

ВИКОРИСТАННЯ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ В СИСТЕМАХ АНАЛІЗУ ХІМІЧНО-НЕБЕЗПЕЧНИХ НАДЗВИЧАЙНИХ СИТУАЦІЙ НА ЗАЛІЗНИЧНОМУ ТРАНСПОРТІ

У статті розглянуто доцільність використання OLAP-технологій для аналізу хімічно-небезпечних надзвичайних ситуацій на залізничному транспорті. Розглянуто недоліки класичних підходів до аналізу та зберігання даних, а також переваги багатовимірних моделей даних і технологій що використовуються для аналізу інформації, що у них зберігається. Проаналізовано доцільність використання у даній предметній області сховищ даних та інтелектуального аналізу цих даних. Розглянуто схеми подання інформації у сховищах даних та проаналізовано переваги та недоліки кожної.

In the article expedience of the use of OLAP-technologies is considered for an analysis chemically of extraordinary near-accidents on a railway transport. The lacks of the classic going near an analysis and storage are considered information, and also advantages of multidimensional models of data and technologies that utilized for the analysis of information, that for them is saved. Expedience of the use in this subject domain of depositories of information and intellectual analysis of these data is analysed. The charts of presentation of information are considered in the depositories of information and advantages and failings are analysed to each.

Вступ

На теперішній час, за рахунок стрімкого розвитку промисловості, розбудови і розширення міст та розвитку залізничного транспорту, особливо того, що орієнтований на перевезення небезпечних вантажів, зростає можливість виникнення під час їх перевезень надзвичайних ситуацій. Особливої уваги потребують надзвичайні ситуації, що мають місце при перевезенні залізницею хімічно-небезпечних вантажів, оскільки значна частина залізничних шляхів проходять через населенні пункти, де виникнення хімічних аварій може призвести до загибелі великої кількості людей або навіть до екологічних катастроф [1, 2]. Виходячи з цього, постає необхідність у впровадженні технологій проведення інтелектуального аналізу хімічно-небезпечних надзвичайних ситуацій з метою підвищення швидкості їх ліквідації та зменшення негативних наслідків від них в короткий термін часу або запобігання таких наслідків [3]. Це потребує швидкого опрацювання великих об'ємів інформації. Серед новітніх технологій, що здатні швидко та ефективно опрацьовувати великі об'єми інформації, виділимо такі: OLAP, KDD, Data Mining.

KDD (Knowledge Discovery in Databases) – аналітичний процес дослідження значних обсягів інформації із залученням засобів автоматизації, що має на меті виявити приховані у множині даних структури, залежності й взаємозв'язки. При цьому передбачається повна чи часткова відсутність апріорних уявлень про характер прихованих структур та залежностей. KDD передбачає, що людина попередньо осмислює задачу й подає неповне (у термінах цільових змінних) її формулювання, перетворює дані до формату придатного для їх автоматизованого аналізу й попередньої обробки, виявляє засобами автоматичного дослідження даних приховані структури й залежності, апробовує виявлені моделі на нових даних, не використовуваних для побудови моделей, та інтерпретує виявлені моделі й результати [4].

Data Mining (дослівно: «Розробка, добування даних» – DM) – дослідження «сирих» даних і виявлення в них за допомогою «машини» (алгоритмів, засобів штучного інтелекту) прихованих нетривіальних структур і залежностей, які раніше не були відомі й мають практичну цінність та придатні для того, щоб їх інтерпретувала людина [4], але це не є ефективним, коли заздалегідь невідомі хоча б деякі алгоритми, які можна було б інтерпретувати в конкретному випадку для аналізу даних визначеної предметної області.

OLAP (online analytical processes, аналітична обробка у реальному часі) – технологія комплексного багатовимірного аналізу даних, що дозволяє виконувати швидкий та ефективний аналіз над потужними множинами даних [4]. При цьому дані зберігаються у вигляді, який відображає стан у реальному масштабі часу, що є доцільним для проведення швидкого та ефективного аналізу інформації.

