичних даних у інформаційно-довідкових системах (ІДС). Розроблено програмний засіб для представлення статистичних даних у графічному вигляді, який відрізняється від відомих більш простим та зручним інтегруванням із середовищем розробки баз даних ІДС. Запропоновано метод мінімізації навантаження віртуальних каналів мультисервісної мережі, яка підтримує інформаційно-довідкову систему з розширеним інтерфейсом, за рахунок створення резервних копій графічних образів на кінцевих пунктах прийому інформації.

Ключові слова: *інформаційно-довідкова система, статистичні дані, засоби графічний образ, віртуальний канал, динамічна база даних.*

Вступ

Інформаційні системи (ІС) здавна знаходять (в тому чи іншому вигляді) досить широке застосування в життєдіяльності людства. Це пов'язано з тим, що для існування цивілізації необхідним є обмін інформацією – передача знань, як між окремими колективами суспільства, так і між різними поколіннями.

Інформаційна система – це система, в якій за допомогою технічних засобів забезпечується зби-

рання, накопичення, обробка інформації і видача результатів людині або групі людей, що беруть участь в управлінні системою [1]. Проте при всьому різноманітті на ринку програмного забезпечення для організації інформаційної системи створення кожної системи носить вузький прикладний характер і у зв'язку з цим, як правило, виникає потреба в додаткових розробках, що враховують особливості специфіки її використання.

Окремий клас інформаційних систем складають інформаційно-довідкові системи (ІДС), спрямовані на надання кінцевому користувачу великих обсягів статистичної інформації (ІДС СІ). Зазвичай програмне середовище, у якому розроблюється така ІДС, спрямоване на швидкий пошук даних та зручний інтерфейс їх висвітлення у оригінальному вигляді. Але інтерфейс графічного надання статистичної інформації такого середовища зазвичай обмежений. Так, наприклад, у середовищі Microsoft Office, ІДС розробляється за допомогою системи управління базами даних Microsoft Access [2], а найбільш сприятливий інтерфейс графічного зображення статистичних даних має комплекс електронних таблиць Microsoft Excel [3]. Середовище Microsoft Office дозволяє інтеграцію доданків, але для кінцевого користувача ІДС потрібен більш зручніший та простіший механізм інтегрування.

Аналогічні проблеми виникають при функціонуванні ІДС СІ у інших програмних середовищах.

Таким чином виникає потреба розширення інтерфейсу користувача ІДС СІ графічними можливостями відображення статистичних даних, які надають спеціалізовані програмні комплекси середовища функціонування

Для вирішення даного завдання на прикладі ІДС СІ, які функціонують у середовищі Microsoft Office та вимагають використання графічних засобів Microsoft Excel, був розроблений програмний засіб інтегрування вищезазначених додатків Microsoft Office (далі – ПЗІ), який відрізняється від відомих більш простим та зручним інтегруванням із середовищем розробки баз даних Microsoft Access та надає широкі можливості представлення даних за допомогою діаграм, що мають дуже широкий спектр налаштувань. Розроблений програмний засіб пройшов дослідну експлуатацію на ряді ІДС СІ.

Аналіз останніх досліджень і публікацій та постановка завдання

В процесі дослідної експлуатації розробленого програмного засобу ПЗІ було виявлено, що його використання у локальному режимі відповідає часовим вимогам відповідних ІДС. Але втілення ПЗІ до розподілених ІДС, які функціонують у середовищі мультисервісної мережі (МСМ), у разі численних одночасних запитів призводило до неприпустимих затримок при отриманні відповідей на запити користувачів. Аналіз трафіка ІДС показав, що причиною затримок у таких випадках було перевищення навантаження каналів МСМ за рахунок великих обсягів передаваної графічної інформації.

