

ІНТЕЛЕКТУАЛІЗАЦІЯ ІНФОРМАЦІЙНИХ СИСТЕМ: ОРІЄНТАЦІЯ НА ФОРМУВАННЯ ЗНАНЬ В ПРОЦЕСАХ АНАЛІЗУ “ІНФОРМАЦІЙНИХ ЗГОРТОК”

Abstract: In the paper one of the aspects of Intellectualization connected with ability of the system to analyze the past activity in the regimes of the interactivity “user-system” to define causes and tendencies of the development of the potential threats of the vital activity, to develop offers on partial or full reorganization (reengineering). Information support of these processes is provided by independent on profile activity operations and procedures of forming “history of activity in the form of Data Storage (DS) of information rollings. On the bases of the analysis DS the possible schemes of the inductive conclusion, forming new functional dependences between parameters-indices are proposed as necessary stage of preparing and planning of reengineering.

Key words: intellectualization, interactivity “user-system”, reengineering, Data Storage, information system, information rollings.

Анотація: В роботі розглядається один з аспектів інтелектуалізації, пов'язаний із здатністю системи аналізувати минулу діяльність у режимах інтерактивної взаємодії “користувач-система”, виявляти причини та тенденції розвитку потенційних загроз життєздатності системи, розробляти заходи щодо часткової або повної її реорганізації (реінжиніринг). Інформаційна підтримка цих процесів забезпечується незалежними від профільної діяльності операціями і процедурами формування історії діяльності системи у вигляді Сховища даних (СхД) інформаційних згорток. На базі аналізу записів СхД пропонуються можливі схеми індуктивного узагальнення та формування нових функціональних залежностей між параметрами-показниками як необхідних етапів підготовки та планування реінжинірингу.

Ключові слова: інтелектуалізація, інформаційна система, інтерактивна взаємодія, реінжиніринг, Сховище даних, інформаційна згортка.

Аннотация: В работе рассматривается один из аспектов интеллектуализации, связанный со способностью системы анализировать прошлую деятельность в режимах интерактивного взаимодействия “пользователь-система”, определять причины и тенденции развития потенциальных угроз жизнеспособности системы, разрабатывать предложения по частичной или полной ее реорганизации (реинжиниринг). Информационная поддержка этих процессов обеспечивается независимыми от профильной деятельности операциями и процедурами формирования истории деятельности в виде Хранилища данных (ХрД) информационных сверток. На основе анализа записей ХрД предлагаются возможные схемы индуктивного обобщения и формирования новых функциональных зависимостей между параметрами-показателями как необходимых этапов подготовки и планирования реинжиниринга.

Ключевые слова: интеллектуализация, информационная система, интерактивное взаимодействие, Хранилище данных, реинжиниринг, информационная свертка.

1. Вступ

Однією з ознак інтелектуальності інформаційної системи (ІС) є її здатність адаптуватися до змін як внутрішніх, так і зовнішніх умов та вимог до результатів діяльності. У роботах [1–3] були окреслені системотвірні механізми та інструментарій відповідного, інтелектуалізуючого систему оснащення ІС. До них відносяться: Сховище даних (СхД) або просторово-часова база впорядкованої сукупності “інформаційних згорток” [4], що характеризують стан системи у вигляді проблемно-орієнтованих груп параметрів-показників, які тематично зв'язані, Олар (онлайніві аналітичні) та ІАД (інтелектуального аналізу даних) – технології.

У даній роботі ми зупинимось на методичних та технологічних засадах формування нових знань та вмінь системи (складових її “континууму розумності” [3]) з використанням схем інтерактивного ситуативного аналізу та індуктивного виводу (узагальнення) на платформі СхД.

2. Постановка проблеми

Необхідність у новій методології та інструментах обумовлена наявністю ряду важливих, але нерозв'язаних на засадах традиційних підходів до організації інформаційних систем (ІС) з

використанням комп'ютерної підтримки діяльності задач [5]. Перш за все, це задачі самоорганізації та адаптації системи до швидкоплинних змін в умовах існування, які визначають нові вимоги до результатів діяльності, нові відносно часу формування та впровадження ІС.

