

ПОБУДОВА МОДЕЛІ ДОКУМЕНТООБІГУ ВІРТУАЛЬНОГО ПІДПРИЄМСТВА НА БАЗІ КОНЦЕПЦІЇ ВИСНОВКУ ЗА АНАЛОГІЯМИ

Розглянуто підхід до побудови моделі системи електронного документообігу віртуального підприємства, що дає змогу автоматизувати на її основі процес управління потоками документів. За основу побудови пропонується використовувати широко апробований апарат, що добре зарекомендував себе, – CBR-метод. Концептуальні положення цієї статті можуть бути використані для розв'язання теоретичних проблем електронного документообігу й створення на їхній основі відповідного прикладного програмного забезпечення.

Ключові слова: електронний документообіг, CBR-метод.

Загальна постановка проблеми

Віртуальна організація становить мережу ділового співробітництва, що охоплює основний бізнес цієї організації, її зовнішнє оточення (постачальників, споживачів тощо), функціонування яких координується і поєднується за допомогою сучасних інформаційних технологій, засобів телекомунікацій. Саме останні разом із мережоподібними принципами організації багато в чому забезпечили формування віртуальних організацій, оскільки вони уможливають принципово не обов'язковою фізичну наявність менеджерів на робочих місцях. Віртуальні колективи групують людей за мірою виникнення необхідності створювати певну вартість для задоволення специфічних потреб. При цьому не виникає фізичного колективу як організації, а відбувається лише об'єднання особливих відмітних здібностей у систему, що стає здатною створювати певну вартість. Мережоподібний характер віртуальних підприємств виявляється в тому, що глобальна мережа використовується для обміну необхідною інформацією між віддаленими, зокрема й географічно, підрозділами компанії. Це мережева організаційна форма, що базується на єдиній інформаційній системі (ІСВП), яка є сукупністю інформаційних систем кожного підприємства-партнера, об'єднаних загальним документообігом, таких, що кожна з систем виконує частину завдань з керування прийняттям рішень, а всі системи разом забезпечують функціонування підприємства [6; 4]. Таким чином, одна з ключових частин в ІСВП віртуального підприємства – система електронного документообігу [5].

Постановка завдань дослідження

Для розроблення системи електронного документообігу віртуального підприємства, якою можливо керувати, необхідно:

1. Дослідити, виявити та сформулювати специфічні характеристики руху документів у віртуальному підприємстві.

2. Побудувати модель документообігу віртуального підприємства, яка забезпечить створення вимірних систем, а також можливість швидко та ефективно керувати потоками документів, й тим самим змінювати бізнес-процеси, що, забезпечить компанії підтримку в безперервній адаптації до мінливої кон'юнктури ринку [3].

Розв'язання завдань дослідження

Як правило, партнерство учасників ВП укладається на певний термін або до досягнення визначеного результату (наприклад, виконання замовлення). Інакше кажучи, партнерство тимчасове і, наприклад, на певних етапах життєвого циклу виробу або при зміні ринкової ситуації можна залучати нових партнерів або позбуватися старих, що потребує змін в системі документообігу.

Природно, що підприємства-партнери для ефективного функціонування всієї мережі повинні базуватися на погодженому господарчому процесі. Коли ж, наприклад, для найкращої відповідності ринковим потребам у віртуальне підприємство об'єднується велика множина підприємств, найчастіше віддалених географічно, тоді таким підприємствам важко узгодити свої дії без оперативної комунікації документів [7].

Отже, для розв'язання інформаційних проблем ВП повинно мати єдину систему електронного документообігу, засновану на широкому застосуванні нових комунікаційних технологій, що робить актуальними проблеми автоматизації зберігання пошуку й обробки документів, гарантування безпеки їхньої передачі й можливості спільного використання.

Дослідження систем документообігу ВП за-свідчило, що їх можна розділити на дві групи (відповідно до інформаційних об'єктів):

- 1) тільки електронні документи;
- 2) як електронні документи, так і бізнес-процеси.

Інтерес представляють системи, що належать до другої групи, тому що, на думку багатьох експертів, розглядання разом як документів, так відображених в них бізнес-процесів є одночасно конструктивним і універсальним, бо забезпечує автоматизацію документообігу й усіх бізнес-процесів підприємства в рамках єдиних концепцій і програмного інструментарію [7].

