

УДК 504.05: 004.047

**МОДЕЛІ БАГАТОВИМІРНИХ СТРУКТУР ДАНИХ ДЛЯ АНАЛІЗУ
ПРИРОДООХОРОННОЇ ДІЯЛЬНОСТІ ПРОМИСЛОВИХ ПІДПРИЄМСТВ**

Скарга-Бандурова І. С., Барбарук Л. В., Сіряк Р. В.

**МОДЕЛИ МНОГОМЕРНЫХ СТРУКТУР ДАННЫХ ДЛЯ АНАЛИЗА
ПРИРОДООХРАННОЙ ДЕЯТЕЛЬНОСТИ ПРОМЫШЛЕННЫХ
ПРЕДПРИЯТИЙ**

Скарга-Бандурова И. С., Барбарук Л. В., Сиряк Р. В.

**MULTIDIMENSIONAL DATA STRUCTURES FOR ANALYSING
ENVIRONMENTAL ACTIVITY OF INDUSTRIAL ENTERPRISES**

Skarga-Bandurova I., Barbaruk L., Siryak R.

**Технологічний інститут Східноукраїнського національного університету
ім. В. Даля, м. Сєвєродонецьк, Україна
skarga_bandurova@ukr.net**

Розглянуто задачу використання багатовимірних структур для аналізу даних інвентаризації відходів промислового підприємства. Визначені основні вимоги та запропоновано модель даних, що враховує ієрархії і забезпечує побудову інформаційних відносин вимірювань і гіперкуба даних, і може бути використана як основа високопродуктивних рішень в промислових системах класу ВІ.

Ключові слова: модель, багатовимірні структури, організація даних, куб, вимір, розмірність, ієрархія

Рассмотрена задача использования многомерных структур для анализа данных инвентаризации отходов промышленного предприятия. Предложена учитывающая множественные иерархии модель данных, которая обеспечивает построение информационных отношений измерений и гиперкуба данных и является основой высокопроизводительных решений в промышленных системах класса ВІ.

Ключевые слова: модель, многомерные структуры, организация данных, куб, измерение, размерность, иерархия

The problem of the use of multidimensional data structures for analyzing waste inventory data of industrial enterprise is considered. A data model that takes the hierarchies and provides relationships between dimensions and data hypercube is defined. Proposed model can be used as a basis for high-performance solutions in industrial BI systems.

Keywords: model, multidimensional structure, data cube, dimension, level, hierarchy

1. Вступ

Можливість подання інформації в абстрактній і незалежній від реалізації формі має вирішальне значення в життєвому циклі будь-якої інформаційної системи, не тільки в процесі проектування, але і під час її експлуатації. Це особливо вірно в

контексті створення сховищ даних, оперативної аналітичної обробки даних (OLAP) та вилучення інформації з даних при реалізації природоохоронних заходів, де через складність, розробка додатків і управління ними представляють собою тривалий ітераційний процес, спрямований на постійне вдосконалення існуючої структури і усунення помилок. В основі автоматизації природоохоронної діяльності підприємств лежать концептуальні моделі, разом з тим широко визнається, що традиційні моделі даних, такі як модель сутність-зв'язок для опису багатовимірної і агрегатного характеру додатків OLAP не підходять. З цієї причини, в останні десятиліття були запропоновані різні багатовимірні моделі [1-7]. У той же час, слід зазначити, що, незважаючи на активний інтерес до цієї області, консенсусу з формалізму або навіть єдиної термінології досі не сформувався [8,9]. Складність вирішення завдання розробки моделей обумовлена низкою причин, серед яких як зазначена відсутність єдиної методології, так і широке використання інтуїтивного підходу до їх розробки. Що стосується предметної області цієї роботи, то відносно не вивченими залишаються питання побудови моделі даних для цілей аналітичної обробки екологічної інформації [10], зокрема стосовно задачі інвентаризації відходів промислових підприємств. Причому ці питання набули особливої актуальності після введення в дію в Україні низки нормативно-технічних і правових документів [11-13].

2. Постановка завдання

Метою робіт у зазначеному напрямку є забезпечення реалізації формальних методів багатовимірної моделювання для ефективного аналізу даних і створення механізмів комплексної автоматизації природоохоронної діяльності промислових підприємств в частині реалізації основних аналітичних функцій.

