

УДК 681.3:519.2

В.М. Крищук, Р.С. Сікорський

Запорізький національний технічний університет, кафедра КВР

## ПРИНЦИПИ ПОБУДОВИ СИСТЕМИ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЕКОНОМІЧНОЇ ДІЯЛЬНОСТІ ПРОМИСЛОВОГО ПІДПРИЄМСТВА

© Крищук В.М., Сікорський Р.С., 2001

**Досліджені функціональні особливості існуючих систем аналізу та прогнозування. Описаний підхід до побудови системи аналізу та прогнозування економічної діяльності підприємства.**

**The functional singularities of existing systems of the analysis and prediction are researched. The approach to a building of a system of the analysis and prediction of economical activity of firm is described.**

### Вступ

Стрімкий розвиток інформаційних технологій, зокрема, прогрес у методах збирання, збереження та обробки інформації, дав змогу багатьом організаціям накопичувати величезні масиви даних, котрі необхідно аналізувати. Обсяг цих даних настільки великі, що можливостей експертів уже не вистачає – це призвело до виникнення попиту на системи автоматичного аналізу даних, котрий з кожним роком постійно зростає. Особливо актуальною проблема своєчасного та якісного аналізу і прогнозу є для економіки, бо вони дозволяють виявити несприятливі фактори й уникнути значних збитків. Існуючі системи, якими користуються закордонні фірми, у більшості своїй не відповідають стандартам, встановленим на українських підприємствах, та до того ж мають занадто високу вартість. Крім того, вони здебільшого не орієнтовані на використання у масштабі підприємства та не мають засобів інтеграції у корпоративну систему. У зв'язку з цим виникає потреба у розробці систем, що дозволяють опрацьовувати великі обсяги економічних даних, легко та зручно аналізувати їх, будувати на основі цих даних якісний прогноз, а також мають гнучкий інтерфейс та засоби інтеграції з іншими системами у корпоративній мережі.

### 1. Особливості архітектури систем аналізу та прогнозування

Інформація про економічну діяльність підприємства збирається, зберігається й оброблюється у різних підрозділах часто за допомогою різних засобів. Така ситуація не сприяє проведенню якісного аналізу та прогнозування, у першу чергу через утруднений доступ до необхідних даних, що зберігаються у різних місцях та форматах. Розкиданість даних змушує включати до складу систем аналізу та прогнозування власні засоби збереження інформації, створювати сховища даних чи розробляти спеціалізовані модулі доступу до даних. Найчастіше використовується доступ до баз даних через драйвери ODBC. Такий підхід зобов'язує користувачів зберігати необхідні драйвери на локальних машинах, що призводить до надмірного використання дискового простору та ускладнює настройку системи після інсталяції. Більшість сучасних систем аналізу та прогнозування підтримують технологію OLAP, бо вона дає змогу досить просто вибирати необхідні дані зі спеціально організованого сховища чи багатовимірної бази даних, легко маніпулювати ними та аналізувати з різних поглядів. Проте для поповнення та оновлення даних у сховищі цим

системам також доводиться використовувати драйвери ODBC, тобто проблема прозорого доступу до даних залишається невирішеною.

Під час збирання інформації у дані вносяться помилкові значення, що виникають через неуважність оператора, технічні збої та некоректне функціонування програм. Неправдоподібні значення можуть викликати значні труднощі при наступному аналізі чи зробити його зовсім неможливим. Уникнути появи помилкових даних практично неможливо через те, що неможливо контролювати саму появу помилок. Помилки у даних зумовлюють необхідність включення до системи модулів фільтрації та згладжування даних, що дає змогу провести попередню обробку інформації та значно підвищити якість аналізу та прогнозу.

Після вибірки необхідних даних та їх попередньої обробки система переходить безпосередньо до аналізу чи побудови прогнозу. Сучасні системи у більшості орієнтуються на використання технології OLAP, яка дає змогу досить легко провести аналіз даних на основі представлення їх у вигляді багатовимірного гіперкуба. Поряд із OLAP-системами існують системи, основані на використанні різних статистичних методів аналізу: кореляційного та регресійного аналізу, аналізу тренду та сезонності часових рядів та ін., а також системи, що використовують нейромережеві методи, зокрема, кластерний аналіз на основі мереж Кохонена.

