

АНАЛІЗ МЕТОДІВ ОБРОБКИ КОРПОРАТИВНИХ ДАНИХ В СИСТЕМАХ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ

У статті на основі проведених досліджень науково-методологічного характеру сформульовані основні ознаки та концепції збереження і представлення даних в системах інтелектуальної підтримки прийняття рішень масштабу корпоративного підприємства

Ключові слова: сховища даних, інтелектуальні системи, системи підтримки прийняття рішень, корпоративні системи

Вступ. Еволюція засобів обчислювальної техніки, застосовуваних в управлінні економікою, спрямована на створення систем, які реалізують функції штучного інтелекту: експертні системи реального часу, системи підтримки прийняття управлінських рішень (СППР), інтелектуальні аналітико-інструментальні пакети, тощо.

Розмаїття означень систем підтримки прийняття рішень відбиває широкий діапазон різних форм, розмірів, типів СППР. СППР набули широкого застосування в економіці передових країн світу, причому їх кількість постійно зростає. На рівні стратегічного управління використовується ряд СППР, зокрема, для довго- середньо- і короткострокового, а також фінансового планування. Орієнтовані на операційне управління СППР застосовуються у галузях маркетингу (прогнозування та аналіз збуту, дослідження ринку цін), науково-дослідних та конструкторських роботах.

Постановка завдання. СППР присутні в будь-якій управлінській інформаційній системі. Тому, до задачі створення системи підтримки прийняття управлінських рішень організації приступають відразу після придбання обчислювальної техніки і установки програмного забезпечення. При розвитку бізнесу, впорядкування структури організації і налагодження міжкорпоративних зв'язків, проблема розробки і впровадження СППР стає особливо актуальною.

Виклад основного матеріалу дослідження. Одним з раціональних підходів до постановки задачі розробки системи інтелектуальної підтримки прийняття рішень масштабу корпоративного підприємства є підхід, орієнтований на інтеграцію двох зв'язаних інформаційних технологій: технології побудови інформаційних сховищ і технології інтелектуального аналізу даних.

СППР проводить аналіз причинних зв'язків помилок, визначає їх залежність від інших факторів та розробляє рекомендації з їх усунення. Такий підхід дозволяє запобігти складних та тривалих обчислень, характерних для традиційних методів аналізу, а замість них скористуватися накопиченими знаннями про можливі причини помилок.

В області інформаційних технологій завжди співіснували два класи систем [1]:

1) системи, орієнтовані на операційну (транзакційну) обробку даних. В англійській літературі вони часто називаються терміном OLTP (On-Line Transaction Processing) - оперативна транзакційна обробка, у протилежності OLAP - оперативній аналітичній обробці або терміном "системи обробки даних" (СОД);

2) системи, орієнтовані на аналітичну обробку даних - системи підтримки прийняття рішень (СППР), або Decision Support Systems (DSS).

На перших стадіях інформатизації завжди потрібно навести порядок саме в процесах повсякденної рутинної обробки даних, на що й орієнтовані традиційні СОД, тому випереджувальний розвиток систем цього класу цілком закономірний. СППР є вторинними стосовно них. Часто виникає ситуація, коли дані в організації накопичуються в ряді незв'язаних СОД, багато в чому дублюючи один одного, але не будучи ніяк погоджені. У

такому випадку достовірну комплексну інформацію одержати практично неможливо, незважаючи на її удаваний надлишок.

За критерієм режиму аналізу даних інформаційно-аналітичні системи підрозділяються на дві категорії:

1) статичні, що включають визначений набір сценаріїв обробки даних і складання звітів (інформаційні системи керівника - ІСК);

2) динамічні, що підтримують побудову і виконання нерегламентованих запитів і формування звітів довільної форми.

СППР, створювані в розрахунку на безпосереднє використання особами, що приймають рішення, являються надзвичайно простими в застосуванні, але жорстко обмеженими у функціональності. Такі статичні СППР, або Інформаційні системи керівника (ІСК) [2, 3] - містять у собі певні множини запитів і, будучи задовільними для повсякденної роботи, нездатні відповісти на всі питання до існуючих даних, виникаючих при прийнятті рішень. Результатом роботи такої системи, як правило, є багатосторінкові звіти. Таким чином, зовнішня простота статичних СППР, за яку активно бореться більшість замовників інформаційно-аналітичних систем, обертається втратою гнучкості. Узагальнена структурна модель СППР представлена на рис.1.



