



МОДЕЛИРОВАНИЕ ПРОЦЕДУРЫ ПОИСКА ИНФОРМАЦИИ В ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ УПРАВЛЕНИЯ

УДК 004.896

ДОРОВСКАЯ Ирина Александровна

к.т.н., доцент кафедры информатики и социально-гуманитарных дисциплин Криворожского филиала Европейского университета.

Научные интересы: Информационные технологии мониторинга условий труда рабочих мест.

ДОРОВСКОЙ Дмитрий Владимирович

к.т.н., доцент кафедры информатики и социально-гуманитарных дисциплин Криворожского филиала Европейского университета.

Научные интересы: Интернет, маркетинг, информационные и маркетинговые технологии, мониторинг и диагностика горно-металлургического оборудования.

ФИЛИППЕНКО Анастасия Юрьевна

Старший преподаватель Криворожского филиала Европейского университета

Научные интересы: информационные технологии для нечетких знаний, разработка инновационного программного обеспечения для улучшения процесса производства, изучение теории принятия решений, изучение систем диагностики оборудования, изучение и применение нейронных сетей.

ВСТУПЛЕНИЕ. ПОСТАНОВКА ЗАДАЧИ

Деятельность интеллектуальной системы управления (ИСУ) – процессо-ориентированный поиск требуемых процедурных решений в системе распределенной информации и знаний. Корпоративные знания ИСУ (или интеллектуальные активы) – это многообразная информация, которую необходимо иметь для поддержки на высоком уровне основных бизнес-процессов организации, а также для быстрого и адекватного реагирования на различные воздействия.

Под ИСУ понимаем интеллектуальную информационную систему управления, поддерживающую автоматизацию функций управления и выдающую информацию для углубления интеллектуальных знаний и принятия управленческих решений [5]. Эффективность работы ИСУ зависит от степени формализации процессов

управления, обусловленной степенью участия управленца, при принятии решения на основе полученной информации, степенью достоверности, релевантности информации, на которой основываются принимаемые решения.

Под информационными ресурсами ИСУ понимаются некоторые сведения, данные, оформленные таким образом, чтобы обеспечивать удобство принятия решений в области целевой деятельности, полученные из общего информационного пространства предприятия (ПП) [4]. Общее информационное пространство ПП представляет собой знания, распределенные по всей ИСУ, включающие и интеллектуальный потенциал компании.

Кумулятивный рост релевантной информации, динамизм внешней к бизнесу среды, неопределенность в процессе принятия решений ведут к необходимости использования функциональных систем



управления, ядром которых станут интегрированные интеллектуальные информационные системы, включающие элементы искусственного интеллекта. Такие системы должны быть построены на знаниях экспертов, и обладать компетентностью (искусственно воспроизводить компетентность экспертов), кроме того, такие системы должны не только обрабатывать знания, но и извлекать новые. Наивысшей степени интеллектуализации информационной составляющей обладают корпоративные интеллектуальные информационные системы (КИИСУ), использующие базы знаний.

В связи с тем, что функции интеллектуально - информационного обеспечения корпораций оформились в самостоятельную, но недостаточно структурированную и слабо интегрированную в ИСУ отрасль, то без специальных технологий учета, регистрации, хранения и мобилизации информационных ресурсов, накопленные опыт и знания СУ не могут быть использованы в полной мере для решения насущных задач управления ИСУ, а это определило и цель исследований.

РЕШЕНИЕ ЗАДАЧИ

Таким образом, для увеличения семантической силы существующих моделей представления знаний и расширения размеров моделируемой предметной области (ПО) необходимо применять достижения в области компьютерной лингвистики, с созданием алгоритмов лингвистически и семантически безупречного «понимания» ЭВМ естественного человеческого языка, что в свою очередь, приведет к созданию моделей представления знаний на основе символьных и логических сетей.

Разработанная модель представления знаний ИСУ, лингвистические ресурсы которой имеют открытую архитектуру,

обладают возможностью расширяемости, позволяющую СУ дополнять и расширять модель конкретной ПО, со всеми терминами, объектами и связями. Данная модель обладает возможностью персонификации работы, сохранения и обмена знаниями аналитиков и экспертов организации в виде персональных интеллектуальных ресурсов; создание персональных фильтров и лингвистических ресурсов.