У даному контексті проблемою є надання системі здатності до людино-машинної самоорганізації й адаптації, маючи на увазі включення в її функціонально-організаційну структуру "розумних" механізмів та інструментів, які в інтерактивних процесах ситуативного аналізу діяльності спираються на засади, що допускають можливість часткової або повної реорганізації системи в порівнянні "з першими днями" її експлуатації. Надання ІС таких можливостей забезпечує далекоглядні її переваги, пов'язані саме з наявністю засобів самонавчання, самоорганізації та реорганізації (реінжинірингу).

Процес накопичення знань протягом багатьох століть, включаючи останні 50 років, коли почали говорити про накопичення знань комп'ютерами, підтримувався з використанням схем дедуктивного аналізу та умовиводу, що базувалися на логічно зв'язаній послідовності операцій, які, у свою чергу, визначались вже здобутими знаннями: скінченною множиною сформованих априорі та виведених у процесі розробки системи аксіом аналізу та правил виводу. Оскільки логічні початки та наслідки дедуктивних систем – це множина лише істинних тверджень, то, по суті, такі схеми прийняття рішень *були логічно неповними*. Більш того, процес розширення системи додатковою кількістю нових аксіом, залишаючись на базі логіки дедуктивного виводу, не збільшує її логічної повноти, оскільки потужність множини правдоподібних (можливих) умовиводів від цього розширення аксіом не зменшується. Отже збільшення діапазону вірогідності рішень за рахунок правдоподібних умовиводів можливе лише на платформі використання *систем індуктивного узагальнення*. Виходячи з того, що "істина" одна, а "не зовсім істин" багато, Розробник ІС при формуванні орієнтованого на аналіз діяльності СхД змушений вибирати певну базову модель підтримки та забезпечення правдоподібних умовиводів. Задача, яка при цьому розв'язується, це – "вибрати адекватну ПрО модель, що забезпечує необхідну вірогідність виводу". Найбільш широкий огляд існуючих моделей приводиться в [6].

Отже, процес аналізу результатів діяльності нетрадиційної ІС пропонується розглядати як паралельний процесу функціонування системи процес неперервного, інтерактивного, ситуативного аналізу результатів та проблемних ситуацій минулої діяльності, тенденцій їх розвитку, формування та прийняття на цій основі запобіжних заходів стосовно адаптації та самоорганізації системи в цілому.

Основними етапами цього процесу є:

- безпосереднє сприйняття проблемних ситуацій (подій, явищ);
- співставлення фактографічного опису їх характеристичних показників з існуючими знаннями про предметну область. Цей етап може розглядатися як етап "розуміння" проблемної ситуації, що склалася. При цьому можуть бути сформовані декілька уточнюючих описів (тобто, має місце певна неоднозначність розуміння), можуть пропонуватися варіанти змін та доповнень до функціонально-організаційної структури, а також пропозиції стосовно розширення бази знань;
- планування відповідних операцій реінжинірингу, аналіз та тестування на базі фактографічних записів СхД можливих наслідків;

– вибір цілеспрямованої послідовності операцій, яка найкраще узгоджується з метою реорганізації і забезпечує достатній рівень життєздатності ІС.

Відмітимо, що сьогоденне успішне застосування ІС тісно пов'язане з задачами, в яких чітко визначені й формалізовані критерії оптимізації та обмежень впливу оточуючого середовища. Але разом з цим існує досить широкий пласт проблем-задач, в яких не завжди вдається апіорі перерахувати всі як внутрішні, так і зовнішні фактори впливу, тому, що цьому заважає множина об'єктивно протирічних критеріїв, які повинні бути врахованими і які в той же час не можуть одночасно задовольнятися.

До розширення функціонального наповнення ІС, яке повинно скласти операційну базу саморозвитку та проблемної (об'єктної) орієнтації методів аналізу життєздатності системи в цілому, повинні належати механізми і інструментарій, які реалізують:

– ідентифікацію проблемної ситуації (ПрС) або початкові кроки її розуміння (визначення причин та факту виникнення, тенденцій розвитку, формування додаткових задач, які потребують розв'язку);

– використання системи не тільки для фіксації та відображення ПрС, а й для проектування та реалізації засобів підтримки, подальшої ефективної діяльності;

– опис (планування) послідовності операцій стосовно адаптації та самоорганізації системи з урахуванням таких методичних засад: *“гасіння пожеж”* – це процеси, якими можна керувати та планувати; *“ігнорування дрібниць”*, не звертаючи увагу на те, що *“компанії гублять дрібниці”*; *“штопання дірок”* – з'їдає ресурси, час, але не взмозі розв'язати проблеми адаптації; *“неперервні зміни”* – стиль життя організації; *“тління”* – пульс організації, тому не зупиняйте його.