Для побудови прогнозу система повинна мати модель процесу, що прогнозується. Побудова моделі економічного процесу є досить складною задачею через велику кількість параметрів та недостатню вивченість законів функціонування економічних систем. Тому для побудови моделей у системах аналізу та прогнозування використовуються підходи, що дозволяють зменшити складність моделювання: імітаційне моделювання, регресійні моделі та моделі на основі нейронних мереж. Використання імітаційних та регресійних моделей змушує вбудовувати в систему спеціалізовану мову для їх описання та створювати бібліотеки стандартних моделей. Такий підхід змушує користувача вивчати всі деталі модельованого процесу, а також зумовлює необхідність вивчення мови системи навіть у тому випадку, коли для розв'язання задачі користувачеві досить побудувати просту модель, описання якої займає кілька рядків на мові системи. При побудові моделі на основі нейронної мережі користувачеві досить задати кількість входів, виходів та вибрати вид мережі. Під час навчання нейронна мережа сама виявить залежності між параметрами та параметри, що найбільше впливають на модельований процес.

До особливостей архітектури систем аналізу та прогнозування слід віднести також орієнтованість на використання якоїсь однієї парадигми, наприклад, OLAP чи нейронних мереж. Крім того, більшість існуючих систем не вирішують проблеми інтеграції окремих частин у корпоративну систему, а також не враховують проблему продуктивності роботи такої корпоративної системи.

Для побудови концепції створення системи аналізу та прогнозування проведемо аналіз існуючих систем, розглянемо їх переваги та недоліки.

## 2. Аналіз функціональних особливостей існуючих систем

Сьогодні на ринку систем аналізу та прогнозування присутні різні пропозиції закордонних та вітчизняних фірм, які покликані задовольнити все зростаючий попит на якісні інструменти аналізу та прогнозування.

## 2.1. Система “Deductor”

Deductor – це пакет програм, призначений для швидкого та ефективного аналізу інформації [1]. У ньому зосереджені сучасні методи збирання, маніпулювання та візуалізації даних. З використанням Deductor-а стають доступними моделювання, прогнозування, кластеризація, пошук закономірностей та багато інших технологій виводу даних (Data Mining).

До складу пакета входить настільний OLAP модуль (Cube Analyzer), модуль аналізу на основі дерев рішень (Tree Analyzer), модуль аналізу на основі карт Кохонена (SOMap Analyzer) та модуль прогнозування і моделювання на основі багатопарових нейронних та RBF мереж (Neural Analyzer).

Пакет Deductor поширюється у двох варіантах:

- Deductor Lite – спрощена безкоштовна версія (оброблює максимум 150 записів);
- Deductor Professional – повна комерційна версія.

До Deductor-а включено набір програм, що забезпечують досить глибокий і різнобічний аналіз інформації. Перенесення результатів аналізу між програмами дозволяє використовувати переваги кожної методики без необхідності жертвувати якістю обробки. Наявність безкоштовної версії дозволяє використовувати Deductor на початкових стадіях впровадження, не виділяючи для цього ніяких додаткових коштів, в той же час версія Lite є швидше демонстрацією можливостей пакета, ніж завершеною програмою.

Суттєвим недоліком системи Deductor є те, що пакет не розрахований на використання у локальних мережах, не має вбудованого модуля зв'язку з Web. Крім того, оскільки пакет цілком написаний на Delphi, його можна використовувати лише на машинах, що працюють під управлінням системи MS Windows.

## 2.2. Система “BrainMaker”

BrainMaker – популярний нейромережевий пакет, розроблений компанією California Scientific Software [2]. Пакет складається з кількох модулів, серед яких основними є сам BrainMaker, а також модуль попередньої підготовки й аналізу даних NetMaker. Крім того, до пакета входять засоби визначення корельованості даних та обчислення часового лага між різними параметрами, а також засоби згладжування вхідних даних.

