

УДК 621.3 : 004.1

В.А. Краснобаєв¹, Я.Ю. Стасєва¹, О.С. Луцаєвський²

¹Харківський університет Повітряних Сил ім. І. Кожедуба, Харків

²Харківський національний університет радіоелектроніки, Харків

АНАЛІЗ МАТЕМАТИЧНОГО ЗАБЕЗПЕЧЕННЯ ІСНУЮЧИХ СИСТЕМ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ

У роботі наведено аналіз найбільш широко розповсюджених методів та засобів інтелектуального аналізу даних, що застосовуються при розробці спеціального математичного забезпечення систем підтримки прийняття рішень

система підтримки прийняття рішень, спеціальне математичне забезпечення

Вступ

Сучасні системи підтримки прийняття рішення (СППР), що виникли як природний розвиток і продовження управлінських інформаційних систем та сис-

тем управління базами даних, являють собою системи максимально пристосовані до рішення задач повсякденної управлінської діяльності, є інструментом, покликаним зробити допомогу людям, що приймають рішення (ЛПР).

За допомогою СППР можуть вирішуватися неструктуровані та слабоструктуровані багатокритеріальні задачі. СППР, як правило, є результатом мультидисциплінарного дослідження, що включає теорії баз даних, штучного інтелекту, інтерактивних комп'ютерних систем, методів імітаційного моделювання.

СППР має наступні основні характеристики [5, 7]:

- СППР використовує і дані, і моделі;
- СППР призначені для допомоги менеджерам у прийнятті рішень для слабоструктурованих і неструктурованих задач;

- вони підтримують, а не заміняють, вироблення рішень менеджерами;

- ціль СППР – поліпшення ефективності рішень.

З аналізу робіт присвячених СППР [1 – 8] можна визначити загальні характеристики ідеальної СППР:

- оперує зі слабоструктурованими рішеннями;
- призначена для ЛПР різного рівня;
- може бути адаптована для групового й індивідуального використання;
- підтримує як взаємозалежні, так і послідовні рішення;

- підтримує 3 фази процесу рішення: інтелектуальну частину, проектування і вибір;

- підтримує різноманітні стилі і методи рішення, що може бути корисно при рішенні задачі групою ЛПР;

- є гнучкою й адаптується до змін як організації, так і її оточення;

- проста у використанні і модифікації;

- поліпшує ефективність процесу прийняття рішень;

- дозволяє людині керувати процесом прийняття рішень за допомогою комп'ютера, а не навпаки;

- підтримує еволюційне використання і легко адаптується до вимог, що змінюються;

- може бути легко побудована, якщо може бути сформульована логіка конструкції СППР;

- підтримує моделювання;

- дозволяє використовувати знання.

Ці характеристики визначають як архітектуру СППР, так і структуру її математичного забезпечення (МЗ).

Щодо архітектури СППР, то вона представляється різними авторами по-різному, так відома узагальнена архітектура Маракаса (Marakas) складається з 5 різних частин, що представлені на рис. 1.

Ця архітектура в цілому відображає сучасні тенденції в проектуванні СППР, але якщо розглядати МЗ, то тут ми зіткнемося з великою кількістю різноманітних підходів та аналітичних методів, що використовуються його розробниками. Тому метою цієї статті є розгляд методів, моделей та засобів аналізу даних які найчастіше застосовуються в СППР.

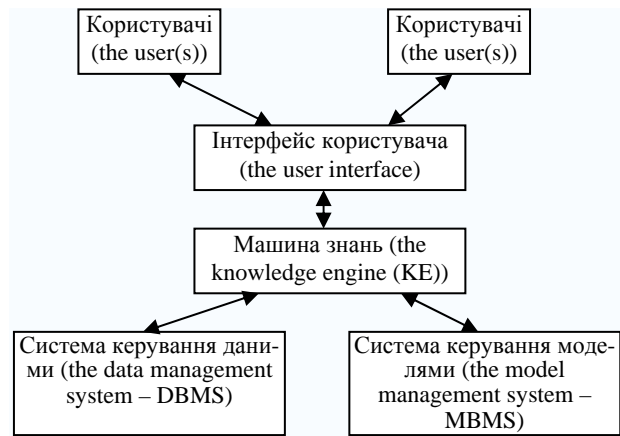


Рис. 1. Архітектура СППР

Викладенн основного матеріалу

Аналітичні методи дають кінцевому користувачеві можливість здійснити весь цикл роботи з вихідними даними, що мають великі обсяги і нез'ясовану статистичну структуру. Цей цикл називається розвідкою даних (Data Mining) і складається з декількох етапів: вибірка, дослідження, модифікація, моделювання, оцінка результатів (Sample, Explore, Modify, Model, Assess).