Мета дослідження

Аналіз можливості та доцільності використання OLAP-технологій для інтелектуального аналізу хімічно-небезпечних надзвичайних ситуацій на залізничному транспорті.

Розв'язування задачі

З урахуванням того, що інтелектуальний аналіз передбачає обробку даних та знань, доцільним є застосування новітніх технологій аналізу даних.

Технології, що базуються на застосуванні класичних статистичних підходів, мають низку недоліків. Відповідні методи ґрунтуються на використанні усереднених показників, на підставі яких важко з'ясувати справжній стан справ у досліджуваній сфері [4].

Окрім того, стандартні статистичні методи відкидають розгляд критичних нетипових ситуацій, хоча такі ситуації можуть становити окремий інтерес для дослідження, характеризуючи виняткові, але важливі явища. Проте окремі нетипові значення можуть становити окремий інтерес для дослідження,

характеризуючи виняткові, але важливі явища. Навіть сама ідентифікація цих спостережень, не говорячи про їх подальший аналіз і докладний розгляд, може бути корисною для розуміння сутності досліджуваних об'єктів чи явищ [3].

Враховуючи вищезазначені недоліки класичних статистичних підходів аналізу та обробки даних і знань, можна стверджувати, що їх застосування до задачі аналізу хімічно-небезпечних надзвичайних ситуацій на залізниці не є доцільним, оскільки:

1. Використання усереднених показників, на яких ґрунтується статистичний підхід, неможливе під час аналізу хімічно-небезпечних надзвичайних ситуацій на залізниці, оскільки аналіз таких ситуацій передбачає точну обробку конкретних показників визначеної надзвичайної ситуації (радіус розповсюдження первинної хімічної хмари, зона зараження тощо).

2. Під час аналізу хімічно-небезпечних ситуацій на залізничному транспорті слід брати до уваги й нетипові дані чи спостереження, так як вони можуть мати в певних випадках вирішальне у надзвичайній ситуації значення.

При проведенні аналізу хімічно-небезпечних надзвичайних ситуацій розв'язують такі інтелектуальні задачі:

- класифікація (виявлення ознак, що характеризують надзвичайну ситуацію, щоб класифікувати її вид, наслідки, загрозу тощо);

- кластеризація (застосовується у випадку, коли групи ознак не визначені заздалегідь, що потребує самостійного виділення подібних групи даних, наприклад, умов виникнення надзвичайних ситуацій);

- виявлення послідовностей (має місце, коли існує ланцюг подій, зв'язаних у часі, наприклад, таких, що передують певній надзвичайній ситуації);

- асоціація (зв'язок змінних між собою, наприклад, зона зараження залежить від кількості і токсичності речовин, що перевозяться);

- прогнозування (базується на історичній інформації, що міститься у сховищах даних).

З урахуванням можливостей OLAP-технології як сукупності засобів багатовимірного експрес-аналізу даних, що містяться у сховищі даних [4, 5], розглянемо можливість їх застосування при аналізі хімічно-небезпечних ситуацій на залізничному транспорті.

Для того, щоб використовувати OLAP-технології, необхідно відмовитись від класичних підходів до інтелектуального аналізу даних, що означає, насамперед, відмову від плоских моделей зберігання інформації, адже їхньої швидкодії та функціональності для застосування швидкодючих методів інтелектуального аналізу недостатньо.

При цьому, необхідно використовувати проти класичної організації баз даних сховища даних, як такі, що передбачають зберігання та швидку обробку потужних інформаційних структур, в тому числі і змінних у часі, що є необхідними під час аналізу хімічно-небезпечних надзвичайних ситуацій на залізниці.

Обґрунтування доцільності використання сховищ даних під час інтелектуального аналізу хімічно-небезпечних надзвичайних ситуацій на залізниці є те, що вони здатні зберігати не тільки статистичні дані, але й знання та додаткові дані в одному місці, відповідають підвищеним вимогам опрацьовувати високо агреговані дані. Це надасть змогу прискорити процес аналізу та задіяти в повній мірі аналіз додаткових даних, що можуть мати вплив на формування кінцевого рішення щодо результату аналізу надзвичайної хімічно-небезпечної ситуації на залізничному транспорті.