Багато операцій з підготовки даних зручніше виконувати не у NetMaker, а в інших пакетах, наприклад, Excel. Більш того, для щоденного використання нейронної мережі, не кажучи вже про реальний час, потрібні спеціальні програми, які вибирають дані з потоку інформації, що надходить, перетворюють їх до потрібної форми й передають на вхід нейронної мережі. Інакше кажучи, тут наявний тісний зв'язок зі сховищами даних.

Таким чином, BrainMaker є системою, цілковито орієнтованою на використання нейронної мережі. У пакеті не реалізовано підтримки технології OLAP, немає засобів роботи у локальній мережі, що є суттєвим недоліком даної системи.

## 2.3. Інформаційно-аналітична система «Оценка и Выбор»

Система розроблена в компанії ГТК, написана на мові FoxPro v3.0. Інформаційно-аналітична система (ІАС) “ОЦЕНКА и ВЫБОР” належить до класу систем підтримки прийняття рішень. Основна функція системи – багатокритеріальна оцінка, порівняння та вибір кращих з можливих варіантів будь-яких рішень.

IAC, по суті, є програмною “оболонкою”, яку користувач може настроїти на будь-яку предметну область, де для аналізованих об’єктів (для варіантів рішень) треба:

- дати оцінку відповідності висунутим вимогам за багатьма критеріями;
- прийняти рішення альтернативного вибору (наприклад “вибрати краще з...” чи “прийняти-відкинути”), або
- прийняти рішення про розподіл ресурсів між групою об’єктів, враховуючи їх поточну пріоритетність.

Функціональні можливості IAC дають змогу використовувати її в широкому діапазоні практичних ситуацій:

- від випадку підготовки рішення лише самим керівником при відсутності об’єктивних даних про характеристики порівнюваних варіантів;
- до випадку підготовки рішення групою керівників та спеціалістів з обробкою великих об’ємів кількісної та якісної інформації про об’єкти, що аналізуються.

Можлива багатокритеріальна оцінка єдиного об’єкта (рішення).

Система не вимагає використання мов програмування та має вбудовані засоби:

- створення та ведення баз даних;
- ведення бази методик (моделей) багатокритеріального аналізу;
- видобутку та формалізації думок експертів з контролем послідовності їх суджень;
- узагальнення та оцінки узгодженості думок групи до 10 експертів;
- експорт та імпорт даних у форматі Excel;
- систему допомоги на російській мові.

Необхідно відзначити такі недоліки, як орієнтованість системи на використання конкретної СУБД (FoxPro), відсутність підтримки OLAP, неможливість роботи з системою у локальній мережі.

#### 2.4. Система “Black Oracle Pro”

Black Oracle Pro призначена для прогнозування значень часових рядів [3]. У пакеті використовується нейронна мережа типу Back Propagation з одним внутрішнім шаром. Навчання можна виконувати трьома алгоритмами (градієнтним, модифікованим градієнтним, модифікованим генетичним) за двома критеріями (точність, тенденція). Окрім цього, Black Oracle Pro має зручний інтерфейс, що дає змогу візуально зберігати результати експериментів та стежити за навчанням за допомогою як графічного, так і числового моніторингу процесу. Система створена з використанням COM технології, що дозволяє “викидати” звіти у Excel одним натисненням кнопки.

Алгоритм тасування точок, що використовується у системі, автоматично виключає можливість перенавчання, й, незважаючи на порівняно погану якість навчання, дає на порядок кращі результати при прогнозуванні, ніж алгоритм, що спирається на мінімізацію сумарної помилки.

Система Black Oracle Pro, як і пакет BrainMaker цілковито орієнтована на використання технології нейронних мереж і має такі ж недоліки, а саме: відсутність підтримки OLAP, неможливість роботи через локальну мережу, відсутність підтримки інших методів аналізу та прогнозування.

## 2.5. Система “Прогноз”

Система “Прогноз” розроблена у пермській філії холдингу “Стерлінг груп” [4]. Комплекс представляє собою пакет програм, спеціалізованих для розв’язування задач загального аналізу, моделювання, прогнозування та планування.

