

*Н.В. ШАРОНОВА*, д-р техн. наук, проф., зав. каф. НТУ "ХПИ"  
(г. Харьков),  
*В.А. ТАРЛОВСКИЙ*, ст. преп. ХЭПИ (г. Херсон),  
*Н.Ф. ХАЙРОВА*, канд. техн. наук, доц. НТУ "ХПИ" (г. Харьков)

## **МЕТОД ПЕРСОНИФИКАЦИИ ИНТЕЛЛЕКТУАЛЬНОГО КОРПОРАТИВНОГО РЕСУРСА КОМПАНИИ**

Проведен анализ существующих проблем представления знаний в современных корпоративных информационных системах. Показана возможность использования математического аппарата алгебры конечных предикатов для моделирования интеллектуальных активов компании. Предложено использование метода компараторной идентификации для разбиения модели знаний предметной области на изменяемые области деятельности менеджера. Библиогр.: 14 назв.

**Ключевые слова:** представление знаний, алгебра конечных предикатов, интеллектуальные активы компании, компараторная идентификация.

**Постановка проблемы.** В связи с лавинообразным ростом объема информации, влияющей на качество принятия решения менеджментом различного уровня, все больше организаций проявляют интерес к программам класса Knowledge Management System (KMS), обеспечивающим интегрированный подход к доступу и использованию информационных ресурсов организации. Из общего ресурсного потока организации, к которому в общем случае относятся структурированные базы данных, текстовая информация и неявные знания, программы KMS выделяют и управляют именно ценными знаниями (интеллектуальными активами) компании.

К интеллектуальным активам компании (ИА) относятся явные знания, обладающие определенным уровнем ценности для компании, т.е. качественная семантическая информация, выраженная знаковой формой и зафиксированная на любом материальном носителе. Для выявления явных знаний из неструктурированной информации, накапливающейся в разных формах и в разных хранилищах, наряду с известными методами полнотекстового поиска применяются аналитические технологии, поддерживающие нечеткий поиск, запросы на естественном языке, автоматическую кластеризацию результатов, а так же автоматическое аннотирование и реферирование.

На сегодняшний день, для реального управления запасами неструктурированных знаний на предприятии, накапливающихся в разных формах и в разных хранилищах, и их использования для принятия верных управленческих решений, менеджеру необходимо локализовать область своей деятельности персональной информационной зоной. Согласно аналитическим исследованиям, проведенным крупнейшей компанией Gartner [1], успех внедрения систем управления знаниями в крупных организациях будет недостаточным, и с 30-процентной вероятностью приведет к

"информационному потоку", при не достаточно четком разграничении рабочей области каждого менеджера или аналитика компании.

**Анализ литературы.** O'Brien I.A. [2] предложил следующее определение "базы знаний" в корпоративной информационной системе: "Хорошо структурированные, рассмотренные в многообразии взаимосвязей данные о свойствах сущностей предметной области и об отношениях между ними". Для представления знаний традиционно используется четыре типа моделей [3, 4, 5]: продукционная модель; формально-логическая модель; фреймовая модель; и семантико-сетевая модель. Основные задачи, которые выходят сегодня на передний план при разработке модели представления знаний современной корпоративной информационной системы:

- лингвистические ресурсы системы должны обладать возможностью расширяемости, позволяющей пользователям дополнять и расширять модель конкретной предметной области [6];

- возможностью персонализации работы, сохранения и обмена знаниями аналитиков и экспертов организации в виде персональных интеллектуальных ресурсов [7];

- возможность работы менеджеров в персональных информационных зонах, проведения исследований по отдельным сферам интересов [6, 7];

- возможность динамической классификации, меняющаяся в процессе получения документов текстовой информации [8].

Для решения вышеперечисленных задач необходимо использовать функциональные системы менеджмента, ядром которых станут интегрированные интеллектуальные информационные системы, включающие элементы искусственного интеллекта [9, 10], основанные на методах и средствах теории интеллекта [11, 12, 13].

**Цель статьи.** Используя средства теории интеллекта, разработать метод, позволяющий локализовать область работы менеджера корпорации индивидуальной информационной зоной, выделяя персональный интеллектуальный ресурс. Использование данного метода позволит разрабатывать и внедрять интегральные системы корпоративного менеджмента, позволяющие разбивать работу менеджеров на отдельные сферы интересов, а также оперативно изменять персональные информационные зоны менеджеров среднего звена корпорации.

**Математический аппарат алгебры конечных предикатов.** Обозначим некий универсум  $U$ , под которым будем понимать все возможные текстовые документы, поступающие менеджеру на обработку (справки, выписки, отчеты, распоряжения, решения и т.д.), понятия и объекты анализа рассматриваемой предметной области, специализированные словари, тезаурусы, отображающие специфику данной предметной области и т.д.