При разработке модели использовались методы и модели, теории интеллекта [3]. Данная модель позволяет реализовать отношения между областью интеллектуальных знаний СУ и достоверными глубинными знаниями, представленными в поступающих на обработку документах.

Введем универсум элементов U , включающий все возможные текстовые документы, поступающие в данной корпоративной интеллектуальной информационной системе СУ на обработку (справки, выписки, отчеты, распоряжения, решения и т.д.), а также понятия и объекты анализа рассматриваемой ПО, специализированные словари, тезаурусы, отображающие специфику данной ПО.

Из элементов универсума образуются подмножества $M_{1i}, M_{2i}, \dots, M_{mi}$, в соответствие с конкретной задачей обработки информации МИСУ. На декартовых произведениях $M_{1i} \times M_{2i} \times \dots \times M_{mi}$ определяются предикаты P_j , характеризующие работу системы. Предикатом P , заданным на U , называется любая функция $\varepsilon = P(x_1, x_2, \dots, x_n)$, отображающая множество U , в множество $\Sigma = \{0, 1\}$. При $n=1$ предикат P является унарным, при $n=2$ — бинарным, при $n=3$ — тернарным. Так как множество U при моделировании информационной системы корпорации, конечно, то и предикат P конечен. Предикаты, обозначаемые 1 и 0, называются тождественно ис-

тинными и тождественно ложными соответственно.

Множество всех n -арных предикатов, заданных на U^n , на котором определены операции дизъюнкции, конъюнкции и отрицания, называется алгеброй n -арных предикатов на U . При этом операции дизъюнкции, конъюнкции и отрицания являются базисными для алгебры предикатов. Алгебра предикатов при любом значении n является разновидностью булевой алгебры, в ней выполняются все основные тождества булевой алгебры [3]. Переменные x_1, x_2, \dots, x_n , называемые предметными, и их значения, называемые предметами, представлены предикатом узнавания предмета a , по переменной x_i , являющегося базисными для алгебры предикатов [3]:

$$x_i^a = \begin{cases} 1, & \text{если } x_i = a, \\ 0, & \text{если } x_i \neq a. \end{cases} \quad (1)$$

где $i = \{1, 2, \dots, n\}$, a – любой элемент универсума.

Будем использовать следующие предметные переменные, определяющие отношение документа, поступающего на обработку в МИСУ, к предметной области деятельности менеджера: l – ключевое слово или словосочетание документа; u – UDK документа и r – предметная рубрика, документа, поступающего для обработки в МИСУ. Данные переменные отражают суть документа, назначение и взаимосвязь его составляющих, т.е. объективно определяют истинные и достоверные глубинные знания документа, предоставляя информацию, необходимую для принятия решений.

Значения соответствующие предметных переменных представлены множествами L , U и R . Множество ключевых слов и словосочетаний определяется статистико-позиционными методами на этапах пред-

лингвистического анализа [1] $L = \{l^i\}$, $1 \leq i \leq 14$: l^1 = депозитные операции; l^2 = депонирование; l^3 = банк; l^4 = вексель; l^5 = акции; l^6 = куксы; l^7 = куртаж; l^8 = маклер; l^9 = пошленные тарифы; l^{10} = транзитная торговля; l^{11} = таможенные тарифы; l^{12} = налогообложение; l^{13} = подоходные налоги; l^{14} = цензовая стоимость.

Иерархическая классификация UDK представлена множеством значений $U = \{u^i\}$, $1 \leq i \leq 5$. Рубрикатор представляет собой универсальную иерархическую классификацию областей знаний, принятую для систематизации всего потока научно-технической информации. Множество значений рубрикатора представлено $R = \{r^i\}$, $1 \leq i \leq 4$, где r^1 = банки и банковская деятельность; r^2 = рынок ценных бумаг; r^3 = экономика фирмы; r^4 = финансовая экономика.

Таблица 1.