Інформаційна підтримка цих процесів забезпечується операціями вилучення поточних (узгоджених в часі) значень певних сукупностей оперативних атрибутів (a, b, c, \dots) , що

характеризують стан системи формуванням тематично або предметно-орієнтованих інформаційних згорток $A_i(t_j)$ та веденням на їх основі *“історії”* діяльності – Сховища даних (СхД).

Отже, кожний запис СхД – це t_j – часовий зріз $\{A_{ij}\} = \{A_i(t_j) | i = 1, n; j = 1, m\}$.

Примітка:

1. A_{ij} – в більш вузькому розумінні, певна функціональна залежність від атрибутів (a_i, b_i, c_i, \dots) або $A_{ij} = f_i(a_{ij}, b_{ij}, c_{ij}, \dots)$.

2. Деякі атрибути можуть входити в різні інформаційні згортки, тобто для незалежних A_k і A_i перетин використаних для їх ідентифікації атрибутів

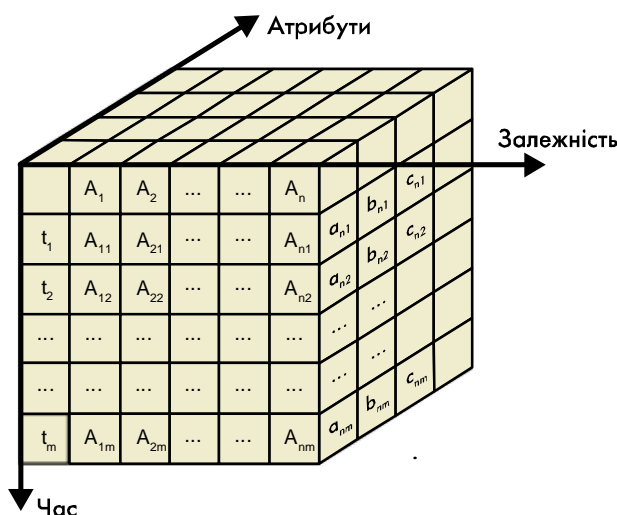


Рис. 1. Сховище даних

$$A_k \cap A_1.$$

Кожному з елементів A_{ij} відповідає домен $[A_{ij} - \varepsilon, A_{ij} + \varepsilon]$. З урахуванням цього будемо говорити, що стовпчик A_i матриці (рис. 1) складається з рівних з точністю ε значень $A_{i1}, A_{i2}, \dots, A_{in}$. Вихід одного із значень за межі ε -інтервалу підключає в підсистемі, яка “веде” цей показник, механізми регулювання. Для аналізу життєздатності системи інтерес представляють ті показники, які у зв'язку зі змінами в умовах існування не регулюються існуючими механізмами, що і потребує їх реорганізації. Крім того, СхД може зберігати факти, які не були відомі Розробникам системи. Наприклад, факт існування прихованої залежності пари інформаційних згорток A_k та A_i , не зважаючи на їх різну тематичну спрямованість. Виявлення таких фактів може стимулювати роботи стосовно більш масштабної реорганізації, пов'язаної, наприклад, з перерозподілом функцій між підсистемами і та ін.

І в першому, і у другому випадках “ідеї реорганізації” підказує аналіз поведінки $\{A_i\}$ або аналіз часових рядів $\{A_i\}$ – стовпчиків матриці “ t_j – зріз СхД”. Їх розкопки (data mining) у процесах “роздумів” над записами СхД дозволяють аналітику виявити факти – моменти істини як не відомі раніше причини або закономірності, яким вони підпорядковуються.

Представимо схему роздумів аналітика в умовах моменту істини у вигляді такої послідовності міркувань:

а) спостерігається ланцюжок минулих явищ та подій діяльності системи або часовий ряд інформаційних згорток $\{A_i\}$ і відповідних часових рядів базових атрибутів $(a_{n2}, b_{n2}, c_{n2}, \dots)$;

в) звертається увага на певні сталі характеристики (інваріанти) явищ або значень показників;

с) висувається апріорна гіпотеза про наявність таких же інваріантів у всіх подібних випадках і встановлюється прототип функціональної залежності між відповідними показниками;

д) на цій основі формуються (модифікуються існуючі) та тестуються на накопиченому фактичному в СхД матеріалі [7, 8] нові *правила виводу та узагальнення даних діяльності* з метою використання їх у подальших прогнозних та управляючих процесах.