Процес електронного документообігу у віртуальному підприємстві можна представити як сукупність деяких елементів та їх відносин між собою.

Ці елементи можна розділити на три категорії: учасники, стани документів і дії учасників. Учасники документообігу – це співробітники партнерів-учасників віртуального підприємства, що здійснюють генерування, рух і термінірування документів [2].

Таким чином, поведінка учасника документообігу може бути презентована як послідовність станів документів. Сукупність усіх станів документів представляє кінцеву множину, яка повністю описує всі можливі сценарії поведінки учасників, враховуючи функції переходів для прийняття рішень про вибір наступного стану. Ці функції реалізуються з допомогою аналізу дій учасників документообігу. Вироблена дія визначає результативний стан, для якого вхідними даними для визначення вибору є поточний стан документа й учасник процесу, тобто реалізується документообіг, у якому на кожному кроці відбувається дія на підставі процесу, а на підставі аналізу поточного стану документа (виконавця) **приймається рішення про наступний стан документа.**

Моделювання такого підходу до розв'язання проблем, заснованого на досвіді минулих ситуацій, призвело до появи технології логічного висновку, заснованого на аналогіях (*Case-Based Reasoning*, або CBR-метод). CBR-методи базуються на простій тезі, що подібні завдання (проблеми) розв'язуються подібним чином.

У ряді ситуацій метод виведення на основі аналогій особливо ефективний, коли:

- основне джерело знань про завдання – досвід, а не теорія;
- рішення не унікальні для конкретної ситуації і можуть бути використані в інших випадках;
- метою є не гарантоване правильне рішення, а краще з можливих.

CBR-методи містять чотири основні етапи, що утворюють так званий CBR-цикл або цикл

навчання за прецедентами [8], структуру якого представлено на рис. 1.

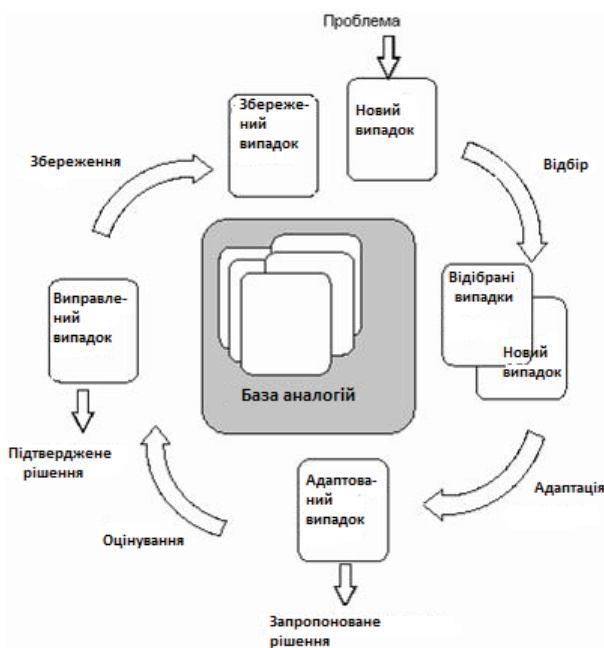


Рис. 1. Цикл висновку на основі аналогій

Таким чином, отримуємо метод, завдяки якому ми можемо робити висновки щодо певної проблеми або ситуації за результатами пошуку аналогій, що зберігаються в базі. Однак цей метод не створює будь-яких моделей або правил, узагальнюючих попередній досвід, – він у виборі рішення ґрунтується на масиві всіх доступних історичних даних, тому неможливо сказати, на основі яких саме факторів, висновуючи за аналогіями, вибудує конкретні рішення.

Проблеми, з якими стикається ця методологія: пошук найпридатніших аналогій і адаптація знайденого рішення надалі. В основі всіх підходів до вибору аналогій лежить той чи інший спосіб вимірювання ступеня близькості між ними і поточного випадку. За таких вимірів обчислюємо чисельне значення певної міри, що визначає склад множини аналогій, що потрібно обробити для досягнення задовільної класифікації або прогнозу.

Концепція виведення

Висновок на основі аналогій – це метод прийняття рішень, в якому використовуються знання про попередні ситуації чи випадки (аналогії) [1].