Отношения между предметными переменными q, l, u и r

r^1	u^1	l^{12}	q^1
r^4	u^2	l^{12}	q^2
r^1	u^1	l^{13}	q^3
r^1	u^1	l^{14}	q^4
r^1	u^3	l^1	q^5
r^2	u^4	l^1	q^6
r^1	u^3	l^2	q^7
r^1	u^3	l^3	q^8
r^1	u^3	l^4	q^9
r^2	u^4	l^4	q^{10}
r^2	u^4	l^5	q^{11}
r^2	u^4	l^6	q^{12}
r^2	u^5	l^7	q^{13}
r^2	u^5	l^8	q^{14}
r^1	u^3	l^8	q^{15}
r^4	u^2	l^9	q^{16}
r^4	u^2	l^{10}	q^{17}
r^4	u^2	l^{11}	q^{18}

Введем также, основное для наших рассуждений, понятие области интеллектуальных знаний СУ q . Под областью интел-

лектуальных знаний конкретного СУ ИСУ мы будем понимать нечетко определенную часть корпоративных знаний, используемую для стандартных управленческих задач данного СУ.

Область интеллектуальных знаний СУ формируется в сфере мышления и имеет внеязыковую природу. Но поскольку мысль не может существовать вне слова, под областью интеллектуальных знаний СУ подразумевается лексическая единица, представляемая значением некоторого словосочетания и выражающая определенное множество управленческих ситуаций. Введем достаточно четко очерченное множество областей интеллектуальных знаний СУ, обуславливающее множеством

рассматриваемых управленческих ситуаций $Q=\{q^i\}$, $1 \leq i \leq 18$.

Построим парадигматическую таблицу, отображающую связь между областью локализации деятельности СУ q^i и предметными переменными l, u и r (таблица 1):

Используя созданную таблицу отношений между предметными переменными q, l, u и r , характеризующими документы, поступающие на обработку в КИС, и локальную область деятельности менеджера, выражаем область локализации интеллектуальной деятельности менеджера q через значения предметных переменных r, l, u :

$$\begin{aligned} r^1 u^1 l^{12} = q^1; r^4 u^2 l^{12} = q^2; r^1 u^1 l^{13} = q^3; r^1 u^1 l^{14} = q^4; r^1 u^3 l^1 = q^5; \\ r^2 u^4 l^1 = q^6; r^1 u^3 l^2 = q^7; r^1 u^3 l^3 = q^8; r^1 u^3 l^4 = q^9; r^2 u^4 l^4 = q^{10}; r^2 u^4 l^5 = q^{11}; r^2 u^4 l^6 = q^{12}; r^2 u^5 l^7 = q^{13}; r^2 u^5 l^8 = \\ q^{14}; r^1 u^3 l^8 = q^{15}; r^4 u^2 l^9 = q^{16}; \\ r^4 u^2 l^{10} = q^{17}; r^4 u^2 l^{11} = q^{18}. \end{aligned} \quad (2)$$

Выполняем операцию почленной дизъюнкции возможно большего числа родственных равенств. Родственными равенствами назовем такие равенства, которые после выполнения над ними операции почленной дизъюнкции приводят к равенствам с левой частью в виде логического произведения, каждый сомножитель которого зависит от одной предметной

переменной [2]. Введение почленной дизъюнкции с использованием родственных равенств обусловлено необходимостью получения локальных областей интеллектуальных знаний СУ, определяемые именем конкретного СУ. Такие области могут включать больше чем одно исчисляемое ограниченное количество рубрик и предметных областей исследований.

$$\begin{aligned} r^1 u^1 (l^{12} \vee l^{13} \vee l^{14}) = q^1 \vee q^3 \vee q^4; \\ r^4 u^2 (l^{12} \vee l^9 \vee l^{10} \vee l^{11}) = q^2 \vee q^{16} \vee q^{17} \vee q^{18}; r^2 u^5 (l^7 \vee l^8) = q^{13} \vee q^{14} \\ r^1 u^3 (l^1 \vee l^2 \vee l^3 \vee l^4 \vee l^8) = q^5 \vee q^7 \vee q^8 \vee q^9 \vee q^{15}; r^2 u^4 (l^1 \vee l^4 \vee l^5 \vee l^6) = q^6 \vee q^{10} \vee q^{11} \vee q^{12}. \end{aligned} \quad (3)$$

Формируем функцию перехода от предметной области интеллектуальных знаний q к локальной области исследова-

ния СУ m , в профессиональную деятельность которого входит данная область исследования q .