3. Моделі даних

Для вибору базової моделі представлення даних у системі аналізу природоохоронної діяльності було розглянуто ряд моделей що використовуються в сховищах даних і визначено такі основні вимоги до моделі:

- 1) присутність явних ієрархій вимірювань;
- 2) підтримка несуворих ієрархій;
- 3) підтримка декількох ієрархій у вимірі для врахування різних шляхів агрегації, наприклад, у вимірі «час» атрибут дати може згорнутися в тижні або місяці;
- 4) симетричне тлумачення вимірювань і мір, наприклад, атрибут «викид» може бути використаний для обчислення середнього значення, а також визначення груп викидів;
- 5) підтримка правильної агрегації даних, тісно пов'язаних з summarizability [14], так щоб дані не підраховувалися двічі, а неаддитивні дані не додавалися. Наприклад, при підрахунку забруднюючих викидів в різних групах, необхідно вважати одну і ту ж речовину один раз у групі, незважаючи на те, що викид може складатися з декількох речовин;
- 6) підтримка можливості зміни даних з плином часу, наприклад, змін в ієрархії.

За результатами аналізу, в якості базової обрано модель, описану в [15]. Багатовимірні структури визначені як «варіації реляційної моделі, які використовують багатовимірні моделі для організації даних і відображають відносини між даними» [16]. Представлена далі модель заснована на кількох узгоджених концепціях і містить наступні структури: багатовимірний простір, вимір і куб даних.

4. Багатовимірний простір

Формально, багатовимірний простір являє собою набір $\langle \Omega, F \rangle$, в якому Ω - простір вимірювань D (dimension), F - сукупність фактів - мір M (measure). Для кожного вимірювання D_i існує набір рівнів $\text{levels}(D_i) \in \Psi$, такий, що будь-яке вимірювання являє собою структуровану решітку рівнів. Для решітки характерний шлях вимірювання рівня $\text{level}(D_{pi})$, який у свою чергу є лінійним, повністю впорядкованим списком рівнів. Повне впорядкування дозволяє використовувати оператори порівняння для рівнів вимірювань. Враховуючи специфіку процесу інвентаризації відходів, приміром лінійного впорядкування може служити ієрархія виміру «виробництво», така, що «цех» $>$ «технологічний процес» $>$ «елементарний технологічний процес» для кожного шляху вимірювання (рис.1).

У [17] вимір визначається як структурний атрибут куба, який являє собою список елементів, усі з яких мають схожий тип з точки зору сприйняття їх користувачем. Наприклад, місяці, квартали, роки складають вимір часу; всі міста, регіони, країни складають вимір географії і т.д. Вимірювання діє як індекс для визначення значення у багатовимірному масиві.

Нехай Ψ - простір всіх рівнів виміру. Вимірювання, якому належить рівень, може бути знайдено через оператор h : $h(DL_i)=D$ якщо $DL_i \in \text{levels}(D)$, де DL_i - рівень шляху i -го виміру. Нехай також для кожного рівня визначено набір значень $\text{dom}(DL_i)$ у просторі значень V . Рівень вимірювання вважається атомарним, якщо його домен є підмножиною V : $\text{dom}(DL_i) \subset V$. Якщо $\text{dom}(DL_i) \subset |V|$, то рівень вимірювання вважається багатозначним.