Розглянемо варіанти подання зв'язків предметної області хімічно-небезпечних надзвичайних ситуацій на залізниці у багатовимірних сховищах даних в рамках вимог OLAP-технологій.

OLAP-технологія передбачає використання схем «зірка» та «сніжинка» при збереженні даних в таблиці фактів і таблицях вимірів.

Таблиця фактів містить один рядок для кожного факту в кубі. Для кожного виміру при цьому вводиться такі стовпці:

1) із значеннями параметрів для конкретного факту;

2) з показником, що вноситься в комірку кубу, та сприймається, як індекс, який використовується для ідентифікації конкретних значень, тобто стовпець для кожного виміру, який містить зовнішній ключ, що посилається на таблицю вимірів для конкретного виміру. Показником виступає поле (як правило це – цифрове поле), значення якого однозначно визначається фіксованим набором вимірів. Залежно від того, як формуються його значення, показник може бути:

- змінною (Variable): значення таких показників один раз вводиться з зовнішнього джерела або формуються програмно, а потім у явному вигляді зберігаються в багатовимірній базі даних;

- формулою (Formula): значення таких показників обчислюються за визначеним виразом.

Кожна схема має свої переваги та недоліки. На вибір схеми подання даних вагомий вплив має рівень агрегації збережених даних

Розглянемо подання зв'язків сховища даних предметної області «Хімічно-небезпечні надзвичайні ситуації на залізниці» за схемою «зірка», що практично є реляційним представленням багатовимірного подання даних – основи OLAP-технології. Схема «зірки» являє собою проекцію метакуба на «реляційну площину». В нашому випадку вона складатиметься з взаємопов'язаних однієї таблиці фактів і трьох таблиць вимірів за рисунком 1.

Таблиця фактів містить по одному рядку для кожного з фактів – мінімальний атом аналізуемого процесу.

Таблиці вимірів містять саме значення вимірів, тобто інформацію, що характеризує факти, для прикладу, це час, вантаж та транспорт (див. рис. 1). Схема «зірка» не відповідає вимогам теорії нормалізації. Такі атрибути, як наприклад назва чи кількість можуть зустрічатись в декількох таблицях. З точки зору теорії нормалізації, таблиці вимірів зберігають надлишкові дані, але ця надлишковість виправдовується двома фактами: схема більш зрозуміла користувачу, запити до бази даних будуть виконуватись швидше за рахунок зменшення кількості таблиць, що слід при їх реалізації опрацювати.

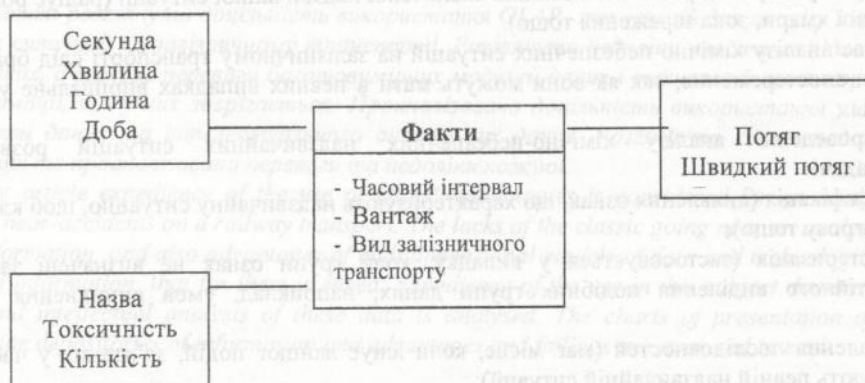


Рис. 1. Подання зв'язків предметної області «Хімічно-небезпечні надзвичайні ситуації на залізниці» у сховищі даних за схемою «зірка»

Схема «сніжинки» – це модифікація схеми «зірка». У цій схемі частина таблиць вимірів розбита на декілька зв'язаних таблиць. Це надає змогу більш гнучко проводити аналіз наявних фактів. Подання зв'язків у сховищі даних за схемою «сніжинка» наведено на рис. 2.