При роботі додатка на етапі вибірки відбувається формування підмножини спостережень з вихідних даних (добір за критеріями або випадковий добір). На етапах дослідження і модифікації можуть бути здійснені: фільтрація даних, відкидання даних з великими викидами, перетворення вихідних перемінних. На етапі моделювання здійснюється побудова регресій і оптимізація підмножини перемінних, прийняття рішень на основі методик нейронних мереж, що реалізують різні алгоритми навчання класифікації об'єктів, побудова класифікаційних дерев для добору оптимального набору перемінних й оптимальні розбивки множини об'єктів, кластеризація й оптимальне угруповання об'єктів. Нарешті, на етапі огляду й оцінки результатів користувач має можливість зіставити різні результати моделювання, вибрати оптимальні класи і параметри моделей, представити результати аналізу в зручній формі.

На етапі підготовки даних забезпечується доступ до будь-яких реляційних баз даних, текстовим і SAS-файлам. Додаткові засоби перетворення й очищення даних дозволяють змінювати вид представлення, проводити нормалізацію значень, виявляти невизначені або відсутні значення. На основі підготовлених даних спеціальні процедури автоматично будують різні моделі для подальшого прогнозування, класифікації нових ситуацій, виявлення аналогій. Дані додатки підтримують структури п'яти різних типів моделей, до яких відносяться:

- нейронні мережі,
- класифікаційні і регресійні дерева рішень,
- найближчі k-околиці,

- байєсівське навчання,
- кластеризація.

Розглянемо більш докладно засоби інтелектуального аналізу даних (ІАД, Data Mining), застосовувані в СППР.

Як перший напрямок розвитку засобів ІАД варто виділити методи статистичної обробки даних, який можна розділити на чотири взаємозалежних розділи:

- попередній аналіз природи статистичних даних (перевірка гіпотез стаціонарності, нормальності, незалежності, однорідності, оцінка виду функції розподілу і її параметрів);
- виявлення зв'язків і закономірностей (лінійний і нелінійний регресійний аналіз, кореляційний аналіз);
- багатомірний статистичний аналіз (лінійний і нелінійний дискримінантний аналіз, кластер-аналіз, компонентний аналіз, факторний аналіз);
- динамічні моделі і прогноз на основі часових рядів.

Серед найбільш відомих і популярних засобів статистичного аналізу варто назвати пакети Statistica, SPSS, Systat, Statgraphics, SAS, BMDP, TimeLab, DataDesk, SPlus, Scenario (BI), "Мезозавр".

Особливий напрямок у спектрі аналітичних засобів ІАД складають методи, засновані на нечітких множинах. Їхнє застосування дозволяє ранжирувати дані по ступеню близькості до бажаних результатів, здійснювати так званий нечіткий пошук у базах даних. Однак платою за підвищену універсальність є зниження рівня вірогідності і точності одержуваних результатів. Тому число спеціалізованих додатків даного методу як і раніше невелике, незважаючи на те, що протягом останніх 35 років математики - прикладники виявляли до нього підвищений інтерес.

Другий великий напрямок розвитку складають кібернетичні методи оптимізації, засновані на принципах систем, що саморозвиваються – методи нейронних та семантичних мереж, еволюційного і генетичного програмування.

Однак нові достоїнства породжують і нові проблеми. Зокрема, рішення, отримані кібернетичними методами, часто не допускають наочних інтерпретацій, що деякою мірою ускладнює життя предметним експертам. До програмних продуктів, що використовують кібернетичні методи ІАД, відносяться системи PolyAnalyst, Neur-oShell, GeneHunter, BrainMaker, OWL, 4Thought (BI).

Безпосередньо до кібернетичних методів ІАД примикають синергетичні методи. Їхнє застосування дозволяє реально оцінити обрїй довгострокового прогнозу. Особливий інтерес викликають дослідження, зв'язані зі спробами побудови ефективних систем керування в хитливих режимах функціонування.

До третього великого розділу ІАД варто віднести сукупність традиційних методів рішення оптимізаційних задач – варіаційні методи, методи дослідження операцій, що включають у себе різні види математичного програмування (лінійне, нелінійне,

дискретне, цілочисельне), динамічне програмування, принцип максимуму Понтрягіна, методи теорії систем масового обслуговування. Програмні реалізації більшості цих методів входять у стандартні пакети прикладних програм, наприклад Math CAD і MatLab.