В основу програмного комплексу покладена технологія оперативної аналітичної обробки реляційних даних (ROLAP). Комплекс орієнтований на багатовимірне представлення даних, але допускається пряма обробка традиційних реляційних таблиць чи логічне перетворення таблиць у багатовимірні об’єкти. Підтримуються шлюзи до SQL-орієнтованих СУБД (Oracle, Informix, MS SQL Server та ін.), СУБД класу персональних (Access, FoxPro, Paradox), електронних таблиць (Excel) та файлів.

Блок моделювання надає користувачеві набір візуальних та програмних засобів для побудови моделей бізнес-процесів. Основою всього блока є макромова, орієнтована на обробку багатовимірних матриць та векторів. Комплексний підхід до побудови моделей бізнес-процесів забезпечується використанням імітаційних, цільових, оптимізаційних та статистичних методів моделювання та прогнозування.

Окрім табличного та графічного представлення, дані та результати розрахунків можуть бути представлені на електронних картах територій, що у ряді випадків значно спрощує їх візуальний аналіз.

Програмні засоби комплексу передбачають механізми захисту об’єктів та даних від несанкціонованого доступу. Права кожного користувача визначаються ступенем його участі в робочому процесі.

Web-розширення надає користувачам можливість незалежно від їхнього географічного положення переглядати звіти та модифікувати вміст інформаційного сховища.

Комплекс підтримує колективну роботу з програмами у рамках архітектури “клієнт-сервер” з використання будь-якої SQL-сумісної СУБД (Oracle, Informix, MS-SQL Server та ін.). До операційної системи та апаратної платформи сервера не висувається ніяких спеціальних вимог, таким чином, можливе використання вказаних СУБД як на платформі Windows NT, так і UNIX.

Клієнтська компонента комплексу виконана на Inprise Delphi, як клієнтська операційна система можуть використовуватися Windows 95/98 чи Windows NT 4.0 та вище.

До недоліків системи “Прогноз” належить наявність досить складної вбудованої мови, спеціалізований формат баз даних системи. До того ж, система зберігає всі дані у власних базах даних, завантажуючи інформацію із зовнішніх джерел кожного разу, коли необхідно оновити дані.

## 3. Принципи побудови системи аналізу та прогнозування

На основі проведеного вище аналізу запропоновано принципи побудови системи аналізу та прогнозування:

1. Забезпечення прозорого доступу до даних та можливості маніпулювати ними. Для розв’язання цієї задачі пропонується використовувати сервер запитів, що підтримує два представлення даних: реляційне та багатовимірне (OLAP) з можливістю логічного перетворення плоских таблиць у куби даних. Для підвищення гнучкості та забезпечення можливості інтеграції з іншими системами планується розробити сервер запитів на основі технології CORBA [5], що дає

змогу взаємодіяти з клієнтами, що написані на різних мовах програмування та виконуються на різних платформах.

2. Забезпечення фільтрації та виправлення даних. До складу системи має бути включений модуль, що містить бібліотеку різних методів фільтрації та згладжування даних. Використання об'єктно-орієнтованого підходу дає змогу легко розширювати бібліотеку методів без необхідності переробки інших модулів системи.
3. Забезпечення аналізу кількома методами. Для цього у систему має бути включена бібліотека методів аналізу. Система орієнтується на використання технології OLAP та нейронних мереж (карти Кохонена), але, завдяки використанню технології CORBA, дозволяє легко підмикати до системи модулі, що реалізують інші технології аналізу даних.
4. Забезпечення побудови моделі процесу та прогнозу майбутнього стану. Система орієнтована на технологію інформаційного моделювання з використанням нейронних мереж. При цьому з'являється можливість не просто побудувати апроксимацію функціональної залежності між даними, а, завдяки спроможності нейронних мереж до навчання, виділити закон взаємодії з наявних даних і передбачити майбутній стан системи. Навчену мережу можливо зберегти для подальшого використання без повторного навчання.
5. Забезпечення зв'язку з іншими програмами. Використання технології CORBA дозволяє об'єднувати існуючі системи у корпоративну систему, при цьому досить легко об'єднувати системи, що працюють на різних апаратних платформах під управлінням різних операційних систем.
6. Забезпечення масштабування та гнучкості системи. Використання об'єктноорієнтованого підходу і технології CORBA дозволяє забезпечити гнучкість системи та досить великі можливості для нарощування, без необхідності перебудови всієї системи.
7. Використання технології розподілених обчислень дає змогу значно підвищити продуктивність системи за рахунок виконання окремих частин системи на різних машинах локальної мережі, і в той же час не втрачати інтегрованості системи.
8. Системи аналізу та прогнозування зазвичай працюють з величезними об'ємами даних, що створює значне навантаження на обчислювальну систему. Використання технології паралельних обчислень, наприклад, об'єднання машин локальної мережі у кластер на основі бібліотеки PVM (Parallel Virtual Machine) [6], дає змогу розподілити навантаження між окремими машинами, а також значно підвищити продуктивність системи в цілому. При цьому кластер є динамічним, тобто можна включати та вилучати машини з кластера під час роботи, а також контролювати роботу кластера з однієї з машин.
9. Забезпечення доступу до системи. Використовуючи технологію CORBA та мову програмування Java, можемо створювати аплети та сервлети, що вбудовуються у HTML сторінку та забезпечують доступ до системи не лише з робочого місця, а з будь-якої машини локальної мережі чи Internet за допомогою Web-браузера.