$$\begin{aligned} q^1 \vee q^3 \vee q^4 \vee q^{13} \vee q^{14} = m^1; q^2 \vee q^{16} \vee q^{17} \vee q^{18} = m^2; \\ q^5 \vee q^7 \vee q^8 \vee q^9 \vee q^{15} = m^3; q^6 \vee q^{10} \vee q^{11} \vee q^{12} = m^4; \end{aligned} \quad (4)$$

Переопределим зависимость локальной области исследования СУ m от переменных r, l, u :

$$\begin{aligned}
 m^1 &= r^1 u^1 l^{12} \vee r^1 u^1 l^{13} \vee r^1 u^1 l^{14} \vee r^2 u^5 l^7 \vee r^2 u^5 l^8 = \\
 &= r^1 u^1 (l^{12} \vee l^{13} \vee l^{14}) \vee r^2 u^5 (l^7 \vee l^8); \\
 m^2 &= r^4 u^2 l^{12} \vee r^4 u^2 l^9 \vee r^4 u^2 l^{10} \vee r^4 u^2 l^{11} = r^4 u^2 (l^{12} \vee l^9 \vee l^{10} \vee l^{11}); \\
 m^3 &= r^1 u^3 l^1 \vee r^1 u^3 l^2 \vee r^1 u^3 l^3 \vee r^1 u^3 l^4 \vee r^1 u^3 l^8 = r^1 u^3 (l^1 \vee l^2 \vee l^3 \vee l^4 \vee l^8); \\
 m^4 &= r^2 u^4 l^1 \vee r^2 u^4 l^4 \vee r^2 u^4 l^5 \vee r^2 u^4 l^6 = r^2 u^4 (l^1 \vee l^4 \vee l^5 \vee l^6).
 \end{aligned} \tag{5}$$

Предикат $P(r, l, u, m)$, описывающий связь локальной области исследования менеджера и переменных, объективно определяющих, глубинные знания документа имеет следующий вид:

$$\begin{aligned}
 P(r, l, u, m) &= m^1 r^1 u^1 (l^{12} \vee l^{13} \vee l^{14}) \vee m^1 r^2 u^5 (l^7 \vee l^8) \vee \\
 &\vee m^2 r^4 u^2 (l^{12} \vee l^9 \vee l^{10} \vee l^{11}) \vee m^3 r^1 u^3 (l^1 \vee l^2 \vee l^3 \vee l^4 \vee l^8) \vee \\
 &\vee m^4 r^2 u^4 (l^1 \vee l^4 \vee l^5 \vee l^6).
 \end{aligned} \tag{6}$$

Определим бинарные отношения P_i , где $i \in \{r, l, u\}$, связывающие переменную m с предметными переменными r, l, u . Для чего произведем бинаризацию отношений (6).

Бинарный предикат P_r определяет отношения переменной m и предметной переменной r :

$$P_r(r, m) = m^1 (r^1 \vee r^2) \vee m^2 r^4 \vee m^3 r^1 \vee m^4 r^2 \tag{7}$$

Рассмотрим графическую интерпретацию модели. Отношение $P_r(r, m)$ можно отобразить в виде двудольных графов:

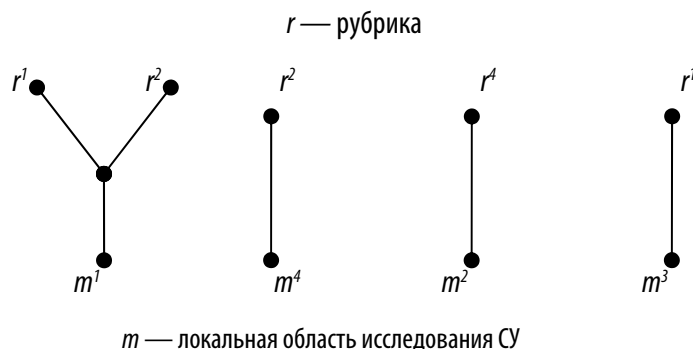


Рис 1. Двудольный граф предиката $P_r(r, m)$

Бинарный предикат P_l определяет отношения переменной m и предметной переменной l :

$$\begin{aligned}
 P_l(l, m) &= (l^{12} \vee l^{13} \vee l^{14} \vee l^7 \vee l^8) m^1 \vee (l^{12} \vee l^9 \vee l^{10} \vee l^{11}) m^2 \vee \\
 &\vee (l^1 \vee l^2 \vee l^3 \vee l^4 \vee l^8) m^3 \vee (l^1 \vee l^4 \vee l^5 \vee l^6) m^4.
 \end{aligned} \tag{8}$$