У четвертий розділ засобів ІАД входять засоби, що назвемо умовно експертними, тобто зв'язаними з безпосереднім використанням досвіду експерта. До їхнього числа відносять метод "найближчого сусіда", що ліг в основу таких програмних продуктів, як Pattern Recognition Workbench або KATE tools.

Інший підхід до вибору рішення зв'язаний з побудовою послідовного логічного висновку - дерева рішень, у кожному вузлі якого експерт здійснює найпростіший логічний вибір ("так" – "ні"). У залежності від прийнятого вибору, пошук рішення просувається по правій або лівій частині дерева і зрештою приходиться до термінальної частини, що відповідає конкретному остаточному рішення. Тут процес статистичного навчання виведений за межі програми і сконцентрований у виді деякого апріорного досвіду, укладеного в наборі галузей-рішень.

Одним з різновидів методу дерев рішень є алгоритм дерев класифікації і регресії, що пропонує набір правил для дихотомічної класифікації сукупності вихідних даних. Даний метод звичайно застосовується для прогнозування того, які послідовності подій будуть мати заданий результат. На основі дерев рішень розроблені такі програмні продукти, як IDIS, C5.0 і SIPINA.

До експертних методів варто віднести і предметно-орієнтовані системи аналізу ситуацій і прогнозу, засновані на фіксованих математичних моделях, що відповідають тій або іншій теоретичній концепції. Роль експерта складається у виборі найбільш адекватної системи й інтерпретації отриманого алгоритму. Достоїнства і недоліки таких систем очевидні – гранична простота і доступність застосування і розплата вірогідністю і точністю за цю простоту. Прикладами програмних продуктів, що відповідають предметно-орієнтованим системам в області фінансів, є Wall Street Money, MetaStock, SuperCharts, Candlestick Forecaster.

У завершенні огляду експертних методів ІАД варто згадати методи візуалізації даних і результатів їхнього аналізу, що дозволяють наочно відобразити отримані висновки для створення у предметних експертів і/або керівників проектів єдиної картини ситуації. До програмних продуктів, що дозволяють формувати попередні звіти і візуалізувати результати, варто віднести системи Mineset і Impromptu (BI). Зокрема, система Mineset містить у собі такі інструменти, як ландшафтний візуалізатор, візуалізатори дисперсії, дерев, правил і свідчень.

Формувати складні нелінійні відображення засобами кольорової графіки дозволяє новий напрямок візуалізації результатів, заснований на ідеях фрактальної математики.

Висновки

Розглянуті методи та засоби аналізу даних складають основу створення спеціального математичного забезпечення СППР. Вибір того, чи іншого математичного засобу аналізу залежить від видів задач, що вирішуються в СППР, цілей їх функціонування та досвіду розробників. Наведений аналіз найбільш розповсюджених засобів ІАД має за мету зорієнтувати розробників МЗ СППР у складному просторі математичних методів аналізу даних.

Список літератури

1. Бадавшин Р.А., Ильясов Б.Г., Черняховская Л.Р. *Проблемы управления сложными динамическими объектами в критических ситуациях на основе знаний*. – М.: Машиностроение, 2003. – 240 с.
2. *Базы данных. Интеллектуальная обработка информации* / В.В. Корнеев и др. – М.: Нолидж, 2000. – 496 с.
3. Брандт З. *Анализ данных. Статистические и вычислительные методы для научных работников и инже-*

неров / Пер.с англ. – М.: Мир, 2003. – 686 с.

4. Гаврилова Т.А., Хорошевский В.Ф. *Базы знаний интеллектуальных систем*. – СПб.: Питер, 2001. – 384 с.

5. Герасимов Б.М., Дивизнюк М.М., Субач И.Ю. *Системы поддержки принятия решений: проектирование, применение, оценка эффективности*. – Севастополь: НИИ ВС, 2004. – 320 с.

6. Джарратано Дж., Райли Г. *Экспертные системы: принципы разработки и программирование*. – М.: Издательский дом "Вильямс", 2006. – 1152 с.

7. *Интеллектуальные системы поддержки принятия решений* / В.А. Геловани, А.А. Баилыков, В.Б. Бритков, Е.Д. Вязилов. – М.: Эдиториал УРСС, 2001. – 304 с.

8. Козлова М.Г. *Системы поддержки принятия решений в современной информатике* // Программы, системы, модели. – Симферополь. – 1996. – № 2. – С. 40-45.

Надійшла до редколегії 13.04.2007

Рецензент: д-р техн. наук, проф. В.С. Харченко, Національний аерокосмічний університет ім. М.С. Жуковського, Харків.