Из элементов универсума образуются подмножества  $M_{1i}, M_{2i}, \dots, M_{mi}$  в соответствии с конкретной задачей обработки информации корпоративной информационной системой. На декартовых произведениях  $M_{1i} \times M_{2i} \times \dots \times M_{mi}$  определяются предикаты  $P_j$ , характеризующие работу системы, выполняя ту или иную интеллектуальную обработку информационных потоков.

Предикатом  $P$ , заданным на  $U$ , называется любая функция  $\varepsilon = P(x_1, x_2, \dots, x_n)$ , отображающая множество  $U$ , в множество  $\Sigma = \{0, 1\}$ . Переменные  $x_1, x_2, \dots, x_n$ , называются предметными, а их значения предметами. Если множество  $U$  конечно, как при моделировании информационной системы корпорации, то и предикат  $P$  конечный [14].

Базисными предикатами для алгебры предикатов будут предикаты вида:

$$\begin{aligned} x_i^a &= 1, \text{ если } x_i = a, \\ x_i^a &= 0, \text{ если } x \neq a, \quad (1 \leq i \leq n), \end{aligned} \quad (1)$$

где  $i = 1, 2, \dots, n$ ;  $a$  – любой элемент универсума. Предикат вида (1) называется предикатом узнавания предмета  $a$  по переменной  $x_i$ . Если универсум конечен и состоит из  $m$  элементов, всего имеется  $m \times n$  различных базисных элементов. Алгебра предикатов полна в том смысле, что любой ее предикат можно представить в виде суперпозиции базисных операций, примененных к базисным элементам. На языке алгебры конечных предикатов могут быть описаны любые конечные отношения, поэтому любой другой математический аппарат, предназначенный для описания произвольных конечных отношений, в логическом смысле, обязательно будет эквивалентен алгебре конечных предикатов [12].

**Описание используемого метода.** Определим на множестве элементов информационной системы  $M = (D, T)$ , где  $D$  – множество документов, поступающих на вход системы, а  $T$  – множество терминологических понятий, определяющих персонифицированные области деятельности менеджеров, универсума  $U$ , отображение  $d \in D \rightarrow T$ . Тогда для любого подмножества  $D_k \subseteq D$  можно задать бинарное отношение неразличимости:

$$\text{int}(D_k) = \{(x, y) \in U \times U | (\forall d \in D_k) d(x) = d(y)\}. \quad (2)$$

Если  $(x, y) \in \text{int } D_k$ , то объекты  $x$  и  $y$  являются неразличимыми по отношению ко всем элементам множества  $D_k$ . На универсуме  $U$  можно ввести бинарное отношение неразличимости  $R$ . Отношение  $R$  эквивалентно отношению толерантности, являющегося экспликацией сходства или неразличимости. Отношение является отношением толерантности, если оно рефлексивно и симметрично [12].

Отношение  $R$  показывает толерантность обрабатываемой менеджером информации, представленной в текстовых документах, относительно интеллектуального ресурса менеджера, представленного в терминологических понятиях имеющегося словаря. Данное отношение позволяет выполнять

операции не с отдельными элементами универсума, а с кластерами толерантных элементов.

Полученные нами приближенные множества документов, содержащих информацию, поступающую в систему для обработки, и терминологических понятий, соответствующих области знаний менеджеров компании, можно описать средствами алгебры логики предикатов.

Введем на декартовом произведении  $D \times D$  универсума  $U$  предикат соответствия документов персонифицированному интеллектуальному ресурсу  $k$ -менеджера. Действительно, если  $G_2(t_1^k, t_2^k) = 1$ , то  $P(d, t_1^k) = P(d, t_2^k)$  для любого документа из области знаний менеджера:

$$G_1(d_1, d_2) = \forall t^k \in T^k (P(d_1, t^k) \sim P(d_2, t^k)). \quad (3)$$

Введем на декартовом произведении  $T^k \times T^k$  универсума  $U$  предикат соответствия терминологических понятий персонифицированному интеллектуальному ресурсу  $k$ -менеджера:

$$G_2(t_1^k, t_2^k) = \forall d \in D (P(d, t_1^k) \sim P(d, t_2^k)). \quad (4)$$

Предикаты  $G_1$  и  $G_2$ , определяемые выражениями (3) и (4) рефлексивны, транзитивны и симметричны, из чего следует, что они являются предикатами эквивалентности и однозначно определяются предикатом  $P$ . Предикат  $G_1(d_1, d_2)$  (1) можно использовать для объективного определения отношения двух любых документов (в данном исследовании текстовых документов)  $d_1$  и  $d_2$ , принадлежащих множеству  $D$ , к одной области знаний менеджера. Если,  $G_1(d_1, d_2) = 1$ , то при любом терминологическом понятии  $t^k$  множества  $T^k$ :  $P(d_1, t^k) = P(d_2, t^k)$ . Т.е. информация данных документов, выражаемая терминологическими понятиями из множества  $T^k$ , позволяет отнести их к области знаний  $k$ -менеджера. В противном случае, если  $G_1(d_1, d_2) = 0$ , то найдется такое терминологическое понятие  $t^k \in T^k$ , для которого  $P(d_1, t^k) \neq P(d_2, t^k)$ . В этом случае документы нельзя отнести к области знаний одного менеджера.