На рисунку подана укрупнена структура розподіленої паралельної системи аналізу та прогнозування.



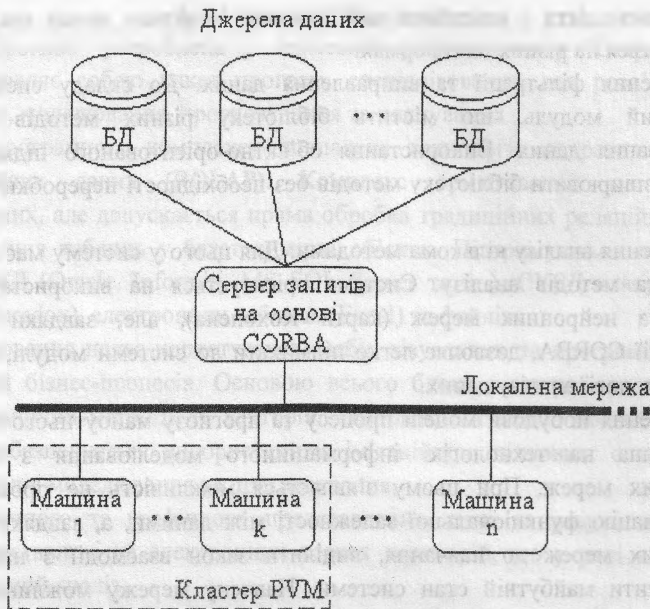


Схема розподіленої паралельної системи аналізу та прогнозування

### Висновки

Запропоновано принципи створення системи аналізу та прогнозування, що вирішує проблеми побудови різнобічного та якісного аналізу і прогнозу, а також дає змогу об'єднати окремі модулі у корпоративну систему. Використання технологій, розподілених та паралельних обчислень, дозволяє вирівнювати навантаження на окремі частини корпоративної системи та забезпечувати безперебійний і безпечний доступ до системи з будь-якої машини локальної мережі чи Internet за допомогою Web-браузера.

1. *Deductor* – описание аналитического пакета. (<http://www.basegroup.ru/deductor/deductor.htm>).
2. Блинов Сергей "BrainMaker - прогнозирование на финансовых рынках" // *Открытые системы*. – 1998. – №4.
3. *Применение нейронных сетей для прогнозирования показателей экономики переходного периода* (<http://www.znanie.kiev.ua/blackoraclepro.html>).
4. Аналитический комплекс "ПРОГНОЗ" (<http://www.prognoz.ru/products/>).
5. Дирк Слама, Джейсон Гарбис, Перри Рассел *Корпоративные системы на основе CORBA*. – М.: Издательский дом "Вильямс", 2000.
6. *Libraries for distributed computing* ([http://www.epm.ornl.gov/pvm/pvm\\_home.html](http://www.epm.ornl.gov/pvm/pvm_home.html))