Рис. 1. Структурна модель СППР корпоративного підприємства

Динамічні СППР, навпаки, орієнтовані на обробку нерегламентованих запитів аналітиків до даних. Робота аналітиків з цими системами полягає в інтерактивній послідовності формування запитів і вивчення їх результатів, кожний з яких може породити нову серію запитів. Але динамічні СППР можуть діяти не тільки в області оперативної аналітичної обробки (OLAP), підтримка прийняття управлінських рішень на основі накопичених даних може виконуватися в трьох базових сферах:

- сфера деталізованих даних - це сфера дії більшості систем, націлених на пошук інформації. У більшості випадків реляційні СУБД відмінно справляються з виникаючими тут задачами. Загальновизнаним стандартом мови маніпулювання реляційними даними є SQL. Інформаційно-пошукові системи, що забезпечують інтерфейс кінцевого користувача в задачах пошуку деталізованої інформації, можуть використовуватися в якості надбудови як над окремими системами обробки даних (СОД), так і над сховищем даних (СД) у цілому;

- сфера агрегованих показників. Комплексний погляд на зібрану в сховище даних інформацію, її узагальнення і агрегація, гіперкубічне представлення та багатомірний аналіз є задачами систем оперативної аналітичної обробки даних (OLAP). Тут можна або орієнтуватися на спеціальні багатомірні СУД, або залишатися в рамках реляційних технологій. В другому випадку заздалегідь агреговані дані можуть збиратися в БД зіркоподібного виду, або агрегація інформації може вироблятися в процесі сканування деталізованих таблиць реляційної БД;

- сфера закономірностей. Інтелектуальна обробка виконується методами інтелектуального аналізу даних (ІАД), головними задачами яких є пошук функціональних і логічних закономірностей у накопиченій інформації, побудова моделей і правил, що пояснюють знайдені аномалії і прогнозують розвиток деяких процесів.

Метою побудови корпоративного сховища даних (data warehouse) є інтеграція, актуалізація й узгодження оперативних даних з різномірних джерел для формування єдиного несуперечливого погляду на об'єкт управління в цілому. При цьому в основі концепції сховищ даних лежить визнання необхідності поділу наборів даних, використовуваних для транзакційної обробки, і наборів даних, застосовуваних у системах підтримки прийняття рішень. Такий поділ можливий шляхом інтеграції роз'єднаних у СОД і зовнішніх джерелах деталізованих даних у єдиному сховищі, їхнього узгодження і, можливо, агрегації. В. Інмон, автор концепції сховищ даних, визначає такі сховища як предметно-орієнтовані, інтегровані, немінливі, підтримуючі хронологію набори даних, організовані з метою підтримки управління, покликані виступати в ролі єдиного джерела істини, що забезпечує менеджерів і аналітиків достовірною інформацією, яка необхідна для оперативного аналізу і підтримки прийняття рішень.

Сховище даних являє собою послідовність моментальних знімків корпоративної інформації через визначені, заздалегідь задані, проміжки часу. Воно може поповнюватися щодня, щотижня або щомісяця - це визначається в процесі його організації. Початкові дані для сховища надходять із БД, що підтримують різні бізнес-процеси компанії, а також із зовнішніх джерел. Вхідні дані попередньо повинні бути оброблені, щоб протиріччя реального часу не перешкоджали використовувати інформацію. Інформація цього типу дозволяє аналізувати тенденції або стан на визначений момент часу. Історичний погляд на події має велике значення для прийняття перспективних рішень.

Середовище сховища даних – це складний конгломерат різноманітних інструментів і функцій, що реалізують безперервний процес створення, експлуатації, поступового розширення і зміни інформаційного складу. Основними компонентами такого середовища є програми доступу до джерел даних, інструменти перетворення оперативних і зовнішніх даних, бази даних сховища, програма керування сховищем (адміністратор).

Концепція сховищ даних припускає не просто єдиний логічний погляд на дані організації, а дійсну реалізацію єдиного інтегрованого джерела даних. Альтернативним, стосовно цієї концепції, способом формування єдиного погляду на корпоративні дані є створення віртуального джерела, що спирається на розподілені бази даних різних СОД. При цьому кожний запит до такого джерела динамічно транслюється в запити до вихідних баз даних, а отримані результати узгоджуються, зв'язуються, агрегуються і повертаються до користувача. Однак, при цьому такий спосіб має ряд істотних недоліків:

- час обробки запитів до розподіленого сховища значно перевищує відповідні показники для централізованого сховища. Крім того, структури баз даних СОД у високому ступені нормалізовані, тому в аналітичному запиті до них потрібно звернення до великого числа таблиць, що також приводить до зниження швидкодії.;

- інтегрований погляд на розподілене корпоративне сховище можливий тільки при виконанні вимоги постійного зв'язку всіх джерел даних у мережі. Таким чином, тимчасова недоступність хоча б одного з джерел може або зробити роботу ІС неможливою, або привести до помилкових результатів;

- виконання складних аналітичних запитів над таблицями СОД споживає великий обсяг ресурсів сервера БД і приводить до зниження швидкодії СОД, що неприпустимо, тому що час виконання операцій у СОД часто досить обмежений;

- різні СОД можуть підтримувати різні формати і кодування даних, дані в них можуть бути неузгоджені. Дуже часто на одне й те саме питання може бути отримано кілька варіантів відповіді, що може бути зв'язане з несинхронністю моментів відновлення даних, відмінностями в трактуванні окремих подій, понять і даних, зміною семантики даних у процесі розвитку предметної області, помилками при введенні, утратою фрагментів архівів і т.д. У такому випадку ціль - формування єдиного несуперечливого погляду на об'єкт керування - може не бути досягнута;

- головним же недоліком варто визнати практичну неможливість огляду тривалих історичних послідовностей, тому що при фізичній відсутності центрального сховища доступні тільки ті дані, що на момент запиту є в реальних БД зв'язаних СОД. Основне призначення СОД - оперативна обробка даних, тому вони не можуть зберігати дані за тривалий (більш декількох місяців) період; у міру старіння дані вивантажуються в архів і віддаляються з транзакційної БД. Що стосується аналітичної обробки, то тут саме найбільш цікавий погляд на об'єкт управління в історичній перспективі.

