

МАТЕМАТИЧНА МОДЕЛЬ ОБРОБКИ ІНФОРМАЦІЇ СКЛАДНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ

Розглядається питання моделювання обробки інформації складної інформаційної системи для підвищення ефективності інформаційного забезпечення правоохоронних органів. Запропоновано застосування кластеризації обчислювальних вузлів для моделювання планування обробки інформації в розподіленій гетерогенній обчислювальній системі.

The article deals with the matter of modeling of information processing of the sophisticated information system for the improvement of effectiveness of information support of the law enforcement authorities. It was suggested to use the clusterization of computational nodes to simulate the planning of data processing within the distributed heterogeneous computing system.

1. ПОСТАНОВКА ЗАДАЧІ

Стрімке впровадження інформаційно-телекомунікаційних технологій управління в усі сфери життєдіяльності суспільства, широке застосування цифрової технології обробки та обміну даними в практику діяльності організацій та установ, обумовлює необхідність пошуку нових підходів до розробки алгоритмів функціонування системи інформаційного забезпечення підприємств та закладів.

Найбільш актуального значення зазначена обставина набуває в діяльності органів внутрішніх справ, які функціонують в умовах обробки значних обсягів даних для ефективного управління підрозділами міліції щодо вирішення завдань охорони громадського порядку та протидії злочинності. Тому науковий пошук у напрямку удосконалення способів та методів обробки даних у складі складної інформаційної системи спеціального призначення є достатньо обґрунтованим.

Дослідженням організаційно-технічних проблем інформаційного забезпечення правоохоронних органів свого часу займалися такі вчені, як Калюжний Р.А., Кондратьєв Я. Ю., Петрачук Л. І., Саницький В. А., Шавиркін Б.Б., Хахановський В. Г. [1-5] та інші. Поряд з цим зазначимо, що питанням моделювання обробки інформації складних розподілених ієрархічних систем спеціального призначення у наукових розробках приділено не достатньо уваги.

¹ Кримський юридичний інститут
Одеського державного університету внутрішніх справ

Мета роботи полягає у проведенні аналізу наукових підходів до моделювання систем обробки інформації складної інформаційної системи для оптимізації комплексу заходів інформаційного забезпечення в діяльності органів внутрішніх справ (ОВС) України.

2. ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Інформаційні підсистеми, як основні складові частини системи інформаційного забезпечення, призначені для збору, накопичення, зберігання та обробки інформації певних обліків, які орієнтовані на використання в діяльності багатьох служб. Найбільша ефективність автоматизованих інформаційних систем досягається при оптимізації планів роботи, швидкому виборі оперативних рішень, чіткому маневруванні матеріальними і фінансовими ресурсами тощо.

Найбільш прийнятною для забезпечення виконання завдань ОВС є ієрархічна структура інформаційної системи, яка має зіркоподібну топологію (рис. 1) [3, 4, 6].

Описуючи більш конкретно інформаційну систему ОВС (рис. 2) [6], слід зазначити, що основою системи збирання, контролю та використання інформації складає нижчий (третій) рівень. На цьому рівні забезпечується первинне накопичення інформації, ведення територіальних банків даних, захист інформації, актуалізація інформаційних обліків та передавання інформації до банків даних другого і першого рівнів.

