

СХОВИЩЕ ДАНИХ ЯК МОДЕЛЬ СИТУАЦІЙНОГО АНАЛІЗУ ДІЯЛЬНОСТІ ІНФОРМАЦІЙНОЇ СИСТЕМИ (ПРОЦЕДУРИ І АЛГОРИТМИ ФОРМУВАННЯ ТА ВИКОРИСТАННЯ)

Abstract: The problem of operational loading of modeling process and analysis of system providing and using history of the past activity of the information system (IS) as specialized Data Warehouse (DW) discussed in the paper. The purpose of the analysis is actualization of knowledge about system on the base of exposure discovery of the new functional dependences between key parameters – indices (their groups) of the activity.

Key words: information system, IS, Data Warehouse, time series, operational loading.

Анотація: В роботі розглядається проблема операційного наповнення процесів моделювання й аналізу системи ведення та використання історії минулої діяльності інформаційної системи (ІС) у вигляді спеціалізованого Сховища фактографічних даних (СхД). Метою аналізу є осучаснення знань про систему за рахунок виявлення/відкриття нових функціональних залежностей між ключовими параметрами-показниками (групами параметрів) діяльності.

Ключові слова: інформаційна система, ІС, Сховище даних, часові ряди, операційне наповнення.

Аннотация: В работе рассматривается проблема операционного наполнения процессов моделирования и анализа системы ведения и использования истории прошлой деятельности информационной системы в виде специализированного Хранилища фактографических данных. Целью анализа является актуализация знаний о системе на базе выявления/открытия новых функциональных зависимостей между ключевыми параметрами-показателями (группами параметров) деятельности.

Ключевые слова: информационная система, Хранилище данных, временные ряды, операционное наполнение.

Майбутнє вже не те, що було раніше
П. Валері

1. Вступ

Слова епіграфа, на нашу думку, можна розглядати як передбачення недостатньої життєстійкості традиційних інформаційних систем управління (ІСУ), яке демонструє, що ці системи виявилися не зовсім готовими підтримувати процеси аналізу поточних результатів своєї діяльності (прийняті рішення та наслідки їх виконання) з метою оновлення знань про себе, обґрунтування адекватності рекомендованих системою рішень реальним потребам об'єкта управління, необхідності адаптації системи до змінених умов існування [1–3].

Майже піввіковий досвід багатопланового використання традиційних ІСУ показує, що в дійсності більшість адаптаційних процесів час від часу виконується, але тільки при виникненні проблем, які заважають продуктивній діяльності системи, хоча головне завдання цих процесів є попередження та запобігання виникненню таких ситуацій з можливими варіантами перебудови системи або її реорганізації (реінжинірингу). Саме тому сучасні технології адаптації ІСУ повинні мати більш складний, людино-машинний характер, який визначається тим, що інтерактивні (людино-машинні) моделі ситуаційного аналізу повинні бути вбудовані безпосередньо у процеси функціонування системи на протязі всього життєвого циклу “об'єднання” Об'єкт&ІСУ, підтримуючи та забезпечуючи необхідний і достатній рівень його “розумної”, адекватної поведінки в умовах існування, які динамічно змінюються.

Основні системотвірні засади моделей абстрагування від проблем “сьогодення” з метою виявлення/відкриття тенденцій впливу зовнішніх і внутрішніх факторів на формування задач “дня завтрашнього” сформулюємо, виходячи з трактування “постулатів бачення майбутнього” ІСУ [6]:

1. Майбутнє базується на історії існуючого – тому необхідно зосереджуватися на аналізі динаміки отриманих результатів діяльності.

2. “Важливо не бути абсолютно правим” – слід пам’ятати про відносність правоти.

3. Для досягнення бажаного рівня життєздатності системи не слід використовувати лише закладені Розробником (регламентовані ним) алгоритми та процедури, а необхідно шукати, відкривати й використовувати нові, більш адекватні поточним умовам.

У даній роботі автори розглядають операційну наповненість моделі “розумної” поведінки системи на базі використання апарата часових рядів [8].

2. Постановка проблеми

Отже, методологія “розумного” характеру функціонування ІСУ як системи ситуаційного управління (ССУ) повинна базуватися на засадах, що аналіз минулої діяльності спирається на інформаційно-логічні моделі, які більш ефективно представляються в термінах “впорядкованих послідовностей (потоків) минулих значень ключових параметрів-показників (ПП)”, а ніж в термінах “поточних їх значень (як оперативних даних виконання регламентованих функцій та задач)”. Спрямованість на потокову структуру моделей вхідних даних дозволяє аналітику в процесах одночасного аналізу тематичних сукупностей ПП різних підсистем (або “інформаційних згорток” [1]) природним чином обминати внутрішні функціонально-організаційні кордони системи, а отже аналізувати стан системи як єдиного цілого з метою виявлення нових (невідомих під час створення системи) залежностей та закономірностей у поведінці ПП (груп ПП).