Отже, вся інформація, яка необхідна для процедури розкопок $A_i(t_j)$, представлена в неявному виді в сукупності m – спостережень. Ключовим моментом обґрунтування процесу індуктивного виводу є структура навчальної вибірки та її імовірнісні розподіли.

Засади індуктивного виводу та пошуку нових функціональних залежностей

СхД (у більш широкому розумінні, просторово-часовий образ IC) – множина інформаційних згорток $A_i(t_j)$, які можуть належати певному класу проблемних ситуацій X.

З використанням типових схем індуктивного узагальнення [8]: (1) подібності, (2) відміни, (3) залишки в табл. 1 представлені приклади логіки можливих індуктивних висновків за результатами аналізу записів СхД.

Таблиця 1. Схеми індуктивного виводу

№ з/п	Вхідні дані спостережень	Індуктивне узагальнення
1	$\{a, b, c\} \subseteq A_1$	<p>Для всіх A з властивостями</p> $a \subseteq A_i \rightarrow \{A_i\} \subseteq X$ <p>Примітка: значення n не визначає "точності" виводу. Інтуїтивно відчувається, що n повинно бути великим</p>
	$\{a, e, d\} \subseteq A_2$	
	$\{a, f, g\} \subseteq A_3$	
	
	$\{a, h, j\} \subseteq A_n$	
2	$\{a, b, c\} \subseteq A_1$	$a \subseteq A_k \rightarrow \{A_k\} \subseteq X$ $(1 \leq k \leq n)$ <p>(інформаційні згортки без властивостей b і c не належать класу X)</p>
	
	$\{a, e, d\} \subseteq A_k$	
	$\neg\{b, c\} \subseteq A_{k-1} \rightarrow A_{k-1} \notin X$	
3	$\{a, b, c\} \subseteq A_1 \rightarrow A_1 \subseteq X_1 \cap X_2 \cap X_3$	$c \subseteq A_i \rightarrow \{A_i\} \subseteq X_3$
	$\{a\} \subseteq A_2 \rightarrow A_2 \subseteq X_2$	
	$\{b\} \subseteq A_3 \rightarrow A_3 \subseteq X_3$	

Для визначення виду (з певною мірою вірогідності $\mu(t)$) функціональної залежності аналізуються на "схожість поведінки" часові ряди $A_k(t)$ та $A_i(t)$ або їх відповідні похідні (перші, другі різниці). Для оцінки "схожості" пропонується багатоетапна процедура, на кожному з етапів якої виконується оцінка певних характеристик поведінки часових рядів.

Як приклад розглянемо наступні етапи аналізу емпіричних даних спостережень:

– виявлення тенденцій або тренду поточного середнього та дисперсії рядів, пов'язаних з аналізом $A_k(t)$ та $A_i(t)$. Якщо один з них є з виявленим трендом, а другий – ні, робиться висновок про "несхожість їх поведінки", і процедури порівняння продовжуються на наступних етапах;

– знаходження числа локальних екстремумів – при виявленні певних закономірностей у поведінці точок локального екстремуму показників вхідних рядів $A_k(t)$ та $A_i(t)$, що змістовно характеризує їх "схожість" в розподілі "вершин" та "низин", робиться висновок: "часові ряди схожі";

– розподіл інтервалів між локальними екстремумами – ряди з однаковими характеристиками розподілу фаз монотонності розвитку показників визначаються як ряди зі "схожістю поведінки".

За оцінкою "схожості" часових рядів $A_k(t)$ та $A_i(t)$ або їх похідних визначається вид можливої функціональної залежності між ними (табл. 2).