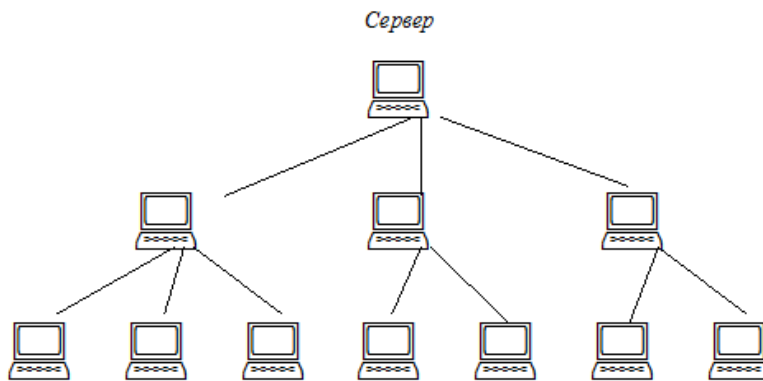


Рис. 1. Ієрархічна структура інформаційної системи ОВС

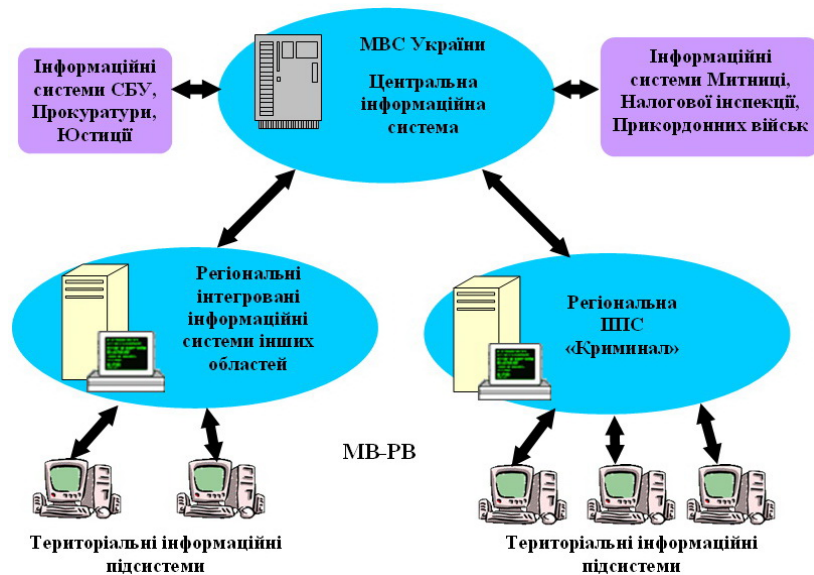


Рис. 2. Система інформаційного забезпечення ОВС України

На другому рівні здійснюється формування регіональних банків даних, опрацювання інформації, забезпечується доступ територіальних підрозділів для авторизованого використання інформації, інформаційний зв'язок із територіальними правоохоронними органами та іншими установами держави. На першому рівні забезпечується інтеграція та опрацювання інформації для формування банків даних інформаційних підсистем, міжвідомчий та міжрегіональний інформаційних зв'язок, управління системою інформаційного забезпечення ОВС та дотримання стратегії її розвитку, розробка нормативно-правової бази та інформаційних технологій.

Розглядаючи процедуру обробки інформації і складних ієрархічних системах ОВС (рис. 3), слід зазначити наявність низки проблем, пов'язаних зі значною кількістю хибних інформаційних пакетів, поява яких визначена причинами недосконалості програмного забезпечення, низькою кваліфікацією персоналу та ін.



Рис. 3. Технологія обробки інформації в інформаційних підсистемах ОВС.
(БД – база даних; ЛБД – локальна база даних)

У такому випадку для вирішення зазначеної проблеми доцільним є застосування моделі обчислювальної системи та вхідного потоку інформаційних даних (запитів) в складній інформаційній системі, а також алгоритму оптимізації. Тоді ефективність планування обробки інформації буде визначатися адекватністю використаних моделей в обчислювальній системі [7].

Існуючі моделі обчислювальних систем, які використовуються для планування обробки інформації в складних інформаційних системах, можна розглядати з трьох позицій:

- 1) Адекватність моделі реальному об'єкту інформатизації;
- 2) Ресурсна ємність алгоритму оптимізації при використанні моделі;
- 3) Ресурсна ємність отримання інформації стану.

Особливістю процесу моделювання є те, що при визначенні класу моделі обчислювальної системи для використання в плануванні обробки інформації на частину показників накладаються обмеження, а інші оптимізуються. Одним випадком зазначеного є більшість систем планування обробки інформації [8], де обчислювальний вузол моделюється процесом, який моделює інформаційний потік в монопольному режимі, що суттєво знижує адекватність моделі.

Тому в умовах застосування складних розподілених ієрархічних інформаційних систем доцільно застосовувати розподілену гетерогенну обчислювальну систему [9], яка обробляє вхідний нестаціонарний інформаційний потік. Метою функціонування системи планування

обробки інформації є максималізація задоволення потреб користувачів по своєчасності отримання результатів.