Предикат  $G_2(t_1^k, t_2^k)$  можно использовать для определения соответствия любых двух терминологических понятий, принадлежащих множеству  $T^k$ , к терминологическим понятиям персонифицированного интеллектуального ресурса одного менеджера. Действительно, если  $G_2(t_1^k, t_2^k) = 1$ , то  $P(d, t_1^k) = P(d, t_2^k)$  для любого документа  $d \in D$ . То есть терминологические понятия  $t_1^k$  и  $t_2^k$  либо одновременно относятся к области знаний менеджера, либо не относятся. Иначе говоря, во множестве документов  $D$  нет такого документа  $d \in D$ , который бы и относился и не относился к области знаний  $k$ -го менеджера, выражаемой терминологическими понятиями  $t^k \in T^k$ . Если же  $G_2(t_1^k, t_2^k) = 0$ , то найдется такой документ  $d \in D$ , для которого  $P(d, t_1^k) \neq P(d, t_2^k)$ . Т.е. либо документ  $d$  включает знания, выражаемые терминологическим понятием  $t_1^k$  и не включает знания, выражаемые

терминологическим понятием  $t^k_2$ , либо наоборот документ  $d$ , включает знания, выражаемые терминологическим понятием  $t^k_2$  и не включает знания, выражаемые терминологическим понятием  $t^k_1$ . В обоих случаях терминологические понятия  $t^k_1$  и  $t^k_2$  будут относиться к персонифицированному интеллектуальному ресурсу разных менеджеров.

Ясно, что документы, входящие в полученные классы эквивалентности не могут быть тождественными по смыслу безотносительно к терминологическим понятиям, отображающим область знаний того или иного менеджера, они являются эквивалентными относительно персонифицированной области знаний менеджера.

Используя предикат (4), можно определить разбиение  $\psi_1$  множества  $D$  на слои, при котором все документы, относящиеся к одному слою, будут относиться к области знаний одного менеджера. Любые же документы, взятые из разных слоев разбиения, будут содержать информацию, относящуюся к области знаний различных менеджеров.

Предикат (3) позволяет определить разбиение  $\psi_2$  множества терминологических понятий  $T$  на слои персонифицированных интеллектуальных ресурсов менеджеров корпорации. Любые два терминологических понятия, взятые из одного слоя разбиения, будут относиться к интеллектуальному ресурсу одного менеджера, а любые два документа, взятые из разных слоев разбиения  $\psi_2$ , будут относиться к интеллектуальным ресурсам различных менеджеров.

Классу  $\mathcal{G}_c$  всех документов,  $d \in D$ , относящихся к области знаний одного менеджера, включающей документ  $c \in D$ , соответствует предикат:

$$\mathcal{G}_c(d) = G_1(d, a). \quad (5)$$

Классу  $\Lambda_b$  всех терминологических понятий  $t^k \in T$ , относящихся к персонифицированному интеллектуальному ресурсу  $k$ -менеджера, включающему терминологическое понятие  $b \in T$ , будет соответствовать предикат:

$$\Lambda_b(t) = G_2(b, t). \quad (6)$$

Учитывая зависимости (5) и (6), получаем следующие формулы:

$$\mathcal{G}_c(d) = \forall t \in T (P(d, t) \sim P(c, t)), \quad (7)$$

$$\Lambda_b(t) = \forall d \in D (P(d, t) \sim P(d, b)). \quad (8)$$

Формулы (7), (8) выражают разбиение документов, поступающих на обработку, и терминологических понятий, относящихся к области интеллектуальных активов менеджеров корпораций, на классы

эквивалентности через предикат персонифицированного интеллектуального ресурса компании, однозначно определяемый классификатором.

**Выводы.** Таким образом, использование математического аппарата алгебры конечных предикатов и, в частности, метода компараторной идентификации, позволяет автоматически разбивать область знаний корпоративной информационной системы на персональные информационные зоны. Кроме того, учитывая динамизм внешней по отношению к бизнесу среды и возможность последующей аппаратной реализации данного метода, его использование в базе знаний предметной области корпоративной информационной системы позволит обеспечить менеджмент среднего звена возможностью эффективной работы с информацией, с более релевантными и обоснованными подходами при принятии решений.