Наведена вище оцінка функціональної неповноти процесів адаптації традиційних ІСУ стосовно ситуативного забезпечення та підтримки власного рівня життєздатності й життєстійкості ініціювала подальший розвиток додаткових можливостей ССУ за рахунок системотехнічних механізмів та інструментів її “континууму розумності” – організованої сукупності операційних засобів, орієнтованих на формування, ведення та ситуативний аналіз результатів фактичної діяльності системи, зафіксованих у Сховищі даних (СхД) [4] як історії діяльності.

В інтеграції з технологіями Олар (онлайніві аналітичні процеси) та Інтелектуального аналізу даних, або Data Mining (=“розкопки даних”) використання механізмів СхД розширює можливості інтерактивних процесів взаємодії аналітик-система засобами інформаційної підтримки більш повного та поглибленого аналізу стану за рахунок співставлення “поведінки” сукупностей ПП з метою пошуку та обґрунтування на цій платформі шляхів реорганізації та подальшого вдосконалення базової ІСУ [5, 8].

Функціональність нових складових визначається так (рис. 1):

СхД – консолідовані (узгоджені в часі) значення послідовності значень (часові ряди) ПП та зовнішніх факторів впливу;

Олар – знаходження тенденцій та закономірностей у поведінці ПП (груп ПП);

ІАД – відкриття нових знань або невідомих раніше (на момент розробки системи) залежностей між ПП.