У такому випадку цільова функція має вигляд

$$\sum_{i=0}^n \xi^i (t_{os}^i) \rightarrow \max,$$

де n – кількість отриманих запитів за інтервал часу планування.

Вирішення завдання оптимізації моделі інформаційної системи для використання з метою планування обробки інформації передбачає наявність наступних вимог мінімізації:

1) Розбіжність вихідних реакцій реальної системи та її моделі при обробці інформації

$$Z = \sum_i |(R(Q_i) - R'(Q_i))| \rightarrow \min, \quad (1)$$

де $R(Q_i)$ – витрати ресурсів усіх видів на обробку i -го запиту [10];

2) Витрати ресурсів на забезпечення актуальності моделі та планування

$$Z = r^p + r^a \rightarrow \min, \quad (2)$$

де r^p – ресурсна ємкість планування;

r^a – ресурсна ємкість підтримки актуальності моделі.

Вирішення завдання (1), (2) можливе з застосуванням наступної моделі обчислювальної системи:

$$S_{ГРОС} = \{S_{OC}, S_K, P\},$$

де S_{IN} – модель обчислювальної системи; S_E – модель комунікаційної системи; D – протоколи взаємодії елементів системи.

Модель обчислювальної системи S_{IN} задається набором груп обчислювальних вузлів G : $S_{IC} = \{G_1, \dots, G_m\}$. Кожна група обчислювальних вузлів призначена для обробки інформації одного типу.

Модель обчислювального вузла $V_j \in G_k$ дозволяє описувати як однопроцесорні ЕОМ, так і багатопроцесорні системи, які знаходяться під управлінням операційної системи кластерного типу (наприклад, Windows Computer Cluster [11]). Для цього пропонується модель обчислювального вузла наступного вигляду:

$$V = \langle n, p, o, C, os, Arh, PO \rangle, \quad (3)$$

де n – кількість процесорів (ядер) обчислювального вузла;

p – умовна виробнича потужність по обробці найбільш приватних запитів;

o – обсяг оперативної пам'яті;

$C = \langle k, c, t_p \rangle$ – комунікаційна підсистема (кількість, виробнича потужність і тип мережених інтерфейсів);

os – тип операційної системи;

Arh – архітектура обчислювальної системи;

PO – множина встановленого програмного забезпечення.

Облік виробничої потужності обчислювального вузла на основі запитів, які зустрічаються найбільш часто, дозволяє прогнозувати час завершення обробки запиту, що мінімізує вираз (1).

Для обліку поточного стану обчислювальних вузлів пропонується обліковувати очікуваний час обробки запиту (запитів) $t = t_1, \dots, t_s$, завантаженість $\eta = \eta_m, \eta_o, \eta_C$, а також список ідентифікаторів запитів $Q = Q_{UID_1}, \dots, Q_{UID_s}$. Тоді модель обчислювального вузла набуває наступного вигляду

$$V_j = \langle t, \eta, Q \rangle$$

Для мінімізації ресурсних витрат (2) пропонується визначати вираз (3) тільки одним раз при підключенні обчислювального вузла до системи. При зміні конфігурації вузла в систему планування повинна передаватися тільки зміна параметрів ΔV . Для зменшення ресурсної ємності планування пропонується об'єднувати обчислювальні вузли, які обробляють запити одного типу з приблизно однаковою виробничою потужністю в групи обчислювальних вузлів G.

Об'єднувати обчислювальні вузли в групи пропонується на основі близькості вихідних впливів при обробці запитів з використанням методу k-середніх [12]. Для цього використовується набір моделей обчислювальних вузлів (3) і його вихідна реакція на обробку запиту $R(Q_i)$, яка визначається обсягом витрат ресурсів. Кожна група обчислювальних вузлів визначається моделлю:

$$G = \langle \theta, \bar{p}, \bar{o}, \bar{C}, os, Arh, PO \rangle,$$

де θ - функція прогнозування часу завершення обробки запиту;

$\bar{p}, \bar{o}, \bar{C}$ – середні значення відповідних параметрів.