Тому при експлуатації ІДС в середовищі мультисервісної мережі (МСМ) в режимі одночасного доступу до результатів графічного відображення статистичної інформації актуальним є питання про

зниження завантаження віртуальних каналів (ВіК), що з'єднують вузол збору статистичних даних з вузлами користувачів графічної інформації (ВК ГІ). Відносно низька пропускна спроможність даних каналів призводить як до зниження оперативної обробки транзакцій, так і до створення черг на обслуговування запитів. Не останню роль відіграє і прагнення до зменшення орендної плати за використання виділених каналів зв'язку.

Одним з відомих підходів до вирішення цього питання є введення інформаційної динамічної бази даних графічної інформації (ДБД ГІ) за рахунок створення додаткових копій інформаційних одиниць ДБД на деяких ВК ГІ.

Так, в [4] цей підхід досліджений стосовно фіксованого числа копій. У [5] число копій не фіксується, а як критерій оптимальності приймається середній об'єм даних, що пересилаються, при обробці запитувального і такого, що коригує повідомлень. У [6] мінімізується оплата оренди.

Але в усіх вищеперелічених роботах не враховується різноманітність каналів зв'язку, а питання підвищення оперативності обробки транзакції не досліджується спільно з вимогою зниження витрат на передачу інформації. Тому пропонується при дослідженні необхідності створення копій ГО на різних ВК ГІ обчислювальної мережі мінімізувати сумарні часові витрати усіх ВіК мережі при обробці запитів транзакцій на отримання даних і при коригуванні додаткових ГО.

Підсумовуючи вищезазначене, визначимо, що **метою даної статті** є розробка методу мінімізації навантаження віртуальних каналів мультисервісної мережі, яка підтримує інформаційно-довідкову систему, що функціонує у середовищі підтримки системи управління базами даних, з використанням ПЗІ.

Результати досліджень

Для формалізації постановки завдання опишемо початкові дані, причому розподіл базових ГО по ВК ГІ припускаємо зафіксованим.

Визначимо кількісний склад досліджуваної ІДС. Для цього введемо такі позначення:

N – кількість ВКГІ, що входять до складу досліджуваної системи;

M – кількість базових реляційних таблиць ДБД ГІ;

K – кількість транзакцій ДБД ГІ.

Кожному ВК ГІ, ГО і транзакції, відповідно, надамо порядковий номер:

$$n \in \overline{1, N}, m \in \overline{1, M}, k \in \overline{1, K}.$$

Структуру зовнішньої пам'яті ІДС опишемо булевою матрицею розподілу ГО ДБД ГІ по ВК ГІ $D = (d_{nm})$, елементи якої відрізняються від нуля тоді і тільки тоді, коли n -та ВК ГІ містить m -й базовий ГО. Визначимо пропускну спроможність (тобто об'єм інформації, яка приймається за одини-

цю часу) ВіК $H_{n_1 n_2}$ між n_1 -м та n_2 -м ВК ГІ як $\beta_{n_1 n_2}(n_1, n_2 \in \overline{1, N})$.

Без обмеження спільності закріпимо кожну транзакцію за конкретним ВК ГІ і зафіксуємо це закріплення булевою матрицею

$$Q_{nk} = (q_{nk}), \quad n \in \overline{1, N}, \quad k \in \overline{1, K},$$

де $q_{nk} = 1$ тоді і тільки тоді, якщо за n -м ВК ГІ закріплена k -та транзакція.

Для визначення необхідного числа додаткових копій кожного з N графічних образів введемо булеву матрицю $X = (x_{nm})$ розмірністю $N \times M$, причому $x_{nm} = 1$ тоді і тільки тоді, коли n -та ВК ГІ містить копію m -го базового ГО.

Напевне, що

$$d_{nm} + x_{nm} \leq 1, \quad \forall n \in \overline{1, N}, \quad \forall m \in \overline{1, M},$$

а кількість копій конкретної таблиці m визначається як $\sum_{n=1}^N x_{nm}$.