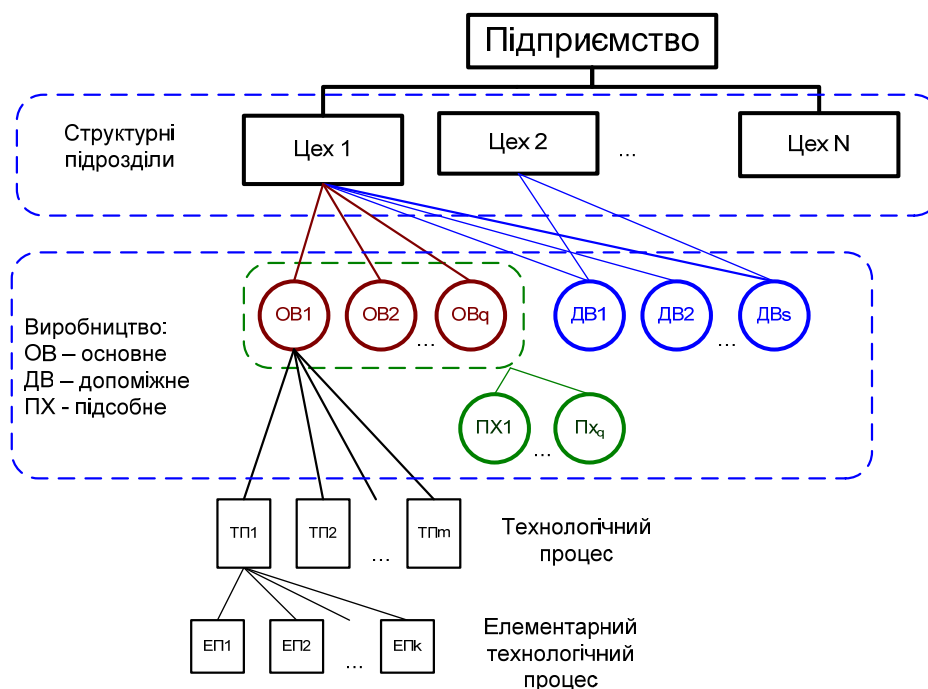


Рис.1. Ієрархія вимірювання «Виробництво»

Нехай також значення x має попередників (ancestor) і нащадків (descendant), $x \in L_0$. Тоді $\text{ancestor}(x, DL)=y$, $y \in \text{dom}(DL)$ і $\text{descendants}(x, DL)=\{x_1, \dots, x_k\}$, де $x_1, \dots, x_k \in \text{dom}(DL)$, $DL < DL_0$. За умови, що відношення ancestor має такі властивості:

- 1) $\text{ancestor}(x, DL) = x$, якщо $x \in \text{dom}(DL)$;

- 2) якщо $x = \text{ancestor}(y, DL) = y$ і $y = \text{ancestor}(x, DL) = y$, тоді $x = y$;
- 3) якщо $x = \text{ancestor}(y, DL_1)$ і $y = \text{ancestor}(z, DL_2)$, тоді $x = \text{ancestor}(z, DL_1)$.

З визначення виміру випливає, що вимірюванню в багатовимірному просторі даних відповідає інформаційне відношення реляційної моделі, а куби будуються з основної бази даних, представленої у вигляді моделі розмірностей.

Модель розмірностей складається з таблиць фактів і таблиць розмірностей, що пов'язує її з базовими сутностями куба. Вимірювання містяться в таблицях фактів. У таблицях розмірностей зберігаються описові ознаки, такі як назви речовини, дати і назви структурних підрозділів. На рис. 2 зображений приклад інформаційних відносин вимірювань і куба даних у вигляді схеми «зірка» [18]. Таблиця фактів IW_F_WASTE_PASSPORT містить унікальний ключ, який об'єднує первинні ключі таблиць вимірів, в яких знаходяться незмінювані, або рідко змінювані дані. У більшості випадків ці дані представляють собою по одному запису для кожного члена нижнього рівня ієрархії виміру [19].

Для реалізації основних операцій аналізу даних у системі аналізу природоохоронної діяльності підприємств описані алгебраїчні операції придатні в гіперкубічних моделях даних для багатовимірних баз даних [20].