В результаті кластеризації зменшується кількість параметрів моделі обчислювального вузла:

$$V = \langle n, p, o, C, os, Arh, PO \rangle \Leftrightarrow V = \langle n, k_p, k_o, k_C, G \rangle,$$

$$\text{де } k_p = \frac{p}{\bar{p}}, \quad k_o = \frac{o}{\bar{o}}, \quad k_C = \frac{C}{\bar{C}}.$$

Таким чином, модель обчислювального вузла представляє собою об'єднання моделей V та V^o :

$$Vp = \{V, V^o\} = \{ \langle n, k_p, k_o, k_c, G \rangle, \langle t, \eta, Q \rangle \}$$

Щодо моделі комунікаційної системи, то вона представляється у вигляді графу, вершинами якого є обчислювальні вузли, а ребрами – комунікаційні лінії:

$$S_K = (\{V\}, \{C\}).$$

При цьому кожне ребро має вагову характеристику, яка відображає пропускну спроможність та латентність лінії. Протоколи взаємозв'язку Р елементів системи визначаються у вигляді набору записів, які визначають правила та механізми взаємозв'язку.

Для перевірки ефективності моделі доцільно застосовувати експериментальний метод порівняння підходів до моделювання за показниками адекватності моделі обчислювальної системи, ресурсної ємкості вирішення завдання планування, а також ресурсної ємкості підтримки моделі в актуальному стані.

Зокрема, оцінка адекватності моделі проводиться за показником похибки прогнозування часу обробки запиту $\varepsilon = \frac{1}{n} \sum_{i=0}^n |t - t_i| / t_i$.

3. ВИСНОВКИ

Підсумовуючи, слід зазначити, що застосування запропонованої моделі дозволяє підвищити адекватність моделі та зменшити її ресурсну ємкість. Зазначене є запорукою підвищення ефективності обробки інформації складної інформаційної системи спеціального призначення.

1. Калюжний Р. А. Автоматизовані системи і засоби обчислювальної техніки у організації і функціонуванні органів внутрішніх справ / Калюжний Р. А. – К.: КНУВС, 1990. – 193 с. 2. Кондратьєв Я. Ю. Нормативно-правова база інформаційно-аналітичного забезпечення діяльності оперативних підрозділів міліції / Я. Ю. Кондратьєв, В. Г. Хахановський // Науковий вісник НАВС України, 2003. – № 3, ч. 2. – С. 53-59. 3. Саницький В. А. Комп'ютеризовані інформаційні підсистеми МВС України у боротьбі зі злочинністю: монографія / В. А. Саницький. – К.: АНТЕКС, 1998. – 130 с. 4. Саницький В. А. Система інформаційного забезпечення ОВС України: навчально-практичний посібник [для курс. вищ. навч. закл. МВС] / В. А. Саницький. – К.: АНТЕКС, 2000. – 144 с. 5. Хахановський В. Г. Правові та організаційно – технічні проблеми захисту конфіденційної інформації в автоматизованих системах ОВС України / В. Г. Хахановський // Науковий вісник НАВС України, 2002. – № 1, ч. 1. – С. 39-45. 6. Шавиркін Б. Б. Застосування ЕОМ в ОВС: навчальний посібник [для курс. вищ. навч. закл. МВС] / Б. Б. Шавиркін, Л. І. Пет-

- рчук. – Донецьк: ДЮІ, 2007. – 178 с. 7. Топорков В. В. Модели распределенных вычислений: монография / В.В. Топорков. – М.: Физматлит, 2004. – 185 с.
8. Якововский М. В. Распределенные системы и сети: учеб. пособие [для студ. высш. учебн. завед.] / М. В. Якововский. – М.: МГТУ «Станкин», 2000. – 275 с.
9. Лебедеко Е. В. Оптимизация модели распределенной вычислительной системы, используемой для планирования обработки запросов / Е. В. Лебедеко, И. В. Логвинов // Информатика и системы управления. - Благовещенск: АГУ, 2009. – № 3(21). – С. 118-124. 10. Alur R. Model-checking for real-time systems / Alur R., Coucorbetis C., Dill D. // Proceedings of LICS 90. – 1990. – P. 414-425.
11. Windows Computer Cluster Server 2003 [Электронный ресурс]. – Режим доступа : <http://www.microsoft.com/rus/win-dowsserver2003/ccs/default.aspx>.
12. Мандель И. Д. Кластерный анализ: монография / И. Д. Мандель. – М.: Финансы и статистика, 1988. – 174 с.