Під час розгляду поточного випадку відшукується схожий випадок як аналог. Замість того, щоб шукати розв'язання кожного разу, спочатку можна намагатися використати рішення, прийняте у схожій ситуації, можливо, адаптувавши його до ситуації, що змінилася у поточному випадку. Після того, як поточний випадок буде об-

роблений, він вноситься в базу аналогій разом з рішенням для можливого майбутнього використання. Більш формальне визначення дано в [10].

Аналогія – це опис ситуації в сукупності з докладним зазначенням дій, що застосовуються у цій ситуації чи для розв’язання певної проблеми. За [11], аналогія містить:

- Опис проблеми.
- Розв’язання цієї проблеми.
- Результат (обґрунтованість) застосування рішення.

В описі проблеми є вся інформація, необхідна для досягнення мети виведення (вибору найбільш відповідного рішення). Всі етапи застосованих дій зберігаються в описі рішення. Існує кілька способів відображення аналогій: від записів у базах даних, деревовидних структур – до предикатів і фреймів. Конкретне відображення має відповідати загальним цілям системи. Проблема подання аналогії – це насамперед вибір інформації, яку треба включати в опис аналогії, знаходження відповідної структури для опису змісту аналогії, а також визначення, яким чином повинна бути організована, та індексована база знань аналогій для ефективного пошуку і багаторазового використання.

Підхід, заснований на аналогіях, в цілому складається з таких компонентів [8]:

- Витяг з бібліотеки аналогій найбільш релевантних аналогій для поточного випадку.
- Адаптація обраного рішення для поточного випадку, якщо це необхідно.
- Застосування рішення.
- Оцінка застосування (перевірка коректності).
- Збереження. Додавання поточного випадку в базу аналогій.

Ефективність пошуку аналогії для поточного випадку багато в чому залежить від того, за якими ознаками організовано індекс у базі аналогій. Існують такі властивості ознак, які зроблять якісними ці індекси [12]:

Спрямованість: індекси мають бути спрямовані на досягнення мети.

Абстрактність: індекси мають бути досить абстрактні, щоб аналогія могла бути використана в різних запитах.

Конкретність: індекси мають чітко позначатися в інших ситуаціях без подальшої обробки.

Повноцінність: індекси мають бути здатні диференціювати аналогії.

Після того як аналогії витягнуті, потрібно вибрати з них найбільш придатну. Це визначається порівнянням ознак поточного випадку і вибраних аналогій. Для пошуку міри схожості аналогій можна використовувати метод «найближчого сусіда» (*nearest neighbour*) [9]. У його основі лежить той чи інший спосіб вимірювання ступеня близькості аналогії і поточного випадку за кож-

ною ознакою (будь це текстовий, числовий або булевський), який є корисним для досягнення мети, тобто вводимо метрику у простір всіх ознак, у цьому просторі визначаємо точку, що відповідає поточній нагоді, і в рамках цієї метрики знаходимо найближчу до неї точку з тих, що представляють аналогії. Зазвичай прогнозуємо на основі кількох найближчих точок, а не однієї (*K-nearest neighbours*). Такий метод стійкіший, оскільки дає змогу вирівняти окремі викиди, випадковий шум, завжди наявний у даних.

Для кожної ознаки визначено вагу, що враховує його відносну цінність. Повністю ступінь близькості аналогії за всіма ознаками можна обчислити, використовуючи узагальнену формулу виду:

$$\frac{\sum_j W_j \cdot \text{sim}(x_{ij}, x_{kj})}{\sum_j W_j},$$

де w_j – вага j -ї ознаки, *sim* – функція подібності (метрика), x_{ij} та x_{ik} – значення ознаки x_j для поточного випадку й аналогії, відповідно.

Після розрахунку ступенів близькості всі аналогії вишиковують в єдиний ранжирований список.

Вибір метрики (або ступеня близькості) – вузловий момент, від якого вирішальним чином залежить пошук відповідних аналогій. У кожній конкретній ситуації цей вибір робимо з урахуванням статистичної природи інформації.