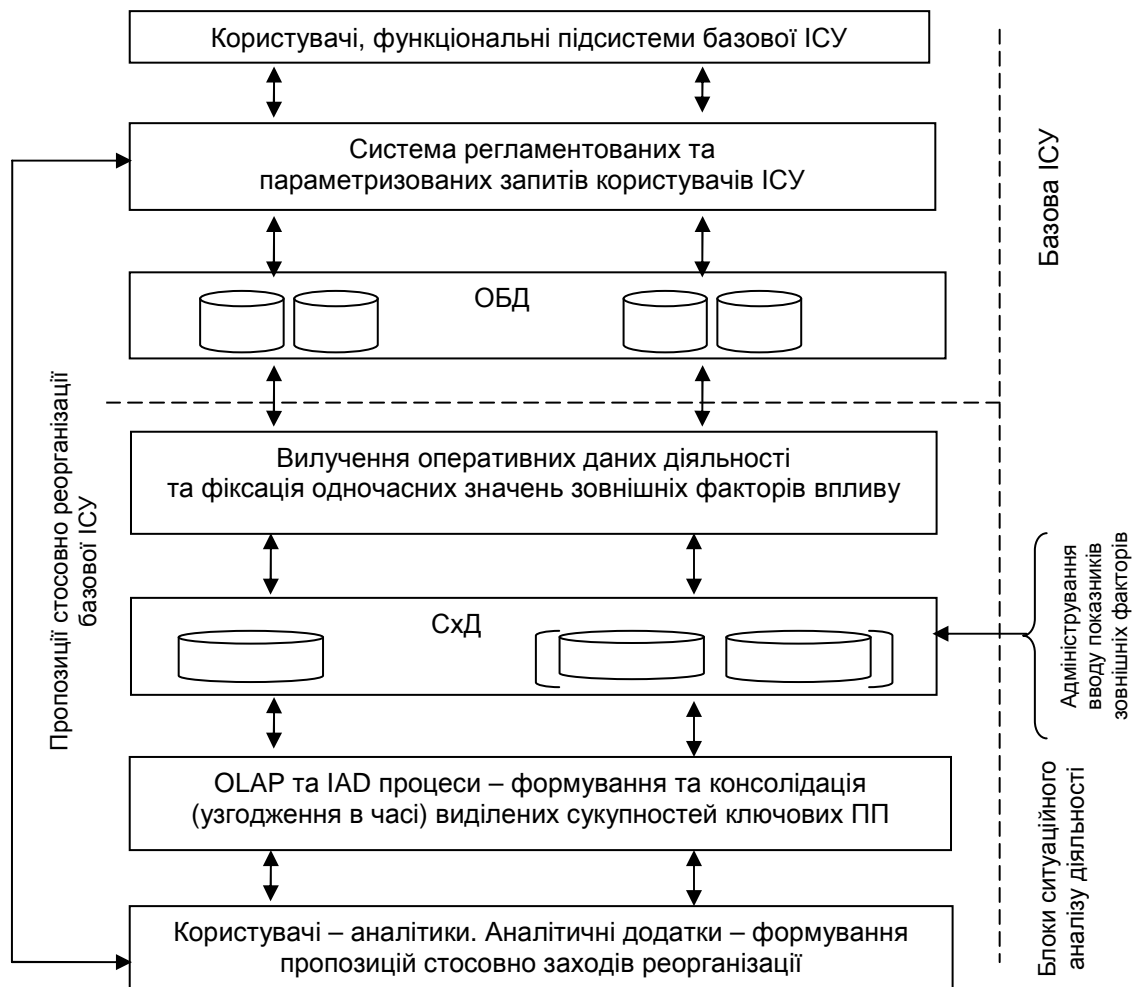


Рис. 1. Інтеграційна схема взаємодії ССУ та базової ІСУ

3. Операційне наповнення “ситуаційних роздумів”

3.1. Інформаційна платформа “роздумів”

1. Вибір та оцінка стратегічних альтернатив самовдосконалення та саморозвитку системи базується на дослідженні потоків ПП, зафіксованих у СхД як інформаційної бази відображення подій, явищ, що фактично мали місце. Ці дослідження мають на меті не знаходження “єдино вірної” стратегії, а формування з урахуванням результатів фактичної діяльності її актуальних (осучаснених) засад.

2. Кожний запис СхД – це t_j часовий зріз множини ключових ПП або сукупність

$\{ПП_{ij}\} = \{ПП_i(t_j) \mid i = \bar{1}, n, j = \bar{1}, m\}$. $ПП_{ij}$ – у більш широкому розумінні, певна функціональна залежність від первинних атрибутів (x_i, y_i, z_i) , або $ПП_{ij} = f_i(x_i(t_j), y_i(t_j), z_i(t_j), \dots)$.

Формалізовану операційну наповненість тріади (СхД&Олар&ІАД) представимо у вигляді [10]

$$ССУ \equiv (\text{базова ІСУ} \cup \text{додатки}((\text{СхД} \times \text{БЗн}), \text{ОП}_{ІАД})),$$

де $(\text{СхД} \times \text{БЗн})$ – інформаційна платформа аналізу;

$\text{ОП}_{ІАД}$ – сукупність процедур-операцій ІАД у складі $\{\text{ОП}_{ДД}, \text{ОП}_{ВЗн}, \text{ОП}_{РЕПН}\}$;

$OP_{\text{ДД}}$ – “добування даних”, їх часове та змістовне узгодження;

$OP_{\text{ВЗн}}$ – інтелектуальне опрацювання вилучених даних, відкриття нових знань та правил виводу;

$OP_{\text{РЕИ}}$ – формування та планування заходів реорганізації (реінжинірингу), верифікація та консолідація (взаємна узгодженість) узагальнень і правил виводу.

3. Для визначення виду (з певною мірою вірогідності $\mu(t)$) можливої функціональної залежності між довільною парою (трійкою, ...) $PPP_k(t) = \mu(t)f(PPP_1(t))$ аналізуються на “схожість поведінки” синхронізовані в часі (з однаковим періодом дискретизації) ряди $PPP_k(t)$ та $PPP_1(t)$ або їх відповідні перші чи другі різниці. Для оцінки “схожості” використовується багатокрокова процедура, на кожному етапі якої виконується оцінка певних ознак “поведінки” ПП як ЧР [9] *тренд* або стійкої закономірності змін певної ознаки; число *поворотних точок* або *довжина проміжків монотонності* та ін.

Наприклад, порівняльний аналіз пари (PPP_1, PPP_k) , який виконується на часовому проміжку $[t_0, T]$ полягає у співставленні значень рядів (PPP_1, PPP_k) , які задовольняють умовам $(t_0 \leq t_j^1 \leq T)$ та $(t_0 \leq t_j^k \leq T)$ відповідно. Мета аналізу – “відкриття” невідомої раніше залежності між цими ПП. Тобто “відкриття”, яке може покласти початок процесам “народження” нової організаційної структури за рахунок перерозподілу або заміні діючих функцій.

Вважаючи, що PPP_{1r} – незалежна змінна, а PPP_1 – залежна і у випадку різних значень періодів дискретизації аналіз на схожість поведінки необхідно провести для відкоригованого (з вирівненим періодом дискретизації) ряду $PPP_1(t)$, який отримується за допомогою інтерполяційного многочлена ПП₁ операціями обчислення нових значень $PPP_1(t_i^1)$ за відмітками часу (t_j^k) .