Сховище даних функціонує у такий спосіб: за заданим регламентом в нього збираються дані з різних джерел - баз даних систем оперативної обробки. У сховищі підтримується хронологія: нарівні з поточними зберігаються історичні дані з указанням часу, до якого вони відносяться. У результаті необхідні дані про об'єкт управління збираються в одному місці, приводяться до єдиного формату, узгоджуються і, у ряді випадків, агрегуються до мінімально необхідного рівня узагальнення.

Спрощеним варіантом корпоративного сховища даних можуть бути вітрини даних, що містять інформацію, яка відноситься до окремих аспектів діяльності організації. Концепція вітрин даних була запропонована у 1991 році [4]. При цьому головна ідея полягає в тому, що вітрини даних містять тематичні підмножини заздалегідь агрегованих даних, за розмірами набагато меншими, чим загальне корпоративне сховище даних, і, тим самим вимагають менш продуктивної техніки для їх підтримки. У 1994 році М. Демарест запропонував об'єднати дві концепції і використовувати сховище даних у якості єдиного інтегрованого джерела для численних вітрин даних. У такому варіанті корпоративна інформаційно-аналітична система має трирівневу структуру:

- загальнокорпоративне централізоване сховище даних;
- тематичні вітрини даних на рівні підрозділів;
- робочі місця кінцевих користувачів.

Розглянута концепція орієнтована винятково на збереження, а не на обробку корпоративних даних. Вона не визначає архітектуру цільових аналітичних систем, а тільки створює поле діяльності для їх функціонування, концентруючи увагу на вимогах до даних. Таким чином, вона залишає свободу вибору в усьому, що стосується способів представлення даних у цільовому сховищі (наприклад, реляційний, багатомірний) та режимів аналізу даних сховища. Для того, щоб існуючі сховища даних сприяли прийняттю управлінських рішень, інформація повинна бути представлена аналітику у потрібній формі, тобто він повинен мати розвинуті інструменти доступу до даних сховища та їх обробки.

Висновки. Створення СППР на основі СД – складний та важливий процес, що вимагає знання бізнесу, програмно-технічного інструментарію і досвіду виконання корпоративних проектів. Разом з тим впровадження подібних систем може дати переваги в бізнесі, які будуть тим відчутніші, чим раніше організація почне створення СППР. За прогнозами консалтингової фірми Gartner Group, до 2015 року приблизно 90-95% компаній використовуватимуть СППР.

Значущість інформаційних систем подібного рівня признається і представниками більшості вітчизняних компаній. Проте через низку обставин, ініціативні або замовлені

роботи по розробці та впровадженню СППР ведуться часто достатньо безсистемно, в основному в двох напрямках:

- закупівля і тестування різноманітних продуктів, що використовуються при створенні СППР і СД;
- підвищення продуктивності існуючих систем за рахунок локального перепроектування структури зберігання або використання сучасніших програмних засобів.

ЛІТЕРАТУРА:

1. Огнева А.М. Розробка структури та методики побудови тематичних сховищ даних / Огнева А.М. – Вісник Технологічного університету Поділля. – 2003. – №2. – Т2(50).
2. Ситник В.Ф. Системи підтримки прийняття рішень / Ситник В.Ф., Олексюк О.С. та ін. – К.: Техніка. – 1995. – 162с.
3. Дэниэл Е. Управление корпоративными знаниями / Дэниэл Е. – О’Лири. – Открытые системы. – 1998. – № 4. – С.20-22.

Рецензент: д.т.н., проф. Сбітнєв А.І., провідний науковий співробітник науково-дослідного центру Військового інституту Київського національного університету імені Тараса Шевченка

д.т.н., доц. Шкулипа П.А., к.т.н. Огневый А.В.,
к.т.н. Ленков Е.С., Ищук О.И.

АНАЛИЗ МЕТОДОВ ОБРАБОТКИ КОРПОРАТИВНЫХ ДАННЫХ В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

В статье на основе проведенных исследований научно-методологического характера сформулированы основные признаки и концепции сохранения и представления данных в системах интеллектуальной поддержки принятия решений масштаба корпоративного предприятия

Ключевые слова: хранилища данных, интеллектуальные системы, системы поддержки принимаемых решений, корпоративные системы

Prof. Shkulipa P.A., Ph.D. Ohnevyi A.V., Ph.D. Lenkov E.S., Ishchuk O.I.
ANALYSIS METHODS FOR PROCESSING CORPORATE DATA SYSTEMS SUPPORT
DECISION-MAKING

In this article on the base of research of scientific and methodological nature formulated main features and concepts of data saving and presentation in the smart decision support systems on the scale of a corporate enterprise.

Keywords: data storage, intelligent systems, backed up by a system of decisions, enterprise systems