Нехай є зразки X_i і X_k в N -мірному просторі ознак. Основні метрики, які традиційно використовують для вибору аналогій, вказано у таблиці 1. Після того як обрано відповідну аналогію, під час пошуку рішення для цільової проблеми виконується адаптація – модифікація наявного в ньому рішення задля його оптимізації. Неможливо виробити єдиний варіант для такої адаптації, як це здебільшого залежить від предметної галузі. Якщо існують алгоритми адаптації, вони зазвичай допускають наявність залежності між ознаками аналогії і ознаками рішення, які містяться в них. Такі залежності можуть задаватися людиною у побудові бази аналогій або виявлятися в базі автоматично. Процес модифікації рішення для його адаптації до поточної нагоди може містити кілька кроків, від простої заміни деяких компонентів у наявному рішенні, коригування або інтерполяції (числових) ознак або зміни порядку операцій, до більш суттєвих. Є й інші підходи:

- Повторна конкретизація змінних у чинній аналогії і присвоєння їм нових значень.

- Уточнення параметрів. Деякі аналогії можуть містити числові значення, наприклад, час виконання будь-якого етапу плану. Це значення має бути уточнено відповідно до нового значення іншої властивості.

• Пошук у пам'яті. Іноді потрібно знайти спосіб подолання труднощів, що виник як побічний ефект заміни одних компонентів рішення іншими. Зворотний зв'язок, що виникає під час збереження розв'язань для нових проблем, означає, що за своєю суттю висновок за аналогіями є самонавчальною технологією, завдяки чому робочі характеристики кожної бази аналогій з плином часу і накопиченням досвіду безперервно поліпшуються.

Висновки

У статті виконано дослідження функціонування системи електронного документообігу у

Таблиця 1. Основні типи метрик

Найменування метрики	Тип ознак	Формула для оцінки міри близькості (метрики)
Евклідова відстань	Кількісні	$d_{ik} = \left(\sum_{j=1}^N (x_{ij} - x_{kj})^2 \right)^{1/2}$
Манхеттенська метрика	Кількісні	$d'_{ik} = \sum_{j=1}^N x_{ij} - x_{kj} $
Міра подібності Хеммінга	Номинальні (якісні)	$\mu_{ij}^H = \frac{n_{ik}}{N}$ де n_{ik} – число збіжних ознак у зразків X_i й X_k
Міра подібності Роджерса-Танімото	Номинальні шкали	$\mu_{ij}^{R-T} = n_{ik}' (n_i' + n_k' - n_{ik}')^{-1}$, де n_{ik} – число збіжних одиничних ознак у зразків X_i й X_k ; n_i, n_k – загальне число одиничних ознак у зразків X_i й X_k , відповідно
Відстань Махалонбиса	Кількісні	$d_{ik}^M = (x_{ij} - x_{kj})^T W^{-1} (x_{ij} - x_{kj})$, W – ковариційна матриця виборки $X = \{x_1, x_2, \dots, x_n\}$
Відстань Журавльова	Змішані	$d_{ik} = \sum_{j=1}^N I_{ik}^j$, де $I_{ik}^j = \begin{cases} 1, & \text{якщо } x_{ij} - x_{kj} < s \\ 0 & \text{інакше} \end{cases}$

Література

1. Варшавский П. Р. Методы правдоподобных рассуждений на основе аналогий и прецедентов для интеллектуальных систем поддержки принятия решений / П. Р. Варшавский, А. П. Еремеев // Новости искусственного интеллекта. – 2006. – №3. – С.39–62.
2. Круковский М. Ю. Концепция построения моделей композитного документооборота / М. Ю. Круковский // Математичні машини і системи. – 2004. – № 2. – 163 с.
3. Круковский М. Ю. Методология построения композитных систем документооборота. Математичні машини і системи. – 2004. – № 1. – К. : ИПММС НАНУ 101 – 114 с.
4. Рамазанов С. К. Организация виртуальных предприятий / С. К. Рамазанов, Л. А. Тимашова, С. А. Бондар, В. А. Лещенко. – Луганськ : Вид-во СЧУ ім. В. Даля, 2004. – 368 с.
5. Саттон М. Дж. Корпоративный документооборот / Майкл Дж. Д. Саттон. – М. : Азбука, 2002. – 448 с.
6. Тарасов В. Б. Виртуальные предприятия : свойства, технологии создания, компоненты инфраструктуры / В. Б. Тарасов, П. С. Шильников // Информационные технологии. – 2000. – № 9. – С. 40–84.
7. Тимашова Л. А. Інформаційні системи для сучасних бізнес-аналітиків / Л. А. Тимашова, Л. А. Бондар, В. А. Лещенко, Т. В. Ткаченко, А. Г. Кондріна. – К. : Вид-во Академії праці і соціальних відносин, 2004. – 483 с.
8. Aamodt A. Case – Based Reasoning : Foundational Issues, Methodological Variations, and System Approaches / A. Aamodt, E. Plaza // Artificial Intelligence Communications. IOS Press. – 1994. – Vol. 7, № 1. – P. 39–59.
9. Anand S. S., Hughes J. G., Bell D. A., Hamilton P. Utilising Censored Neighbours in Prognostication : Workshop on Prognostic Models in Medicine, Eds. Ameen Abu-Hanna and Peter Lucas, Aalborg, (AIMDM'99). – Denmark. – 1999. – P. 15–20.
10. Artificial Intelligence Techniques. Springer Verlag / Ed. A. Bundy, 1997.