4. Висновки

Перехід від регламентованого (в рамках традиційних ІСУ) “аналізу результатів діяльності лише в розрізі поточних значень ОБД” до “системного (ситуаційного) аналізу минулої діяльності об’єкта інформатизації в цілому”, тобто до дослідження зв’язків між ключовими ПП різних підсистем на спільних проміжках часу, зафіксованих в СхД, та пошуку з використанням операцій $OP_{\text{ДД}}$ нових залежностей між ними, визначає методологічні засади інформаційної підтримки та стимулювання “роздумів” (інтелектуального аналізу) аналітика у процесах обґрунтування можливої модифікації базової ІСУ і її бази знань і формування пропозицій стосовно планів та заходів їх реорганізації.

Переваги людино-машинної інтелектуалізації не обмежуються наданням аналітикам наочної і актуальної картини різних аспектів діяльності об’єкта інформатизації на основі виявленої динаміки ПП. До переваг слід віднести нові операційні можливості стосовно пошуку прихованих в СхД об’єктивно існуючих закономірностей поведінки ПП (груп ПП) та виявлення вірогідних тенденцій їх розвитку та становлення.

Встановлюючи порогові значення ключових ПП та фіксуючи тенденції погіршення діяльності системи в цілому до того, як вона стане проблемною, аналітикам вдається обґрунтувати запобіжні заходи аварійного розвитку подій, які передбачають перерозподіл ресурсів (задач, функцій).

Серед відкритої множини типових ІТ у складі базових ІСУ, які потребують такої інтелектуалізації, а, отже, і запропонованих в роботі процедур аналізу минулої діяльності, можна назвати [12]:

- фінанси, банки, банківська діяльність – внутрішня звітність, бюджет, активи, пасиви, база клієнтів, робота філіалів;
- торговельні організації – продаж товарів у розрізі номенклатури, тенденції розвитку динаміки продаж;
- промислові підприємства – контроль відправки товарів, податкова звітність, аналітика відношень з клієнтами та конкурентами, аналітика власного персоналу;
- державні органи – аналіз надвеликих СхД з інформацією про населення, житлово-комунальні послуги, кадастри землі та підприємств;
- та ін.

СПИСОК ЛІТЕРАТУРИ

1. Морозов А.О. Ситуаційні центри – основа керування організаційними системами // Математические машины и системы. – 1997. – № 1. – С. 1–10.
2. Тарасов В.Б. Тектология А.А. Богданова и неоклассические теории организации // Программные продукты и системы. – 1998. – № 2. – С. 43–47.
3. Теслер Г.С. Новая кибернетика. – Киев: Изд-во Логос, 2004. – 396 с.
4. Асельдеров З.М. “Континуум розумності” інформаційних систем / З.М. Асельдеров, А.О. Морозов, В.І. В’юн // Матеріали конференції “Штучний інтелект”. – Донецьк, Кацівелі, 2004. – Вересень.
5. Банасевич А., Гурленко А. Хранилище данных и OLAP // Корпоративные системы. – 2002. – № 4. – С. 44–51.
6. Нэбитт Джон Установка ума. – М., 2003. – 145 с. (Препринт / Изд-во МГУ).
7. Морозов А.О. Інтелектуалізація інформаційних систем: орієнтація на формування знань у процесах аналізу “інформаційних згорток” / А.О. Морозов, В.І. В’юн, Г.Є. Кузьменко // Математичні машини і системи. – 2005. – № 2. – С. 140–146.
8. Бідюк П.І. Часові ряди: моделювання та прогнозування / П.І. Бідюк, О.І. Савенков, І.В. Баклан. – Київ: Екмо, 2003. – 144 с.
9. Балабанов А.С. Проблема вывода знаний о структуре зависимостей между переменными из данных больших объемов в условиях помех // Проблемы программирования. – 2000. – № 1–2. – С. 527–535.
10. В’юн В.І. “Нова кібернетика” та сучасні інформаційні системи управління // Математичні машини і системи. – 2006. – № 4. – С. 36–41.
11. Берштейн Л.С. Интеллектуальные системы / Л.С. Берштейн, П.М. Ильягуев, В.Б. Мелехин. – Махачкала: Дагкнигоиздат, 1996. – 67 с.
12. Андрусенко Т. Современный анализ данных: общие вопросы Business Intelligence // Корпоративные системы. – 2005. – № 2. – С. 59–63.

Стаття надійшла до редакції 19.06.2007