Виходячи із статистики експлуатації ІДС, можна визначити такі її функціональні характеристики:

a_{mk} – об'єм запиту даних, що передаються для одноразової корекції m -го ГО;

b_{mk} – усереднений об'єм даних, які запрошуюються транзакцією k для m -го ГО;

c_m – усереднений об'єм даних, що передаються для одноразової корекції m -го ГО;

$\lambda_m^{(k)}$ – інтенсивність одноразових корекцій m -го ГО;

λ – інтенсивність запитів k -ої транзакції до m -го ГО;

V_m – об'єм даних ГО з номером m ;

W_m – об'єм зовнішньої пам'яті, доступної для ГО ДБД ГІ на ВК ГІ з номером n .

Усі вищеперелічені характеристики розглядаються за фіксований інтервал часу, кратний часу періодів основних транзакцій ІДС. Розмірністю об'єму даних ІДС є базова інформаційна одиниця (ІО) мережі передачі даних (МПД).

Розрахуємо ряд величин, що характеризують процес передачі інформації в середовищі даної ДБД ГІ.

Корекція копій m -го ГО з інтенсивністю $\lambda_m^{(k)}$ вимагає передачі

$$u_m^{(k)} = \sum \lambda_m^{(k)} c_m x_{n_1 m} \quad (\text{ІО МПД})$$

за час

$$t_m^{(k)} = \sum_{n_1=1}^N \frac{\lambda_m^{(k)} c_m x_{n_1 m}}{\sum_{n_2=1}^N d_{n_2 m} \cdot \beta_{n_1 n_2}}.$$

Отже, за даний часовий інтервал віддалені канали зв'язку будуть сумарно завантажені корекцією копій ГО протягом такого часу

$$t_m^{(k)} = \sum_{m=1}^M t_m^{(k)} = \sum_{m=1}^M \lambda_m^{(k)} \cdot c_m \cdot \sum_{n_1=1}^N \frac{\lambda_m^{(k)} c_m x_{n_1 m}}{\sum_{n_2=1}^N d_{n_2 m} \cdot \beta_{n_1 n_2}}, \quad (1)$$

а середній час корекції одного ГО

$$t_{cp}^{(k)} = t_{\Sigma}^{(k)} / \sum_{i=1}^M \lambda_m^{(k)}. \quad (2)$$

Оскільки у ряді випадків запити транзакцій ДБД ГІ отримуватимуть дані з копій ГО на "свій" ВК ГІ, то завантаження віддалених каналів зв'язку зменшиться. Розрахуємо $t_{\Sigma}^{(k)}$ – сумарне завантаження каналів запитами транзакцій за даний часовий інтервал. Для цього зафіксуємо такий процес: відбувається запит до m -го ГО із транзакції з номером k . Оскільки транзакція закріплена за конкретним ВК ГІ ($n_k | q_{n_k k} = 1$), то

$$\sum q_{n_3 k} \cdot x_{n_3 m} = 1 \Leftrightarrow x_{n_k m} = 1. \quad (3)$$

У інших випадках сума у виразі (3) набуває тільки нульових значень.

При виконанні умови (3) віддалені виділені канали ІДС в зафіксованому процесі не задіюються, в будь-якому іншому випадку необхідно передати $a_{mk} + b_{mk}$ інформаційних одиниць МПД.

Отже, з урахуванням (3), потрібно передача об'єму даних

$$u_m = \lambda_{mk} (a_{mk} + b_{mk}) \cdot \left(1 - \sum_{n_3=1}^n q_{n_3 k} \cdot x_{n_3 k} \right)$$

по каналу $H_{n_4 n_5}$ з пропускною спроможністю

$$\tilde{q}_{n_4 n_5} = \sum_{n_4=1}^N \sum_{n_5=1}^N q_{n_4 k} \cdot d_{n_5 m} \cdot \beta_{n_4 n_5}$$

за час

$$t_m = \frac{u_m}{\tilde{q}_{n_4 n_5}} = \frac{\lambda_{mk} (a_{mk} + b_{mk}) \cdot \left(1 - \sum_{n_3=1}^n q_{n_3 k} \cdot x_{n_3 k} \right)}{\sum_{n_4=1}^N \sum_{n_5=1}^N q_{n_4 k} \cdot d_{n_5 m} \cdot \beta_{n_4 n_5}}. \quad (4)$$