Таблиця 2. Приклади функціональних залежностей

Вид функціональної залежності	Оцінюється схожість для вхідних рядів	Обчислення коефіцієнтів
Лінійна $A_k(t) = \alpha_0 + \alpha_1 A_1(t)$	$(\Delta^1 A_k(t), \Delta^1 A_l(t))$, де Δ^1 – ряд перших різниць	$\alpha_0 - \alpha_1 A_{k1} = A_n$ $\alpha_0 - \alpha_1 A_{k2} = A_{l2}$
Квадратична $A_k(t) = \alpha A_1^2(t)$ або $A_k(t) = \alpha_0 + \alpha_1 A_1(t) + \alpha_2 A_1^2(t)$	$(\Delta^1 A_k(t) / \Delta^1 A_l(t))$, $\alpha \sum (A_1(t), A_k(t))$ $(\Delta^1 (\Delta^1 Y(t) / \Delta^1 A_1(t)), (\alpha_2 \Delta^2 A_1(t)))$ – де Δ^2 ряд других різниць відповідного ряду	$\alpha_0 + \alpha_1 A_{11} + \alpha_2 A_{11}^2 = A_{kl1}$ $\alpha_0 + \alpha_1 A_{12} + \alpha_2 A_{12}^2 = A_{kl2}$ $\alpha_0 + \alpha_1 A_{13} + \alpha_2 A_{13}^2 = A_{kl3}$
Гіперболічна $A_k(t) = A_1^{-1}(t)$	$(A_k(t), A_1^{-1}(t))$, де $A_1^{-1}(t) = (A_{11}^{-1}, A_{12}^{-1}, \dots)$	
Експоненціальна $A_k(t) = e^{A_1(t)}$	$(\ln A_k(t), A_1(t))$	

Примітка:

1. Під функціональною залежністю двох проблемних ситуацій ми розуміємо можливі взаємні переходи між ними.
2. α – дорівнює з певною мірою вірогідності $\mu(t)$, яка вибирається для кожного ПрО.
3. Змістовно “інтегральна” та “диференціальна” залежності визначають зв’язок швидкостей змін значень показників відповідних часових рядів.

3. Висновки

Відмітимо, що не можуть існувати універсальні правила отримання раціональних знань з емпіричних узагальнень. Але, на нашу думку, мають місце нетривіальні, проблемно (предметно)-орієнтовані правила, які можуть бути використані при певних, добре описаних умовах і можуть бути корисними у процесах інтерактивних суб’єктивних роздумів.

Отримані за розглянутими в роботі схемами індуктивного узагальнення та визначення і подальшого тестування можливих функціональних залежностей “раціональні елементи нових знань” можуть бути на перших порах і некоректними з точки зору *точного логічного виводу*, але арбітром, як завжди, може виступати лише практика їх використання.

Викладене можна розглядати як пропозиції до реалізації у складі ІС інтелектуалізуючих її механізмів та інструментів ситуативного аналізу виробничих процесів ПрО та пов’язаних з ними певних показників оточуючого середовища з використанням накопиченого емпіричного матеріалу діяльності з метою виявлення можливих проблем поточної самоорганізації, адекватної змінним умовам існування.

СПИСОК ЛІТЕРАТУРИ

1. Морозов А.О. Ситуаційні центри – основа керування організаційними системами // Математичні машини та системи. – 1997. – № 1. – С. 7 – 10.

2. Асельдеров З.М., В'юн В.І., Морозов А.О. "Континуум розумності" інформаційних систем // Матеріали конференції "Искусственный интеллект". – Донецьк: Кацивелі. – 2004. – Вересень.
3. В'юн В.І. Інтелектуалізація ІС – механізми та інструментарій інтерактивного ситуативного аналізу // Математичні машини та системи. – 2004. – № 3. – С. 125 –131.
4. Морозов А.О. Базы знаний в системах ситуационного управления коллективного пользования // УСИМ. – 1995. – № 4/5. – С. 91 – 95.
5. Балабанов А.С. Выделение знаний из баз данных – передовые компьютерные технологии интеллектуального анализа данных // ММС. – 2001. – № 1–2. – С. 40 – 54.
6. Саймон Н.А. Склады данных // Открытые системы. – 1997. – № 3. – С. 5 – 15.
7. Спирли Э. Корпоративное хранилище данных. – Москва: Изд. дом "Вильямс", 2001. – 396 с.
8. Ивашко В.Г., Финн В.К. Экспертные системы и некоторые проблемы их интеллектуализации // Семиотика и информатика. – 1986. – Вып. 27. – 164 с.