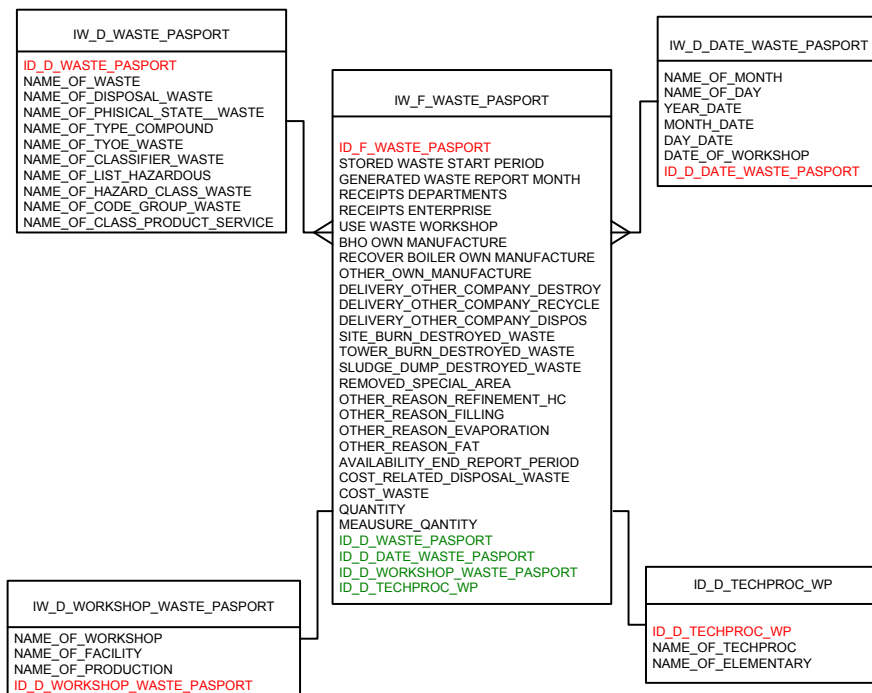


Рис. 2. Структура моделі схеми «зірка» для паспорту відходів

5. Висновки

В статті розглянуто задачу використання багатовимірних структур для інвентаризації відходів і організації ефективного аналізу природоохоронної діяльності підприємств. Визначені основні вимоги та формалізовано модель даних, що враховує множинні ієрархії. Запропоноване рішення передбачає використання механізмів аналітичної звітності, а також сценаріїв вирішення наступних завдань: аналіз впливу змін технологічних умов виробництва, сегментацію відходів, встановлення технологічних процесів, що приносять найбільше екологічне забруднення, прогнозування екологічної шкоди тощо.

Література

1. Балдин, А. В. Адаптируемая модель данных на основе многомерного пространства [Электронный ресурс] / А. В. Балдин, Д. В. Елисеев - Электрон. журн. – М.: «Наука и образование: электронное научно-техническое издание», 2010 – Режим доступа: <http://technomag.edu.ru/doc/161410.html> -12.09.2013 г.
2. Белов, В. Н. Модели многомерного представления и обработки данных на основе алгебры кортежей в информационно-аналитической системе : автореф. дис. канд. техн. наук: 05.13.17, 05.13.01 / В. Н. Белов. – Пенза., 2012. – 20 с.
3. Сергеев, В. П. Векторно-матричная модель представления данных [Текст] / В. П. Сергеев // Вестник компьютерных и информационных технологий. – 2010. – № 6(72). – С.3-14.
4. Томашевський, В. М. Математична модель задачі проектування гібридних сховищ даних з врахуванням структур джерел даних [Текст] / В. М. Томашевський, А. Ю. Яцишин // Вісник НТУУ «КПІ» Інформатика, управління та обчислювальна техніка. – 2011. – №53. – С. 54-61.
5. Boussaid, O. X-Warehousing: An XML-Based Approach for Warehousing Complex Data [Текст] / O. Boussaid, R. B. Messaoud, R. Choquet, R., S. Anthoard // 10th East European Conf. on Advances in Databases and Information Systems (ADBIS). – 2006. – pp.39–54.
6. Torlone, R. Conceptual Multidimensional Models [Текст] / R. Torlone – Chapter 3 in Multidimensional Databases: Problems and Solutions, M. Rafanelli (ed.), Idea Group Inc., 2003. – pp. 69–90.
7. Tryfona, N. starER: A Conceptual Model for Data Warehouse Design [Текст] / N. Tryfona, F. Busborg, J. G. Borch Christiansen // Proc. of ACM 2nd Int. Workshop on Data Warehousing and OLAP (DOLAP). – 1999. – pp. 3-8.
8. Баин, А.М. Современные информационные технологии систем поддержки принятия решений : учеб. Пособие [Текст] / А. М. Баин. – М.: ИД «Форум», 2009. – 240 с.
9. Cuzzocrea, A. CAMS: OLAPing Multidimensional Data Streams Efficiently [Текст] / A. Cuzzocrea // DaWaK 2009. – LNCS 5691. – pp. 48-62.
10. Lepš, J. Multivariate analysis of ecological data using CANOCO [Текст] / J. Lepš, P. Šmilauer. – Cambridge University Press, 2003. – 270 p.
11. Постанова Кабінету Міністрів від 1 листопада 1999 р. N 2034 «Про затвердження Порядку ведення державного обліку та паспортизації відходів».
12. Методическое руководство для проведения национальной инвентаризации опасных отходов в рамках Базельской конвенции [Электронный ресурс] / Серия / СБК No: 99/009 (E) Женева, Май 2000. – Режим доступа: [http://www.basel.int/Portals/4/Basel%20Convention/docs/pub/metological guide-r.pdf](http://www.basel.int/Portals/4/Basel%20Convention/docs/pub/metological_guide-r.pdf) - 10.02.2014 г.
13. Распоряжение Главы областной государственной администрации № 1323 от 2.11.2010 г. «Об утверждении мероприятий по организации ведения государственного учета и паспортизации отходов».
14. Lenz, H.-J. Summarizability in OLAP and Statistical Data Bases [Текст] / H.-J. Lenz, A. Shoshani // Proc. of 9th Int. Conf. on Scientific and Statistical Database Management. – 1997. – pp.132-143.