Виходячи з (4) розраховуємо t_{Σ} як

$$t_{\Sigma} = \sum_{k=1}^K \sum_{m=1}^M t_m = \sum_{k=1}^K \sum_{m=1}^M \frac{\lambda_{mk} (a_{mk} + b_{mk}) \cdot \left(1 - \sum_{n_3=1}^n q_{n_3 k} \cdot x_{n_3 k} \right)}{\sum_{n_4=1}^N \sum_{n_5=1}^N q_{n_4 k} \cdot d_{n_5 m} \cdot \beta_{n_4 n_5}}. \quad (5)$$

Тоді середній час, потрібний транзакції для отримання даних з одного ГО, розраховується як

$$t_{cp} = \frac{t_{\Sigma}}{\sum_{k=1}^K \sum_{m=1}^M \lambda_{mk}} \quad (6)$$

З урахуванням вищевикладеного критерію мінімізації і отриманих виразів (1) і (5) цільова функція завдання набуде такого вигляду:

$$t_{cp}^{(k)} + t_{\Sigma} \rightarrow \min ,$$

$$\sum_{m=1}^M \lambda_m^{(k)} \cdot c_m \cdot \left(\frac{\sum_{n_1=1}^N \lambda_m^{(k)}}{\sum_{n_2=1}^N d_{n_2m} \cdot \beta_{n_1n_2}} \right) +$$

$$+ \sum_{k=1}^K \sum_{m=1}^M \frac{\lambda_{mk} (a_{mk} + b_{mk}) \cdot \left(1 - \sum_{n_3=1}^n q_{n_3k} \cdot x_{n_3k} \right)}{\sum_{n_4=1}^N \sum_{n_5=1}^N q_{n_4k} \cdot d_{n_5m} \cdot \beta_{n_4n_5}} \rightarrow (7)$$

$$\rightarrow \min$$

Найбільш загальні обмеження завдання пов'язані з розміром дозволеної зовнішньої пам'яті для ДБД ГІ, тобто

$$\sum_{m=1}^M V_m x_{nm} \leq W_n, \forall n \in \overline{1, N}. \quad (8)$$

Із запропонованого способу формалізації витікають такі обмеження:

– закріплення кожної транзакції за визначеним ВК ГІ, тобто

$$\sum_{n=1}^N q_{nk} = 1, \forall k \in \overline{1, K}; \quad (9)$$

– фіксований розподіл базових ГО по ВК ГІ, тобто

$$\sum_{n=1}^N d_{nm} = 1, \forall m \in \overline{1, M}; \quad (10)$$

– неприпустимість створення більше однієї копії одного і того ж ГО на одному ВК ГІ:

$$\sum_{n=1}^N x_{nm} \leq M, \forall m \in \overline{1, M}; \quad (11)$$

$$d_{nm} + x_{nm} \leq 1, n \in \overline{1, N}, \forall m \in \overline{1, M}. \quad (12)$$

Сформульована задача (7) – (12) відноситься до класу задач булева програмування. При невеликих значеннях N, M, K вона вирішується стандартними точними методами (наприклад, адитивний алгоритм Балаша [4]). При збільшенні розмірності необхідно шукати спеціальні методи рішень (наближені або точні особливості обмежень завдання, що враховуються).

У разі, коли періоди повторюваності транзакцій визначити достатньо важко або неможливо, фіксований інтервал часу визначається довільно, а цільова функція (7) – виразом, що отримується з (2) і (6):

$$t_{cp}^{(k)} + t_{cp} \rightarrow \min . \quad (13)$$

Тим самим, застосовуючи (13), у критерії мінімізації ми переходимо до обчислення середніх часів корекції одного ГО і отриманню даних з одного ВК ГІ.