Рис. 2. Подання зв'язків предметної області «Хімічно-небезпечні надзвичайні ситуації на залізниці» у сховищі даних за схемою «сніжинка»

В результаті проведеного дослідження з урахуванням вимог до схеми багатовимірної моделі даних, яка є основою OLAP- технології, можна зробити такі висновки щодо питань проектування архітектури багатовимірного сховища даних, призначеного для аналізу хімічно-небезпечних ситуацій на залізниці:

- при збереженні багатовимірної інформації у сховищі даних доцільно будувати куб на основі бази даних, що має структуру «сніжинка». В центрі бази даних знаходиться таблиця фактів, яка містить виміри багатовимірної бази даних. Крім того, до бази входить ще певна кількість таблиць вимірів. В цих таблицях зберігаються записи, що характеризують дані, які знаходяться в таблиці фактів;

- структура бази даних «сніжинка» є найбільш доцільною для аналізу хімічно-небезпечних надзвичайних ситуацій на залізниці. При цьому, в центрі бази даних міститиметься інформація стосовно надзвичайних ситуацій, що можуть виникнути при перевезенні хімічно-небезпечних речовин, а інші таблиці будуть містити інформацію щодо можливих за таких обставин надзвичайних ситуацій.

Слід зазначити, що така структура не є однозначною, це лише фрагмент повної структури. Вона може змінюватись таким чином, як того вимагає конкретна задача (залежно від агрегації даних та параметрів системи).

Рівень агрегації даних при використанні їх в багатовимірних моделях даних досить високий, і тому обсяг самої бази даних є значно меншим відносно об'ємів інформації, яка зберігається, що є досить важливим і актуальним саме для систем аналізу хімічно-небезпечних надзвичайних ситуацій на залізниці.

Проектування багатовимірної моделі даних починається з визначення питань, за якими кінцеві користувачі могли б оперувати системою.

У системах аналізу хімічно-небезпечних ситуацій на залізничному транспорті найбільш ефективним є використання двох основних типів багатовимірних моделей даних:

1) метакубічної моделі, яка передбачає, що всі показники повинні визначатися одним набором вимірів з можливим введенням надлишкових для них вимірів;

2) полікубічної моделі, для якої характерно, що в багатовимірній базі даних може бути визначено кілька гіперкубів з різною розмірністю та вимірами в якості граней, що не зв'язані між собою. Це вимагає формування різних метакубів, які можуть розрізнятися своєю розмірністю

Грані метакуба обираються з урахуванням вихідного вектору, що є визначальним при прийнятті рішення в хімічно-небезпечній надзвичайній ситуації, тому число неіснуючих значень у комірках гіперкуба відносно невелике.

Обравши, таким чином, рівень агрегації у часі, можна проаналізувати швидкоплинну техногенну ситуацію і спрогнозувати динаміку її розвитку як загалом, так і у визначених часових інтервалах, що дозволить запобігти руйнівним наслідкам та людським ураженням.

Вибір рівня агрегації <Хімічні речовини, що входять до складу вантажів> дозволяє перейти на якісно новий рівень аналізу. На цьому рівні можна буде враховувати взаємозв'язки між конкретним <Вантажем>, <Видом залізничного транспорту> і <Часовим інтервалом>. Отже, можна проаналізувати не тільки саму надзвичайну ситуацію в рухомому складі залізниці, але й всебічно проаналізувати розвиток хімічно-небезпечного процесу та вплив на нього складової вантажу, що транспортується, у часі.

Основними операціями маніпулювання вимірами є: зріз, обертання, агрегація та деталізація. Ефективне використання цих операцій можливе лише при використанні об'єктно-орієнтованого програмування, що характеризуються такими його важливими характеристиками як наслідування, інкапсуляція та поліморфізм. Крім того, чітка ієрархія структури дає змогу однозначно визначити класову структуру даних та, використовуючи наслідування, побудувати програмний комплекс системи аналізу хімічно-небезпечних ситуацій на залізничному транспорті.