15. *Vassiliadis, P.* Modeling Multidimensional Databases, Cubes and Cube Operations [Текст] / P. Vassiliadis // Proc. of the 10th Int. Conf. on Scientific and Statistical Database Management. – 1998. – pp. 53-62.
16. *O'Brien, J.* Management Information Systems [Текст] / J. O'Brien, G. Marakas. - McGraw-Hill/Irwin, 2010. – 712 p.
17. *OLAP AND OLAP Server Definitions* [Електронний ресурс] / – Режим доступа: www. URL: <http://olapcouncil.org/research/resrchly.htm> -14.01.2014 г.
18. *Giovinazzo, W. A.* Object-oriented data warehouse design: building a star schema [Текст] / W. A. Giovinazzo. – Prentice Hall PTR, Upper Saddle River, NJ, 2000. – 368 p.
19. *Соловьев, С. В.* Технология разработки прикладного программного обеспечения [Текст] / С. В. Соловьев, Р. И. Цой, Л. С. Гринкруг. – М.: Академия естествознания, 2011. – 407 с.
20. *Скарга-Бандурова И.С.* Методы многомерного моделирования в задачах экологического мониторинга и инвентаризации отходов [Текст] / И. С. Скарга-Бандурова // Радіоелектронні і комп'ютерні системи. – 2013. – № 5(64). – С. 96-99.

УДК 001.573.681

МОДЕЛЬНЕ ДОСЛІДЖЕННЯ ЕВОЛЮЦІЙНИХ ЗМІН ПІДСИСТЕМИ «ГРУНТ» В РЕГІОНАЛЬНИХ ЕКОСИСТЕМАХ

Котовенко О.А., Мірошніченко О.Ю., Березницька Ю.О.

МОДЕЛЬНОЕ ИССЛЕДОВАНИЕ ЭВОЛЮЦИОННЫХ ИЗМЕНЕНИЙ ПОДСИСТЕМЫ «ПОЧВА» В РЕГИОНАЛЬНЫХ ЭКОСИСТЕМАХ

Котовенко Е.А., Мирошніченко Е.Ю., Березницькая Ю.О.

REGIONAL ECOSYSTEMS SUBSYSTEM “SOIL” EVOLUTIONAL CHANGES MODEL RESEARCH

Kotovenko O., Miroshnychenko O., Bereznytskaya J.

Кафедра охорони праці і навколишнього середовища КНУБА, Київ, Україна,
kafopnsknuba@ukr.net

Для дослідження процесів міграції та накопичення речовин у ґрунтах визначеного типу запропонована модель, яка базується на методі імовірісно-автоматного моделювання. Це один з найперспективніших методів при встановленні змін станів підсистеми «ґрунт» та знаходження необхідних її характеристик та показників. Для конкретного регіону проводився аналіз екологічної ємності ґрунтів та коефіцієнту їх вразливості.

Ключові слова: автоматно-імовірносне моделювання, ґрунти, регіональна екосистема, індекс вразливості, потужність техногенного впливу