Висновки

Розроблено програмний засіб для представлення статистичних даних у графічному вигляді, який відрізняється від відомих більш простим та зручним інтегруванням із середовищем розробки баз даних ІДС. Запропоновано метод мінімізації навантаження віртуальних каналів мультисервісної мережі, яка підтримує інформаційно-довідкову систему з розширеним інтерфейсом, за рахунок створення резервних копій графічних образів на кінцевих пунктах прийому інформації. Напрямок подальших досліджень – алгоритмізація запропонованого методу та розробка відповідного програмного забезпечення, яке дозволить автоматично створювати резервні копії графічних образів.

Список літератури

1. Коннолли Т. Базы данных: проектирование, реализация и сопровождение. Теория и практика: пер. с англ.; 2-е изд. / Томас Коннолли, Каролин Бегг. – М.: Издательский дом "Вильямс", 2001. – 1120 с.
2. Боровіков В.У. MS ACCESS 2002. Програмування і розробка баз даних і додатків / В.У. Боровіков. – СО-ЛОН-Р, 2002. – 560 с.
3. Глушаков С.В. Персональный компьютер; 5-е изд., доп. и перераб. / С.В. Глушаков, А.С. Сурядный. – Х: Фолио, 2003. – 500 с.
4. Tabatabai V. Queue theoretic model for file assignment in a distributed date base Network / V. Tabatabai, M. Arozuliah // "COMPCON". Fall 78 IEEE Comp. Soc. Inf. conf. - Washington, 1978. – P. 175-182.
5. Ярмош Н.А. Алгоритм оптимального размещения файлов в распределенной ВС / Н.А. Ярмош, Е.М. Журавский // Вести АН БССР. Сер. физ.-техн. наук. – 1984. – № 1. – С. 78-83.
6. Motzkin D. An optimal data allocation model for distributed / D. Motzkin // Math, and Comput. Modell. – 1988. – Vol. 11. – P. 920-925.
7. Бахтин А.Б. Дискретные задачи производственного типа / А.Б. Бахтин, А.А. Колоколов, З.В. Коробкова. – Н-ск.: Наука, 1988. – 168 с.

Надійшла до редколегії 26.10.2010

Рецензент: д-р техн. наук, проф. І.В. Рубан, Харківський університет Повітряних Сил ім. І. Кожедуба, Харків.

МЕТОД МИНИМИЗАЦИИ СУММАРНОГО ОБЪЕМА ИНТЕГРАЛЬНЫХ ПОТОКОВ РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННО-СПРАВОЧНОЙ СИСТЕМЫ

И.В. Ильина, Г.А. Кучук, Ю.В. Мишунина

Проанализированы средства организации графического представления статистических данных в информационно-справочных системах (ИСС). Разработана программа для представления статистических данных в графическом виде, которая отличается от известных более простым и удобным интегрированием со средой разработки баз данных

ИСС. Предложен метод минимизации нагрузки виртуальных каналов мультисервисной сети, которая поддерживает информационно-справочную систему с расширенным интерфейсом, за счет создания резервных копий графических образов на конечных пунктах приема информации.

Ключевые слова: информационно-справочная система, статистические данные, средства, графический образ, виртуальный канал, динамическая база данных.

**METHOD OF INTEGRAL STREAMS TOTAL VOLUME MINIMIZATION
OF THE DISTRIBUTED I&R SYSTEM**

I.V. Il'ina, G.A. Kuchuk, Yu.V. Mishunina

Facilities of organization of graphic presentation of statistical information are analysed in the I&R systems (IRS). The program is developed for presentation of statistical information in a graphic kind, which differs from known more simple and comfortable integration with the environment of development of databases IRS. The method of minimization of loading of the virtual ductings of multiservice network which supports the информационно-справочную system with the extended interface is offered, due to creation of back-up copies of graphic appearances on the eventual points of reception of information.

Keywords: I&R system, statistical information, facilities, graphic appearance, virtual circuit, dynamic database.