Розглянемо операції з вимірами стосовно їх використання в системах аналізу хімічно-небезпечних ситуацій на залізничному транспорті:

1) Операція "Зріз". Користувача рідко цікавлять усі потенційно можливі комбінації значень вимірів. Більш того, він практично ніколи не працює одночасно відразу з всім метакубом даних. Підмножина метакуба, що утворилася в результаті фіксації значення одного або декількох вимірів, називається зрізом. Наприклад, якщо обрати значення виміру <Часовий інтервал> отримаємо підмножину метакуба, змістом якого буде інформація про всі хімічні процеси, що прогноуються, мали й мають місце у швидкоплинній надзвичайній ситуації, яка виникла при перевезенні хімічно-небезпечного вантажу, у часі.

2) Операція "Обертання". Зміна порядку подання візуалізації вимірів (звичайно застосовується при двовірних моделях даних) називається обертанням. Ця операція забезпечує можливість візуалізації даних у зручній та зрозумілій для користувача формі.

3) Операція "Агрегація". У процесі аналізу користувач не тільки працює з різними зрізами даних і виконує їх обертання, але і переходить від деталізованих даних до агрегованих, тобто виконує операцію агрегації.

4) Операція "Деталізація". Перехід від більш агрегованих до більш деталізованих даних називається операцією Деталізації.

Наприклад, почавши аналіз на рівні <Хімічно-небезпечна надзвичайна ситуація>, користувач може одержати детальну інформацію про <Вантаж> або <Вид залізничного транспорту>.

В багатовимірних моделях даних можна встановити автоматичний тип відношень без їх додаткового описання. Наприклад, для вимірів, що мають тип <Часовий інтервал> автоматично встановлюється відношення щодо <Секунда>, <Хвилина>, <Година>, <Доба>. Множина відношень може мати ієрархічну структуру, що є зручнішим від встановлення нових вимірів. У цьому випадку всі потенційно можливі значення з різних вимірів поєднуються в єдину множину.

Означений підхід дозволяє чітко виділити рівні ієрархії для всіх об'єктів, які описуються у системі, і побудувати чітку та однозначну структуру даних, що є найбільш вдалим варіантом для аналізу та прийняття рішень, оскільки при цьому можна буде використовувати OLAP-технологію та Data Mining (інтелектуальний аналіз даних). Це надає велику кількість переваг порівняно з класичними реляційними базами даних, для аналізу яких використовується мова SQL.

Отже, багатовимірні моделі даних усе частіше використовуються в різних галузях не тільки як самостійний програмний продукт, але і як аналітичні засоби обробки даних, що дозволяють підтримувати значні за об'ємами бази даних, компактно зберігати їх в агрегованому вигляді та легко маніпулювати ними, що має особливе значення для систем аналізу хімічно-небезпечних надзвичайних ситуацій на залізничному транспорті.

Висновки

Таким чином, з урахуванням вимог, що висуваються до процесу прийняття рішень в можливих швидкоплинних хімічно-небезпечних надзвичайних ситуаціях на залізницях при перевезенні вантажів, доцільним є використання OLAP-технологій при організації даних та знань у реальному часі як таких, що характеризуються високою швидкістю їх обробки, пристосованістю до змін та доповнень в будь-який момент часу та можливістю застосування інтелектуальних методів та підходів до їх аналізу з подальшим

прийняттям рішень відповідними системами.

Література

1. Методика прогнозування масштабів зараження сильнодействующими ядовитими веществами при авариях (разрушениях) на химически опасных объектах и транспорте. – М.: ВЦК ГО, 1990. – 28 с.
2. Ліквідація пожеж на залізничному транспорті / Кацман М.Д., Кононов Г.Б., Діденко Т.В., Огороднійчук Н.В.. – К.: вид-во "Основа", 2006. – 215 с.
3. Пупков К.А., Коньков В.Г. Интеллектуальные системы. – М.: Изд-во МГТУ им. Н.Э.Баумана, 2003.
4. Барсегян А.А., Куприянов М.С. Методы и модели анализа данных: OLAP и Data Mining.
5. Codd, S.B. Codd, C.T. Salley. Providing OLAP (On-Line Analytical Processing) to User-Analysts: An IT Mandate. – E.F.Codd & Associates, 1998.