11. Althof K.-D. A Review of Industrial Case – Based Reasoning Tools / K.-D. Althof, E. Auriol, R. Barlette, M. Manago. AI Intelligence, 1995.
12. Kolodner J. L. Maintaining Organization in a Dynamic Long-term memory / J. L. Kolodner // Cognitive Science. – 1983. – № 7(4). – P. 243–280.

О. Костюк

CONSTRUCTION MODELS DOCUMENT VIRTUAL ENTERPRISES ON THE BASIS OF CONCLUSIONS ANALOGOUS

An approach to building a model of virtual enterprise electronic document management to automate the process based on its flow management documents. As the foundation of building is proposed to use the machine extensively tested, well proven, CBR-method. Conceptions of this article can be used to solve theoretical problems of electronic document and a basis for their respective application software.

Keywords: electronic document-over, CBR-method.

Матеріал надійшов 29 березня 2011 р.

УДК 004.4'24

Замковий О. В.

ПРОТОТИПУВАННЯ АВТОМАТИЗОВАНОЇ МУЛЬТИАГЕНТНОЇ ПОШУКОВОЇ СИСТЕМИ УНІФІКОВАНОЮ МОВОЮ МОДЕЛЮВАННЯ АГЕНТІВ

Запропоновано модель мультиагентної пошукової системи по мережі Інтернет засобами уніфікованої мови моделювання агентів (AUML). Змодельована пошукова система класифікує веб-сторінки за певними типами та виконує спеціалізований пошук залежно від типу пошукового запиту.

Ключові слова: агент, пошук, AUML, типізація, мультиагентні системи

Вступ

Через зростання залежності людини від інформації, що надходить з Інтернет (її актуальності, своєчасності, повноти і т. д.) пошук і розпізнавання інформації в мережі перетворюється із разової операції у тривалий або навіть постійний процес. Для автоматизації такого пошуку необхідно використовувати спеціалізовані прикладні програми, які відстежують зміни інформації, що перебуває у сфері інтересів конкретного користувача. Такі програми дістали назву *агенти Інтернет*.

Агентами Інтернет називають програми, що автоматизують пошук, розпізнавання, отримання та аналіз інформації у мережі, орієнтовані на потреби конкретного користувача (або групи користувачів).

Програмування агентів Інтернет є одним із найцікавіших напрямків у сфері пошуку й розпізнавання інформації в мережі, що активно розвивається останнє десятиліття. Перспективність

цього підходу продемонстрована в ряді дослідницьких і комерційних проектів [3, с. 115–118; 6, с. 2–3; 8, с.4–8; 4, с. 9–10].

У цій статті автор моделює основні модулі автоматизованої мультиагентної пошукової системи (АМАПС), яка здійснює спеціалізований пошук на основі визначених типів сайтів у Інтернет [опис системи в статті: 2, с. 234–239]. Новизна роботи полягає в створенні пошукової системи, що здійснює розбиття множини всіх сайтів на спеціалізовані підмножини, які відповідають певному типу (наприклад, магазини, розважальні сайти тощо), та виконує типізований пошук залежно від типу пошукового запиту. Автор створює прототипи основних агентів АМАПС за допомогою уніфікованої мови моделювання агентів (AUML [4, с. 13]), що є стандартом організації зі стандартизації агентних технологій FIPA [4, с. 10–27] для моделювання мультиагентних систем.