**Список литературы.** 1. [http://www.knowledge-portal.com/knowledge\\_management\\_technologies/gartner\\_group\\_perspective.htm](http://www.knowledge-portal.com/knowledge_management_technologies/gartner_group_perspective.htm) [Электронный ресурс]. 2. *O'Brien I.A.* Management Information Systems: A Managerial End User Perspective / *I.A. O'Brien*. – Boston: IRVIN, 1990. – 650 p. 3. *Уотермен Д.* Руководств по экспертным системам / *Д. Уотермен*. – М.: Мир, 1989. – 388 с. 4. *Белоногов Г.Г.* Компьютерная лингвистика и перспективные компьютерные технологии / *Г.Г. Белоногов, Ю.П. Калинин, А.А. Хорошилов*. – М.: "Русский мир", 2004. – 248 с. 5. *Дубровин А.Д.* Интеллектуальные информационные системы: Учеб. пособие. Ч.1 / *А.Д. Дубровин*. – М.: МГУКИ, 2008. – 232 с. 6. *Бажин И.И.* Информационные системы менеджмента / *И.И. Бажин*. – М.: ГУ – ВШЭ, 2002. – 688 с. 7. *Гаскаров Д.В.* Интеллектуальные информационные системы / *Д.В. Гаскаров*. – М.: ФГУП "Издательство Высшая школа", 2003. – 464 с. 8. *Семенов Н.А.* Интеллектуальные информационные системы / *Н.А. Семенов*. – Тверь: Издательство ТГТУ, 2004. – 168 с. 9. *Business Process Reengineering: The Oracle Perspective.* ORACLE CONSULTING, 1994. 10. *Hammer M.* Reengineering the Corporation. A Manifesto for Business Revolutions / *M. Hammer, J. Champy*. – HarperBusiness, 1993. 11. *Шабанов-Кушнарченко Ю.П.* Компараторная идентификация лингвистических объектов: Монография / *Ю.П. Шабанов-Кушнарченко, Н.В. Шаронова*. – К.: ИСДО, 1993. – 116 с. 12. *Бондаренко М.Ф.* Теория интеллекта. Учебник / *М.Ф. Бондаренко, Ю.П. Шабанов-Кушнарченко*. – Харьков: Изд-во СМИТ, 2007. – 576 с. 13. *Шаронова Н.В.* Использование метода компараторной идентификации для разбиения семантического пространства предметной области знание-ориентированных систем / *Н.В. Шаронова, Н.Ф. Хайрова* // Вестн. Херсон. гос. техн. ун-та. – Херсон, 2007 – № 4 (27). – С. 39–42. 14. *Булкин В.И.* Использование метода декомпозиции бинарных предикатов при формализации интеллектуальной деятельности / *В.И. Булкин, Н.Ф. Хайрова, Н.В. Шаронова* // Вестн. Херсон. гос. техн. ун-та. – Херсон, 2005. – № 1. – С. 78–82.

УДК 519.767

**Метод персоніфікації інтелектуального корпоративного ресурсу компанії / Шаронова Н.В., Тарловский В.А., Хайрова Н.Ф.** // Вісник НТУ "ХПІ". Тематичний випуск: Інформатика і моделювання. – Харків: НТУ "ХПІ". – 2009. – № 43. – С. 191 – 197.

Проведений аналіз існуючих проблем представлення знань в сучасних корпоративних інформаційних системах. Показана можливість використання математичного апарату алгебри кінцевих предикатів для моделювання інтелектуальних активів компанії. Запропоновано використання методу компараторної ідентифікації для розбиття моделі знань наочної області в змінні області діяльності менеджера. Бібліогр.: 14 назв.

**Ключові слова:** представлення знань, алгебра кінцевих предикатів, інтелектуальні активи компанії, компараторная ідентифікація.

UDC 519.767

**Method of personification of the company intellectual corporate resource / Sharonova N., Tarlovsky V., Khairova N.** // Herald of the National Technical University "KhPI". Subject issue: Information Science and Modelling. – Kharkov: NTU "KhPI". – 2009. – №. 43. – P. 191 – 197.

In these article the analysis of existing problems of knowledge representation in modern corporate information systems is carried out. Possibility of use of a mathematical apparatus of algebra of finite predicates for modelling of company intellectual assets of the company is shown. The method of comparative identification is used for splitting of data domain knowledge model into dynamic changeable spheres of a manager activity. Refs: 14 titles.

**Key words:** knowledge representation, algebra of finite predicates, intellectual assets of the company, comparative identification.

*Поступила в редакцию 10.09.2009*