Надійшла 8.4.2009 р.

УДК 621.321

І.В. ТРОЦИШИН, О.П. ВОЙТЮК, Л.В. ТРОЦИШИНА

Хмельницький національний університет

ВИМІРЮВАННЯ ЧАСТОТИ ЗА МЕТОДОМ КОІНЦИДЕНЦІЇ ТА ОСОБЛИВОСТІ УТВОРЕННЯ ШКАЛИ ВИМІРЮВАЛЬНОГО ПЕРЕТВОРЕННЯ

Розглянуті питання утворення виміральної шкали вимірювання частоти за методом коінцидентції. Проаналізовано утворення дробової та дробово-раціональної шкал вимірювання. Вперше вказано на особливості поведінки позначок та їх розподіл на відрізках шкали.

The considered questions of formation of measuring scale of measuring of frequency are after the method of coincidence. It is analysis formation of shot and shot-rationals scales of measuring. It is first indicated on the features of conduct of poacher and their distributing on the segments of scale.

Вступ

Важливість вимірювань для розвитку суспільства має визначальну роль, а особливо це стосується наукових досліджень, про що сказав ще Д.І. Менделєєв: "Наука починається лише тоді, коли починають вимірювати". Цей відомий всім вислів є фундаментальним, але досить узагальненим, тому на практиці в галузі метрології та виміральної техніки виникають проблеми, які пов'язані із досить примітивними підходами до процесу вимірювання, і, як правило, до спрощеного їх розгляду. Таких прикладів доволі багато, але ми розглянемо до яких наслідків це може призвести на прикладі вимірювання частоти та її параметрів, які займають майже половину вимірювань в радіотехнічних і телекомунікаційних системах.

В роботах [1-8] питання вимірювання частоти за методом коінцидентції знайшли своє відображення, але частина питань залишилась недослідженою. Тому утворення шкали вимірального перетворення є визначальним для вірного розуміння всієї процедури вимірювань та використання їх результатів.

Метою статті є спроба показати деякі нові особливості утворення та поведінки виміральної шкали для методу коінцидентції, з позицій вимірювання частоти із апіорі невідомим часом існування сигналу, тобто реальних швидкоплинних сигналів та процесів.

Класичні схеми вимірювання відомі [1, 2, 5] і представлені на рис. 1 та рис. 2.

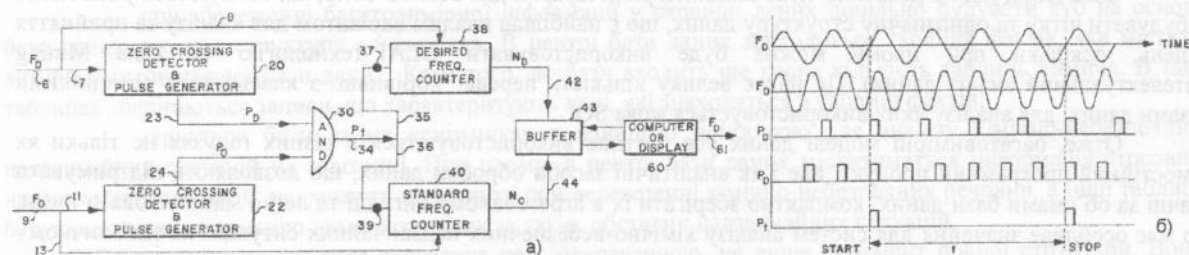


Рис. 1. Запатентована схема вимірювання за методом коінцидентції [1]

Всі схеми досить подібні, але лише за схемою [8], отриманою на основі фазочастотного підходу в рамках фазочастотної теорії вимірювання та перетворення радіосигналів [4, 5, 8, 10], було побудовано та випробувано ряд частотомірів та вимірвальних перетворювачів, які дали унікальні результати [8, 10], але деталі процесу процедури вимірювання залишились поза увагою і не публікувались. Тому наведені нижче результати вперше детально покажуть утворення виміральної шкали (так званої шкали відношень) не