Отношение $P_l(l, m)$ можно отобразить в виде двудольных графов:

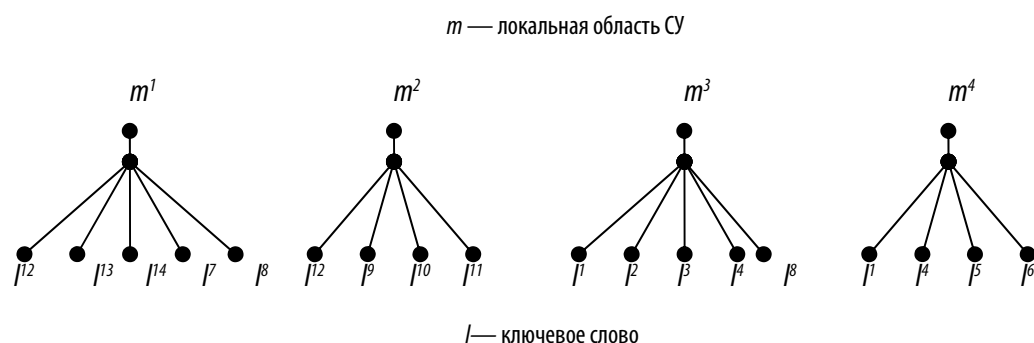


Рис.2. Двудольный граф предиката $P_l(l, m)$

Бинарный предикат P_u определяет отношения переменной m и предметной переменной u :

$$P_u(u, m) = m^1 (u^1 \vee u^5) \vee m^2 u^2 \vee m^3 u^3 \vee m^4 u^4 \quad (9)$$

Отношение $P_u(u, m)$ можно отобразить в виде двудольных графов:

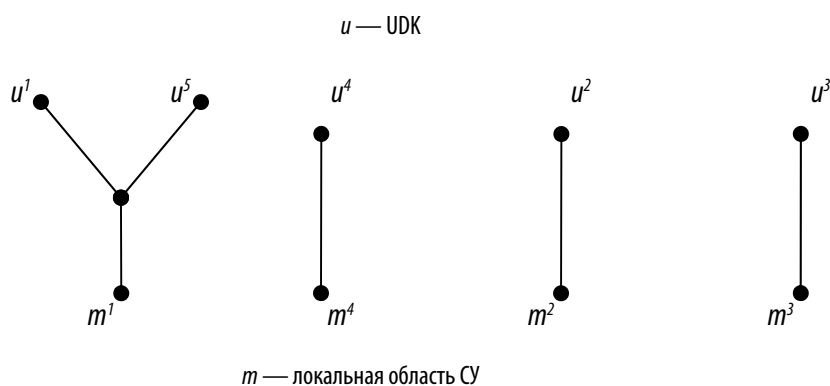


Рис.3. Двудольный граф предиката $P_u(u, m)$

Таким образом, построена математическая модель локализации области исследования СУ, характеризующаяся системой бинарных отношения P_l , P_r , P_u , отображаемыми двудольными графами на рисунках 1-3. Образуя конъюнкцию предикатов (7-9), получим предикат модели P , связывающий между собой предметные переменные r , l , u , m :

$$P(r, l, u, m) = P_r(r, m) \wedge P_l(l, m) \wedge P_u(u, m) \quad (10)$$

Предикат P можно наглядно изобразить в виде логической сети (рис.4), которая является графическим представлением результата бинарной конъюнкции декомпозиции многоместного предиката. Логическая

сеть состоит из полюсов и ветвей. Каждому полюсу логической сети ставится в соответствие своя предметная переменная модели, которая называется атрибутом этого полюса. В нашей сети четыре полюса обозначаются соответствующими предметными переменными r , l , u , m . С каждым полюсом связана область изменения атрибута полюса, т.е. его домен. Любой полюс логической сети в каждый момент времени несет какое-то знание о значении своего атрибута. Оно представляет собой одно из подмножеств домена полюса. Указывая состояние всех полюсов

в данный момент времени, определяем состояние сети в этот момент времени.

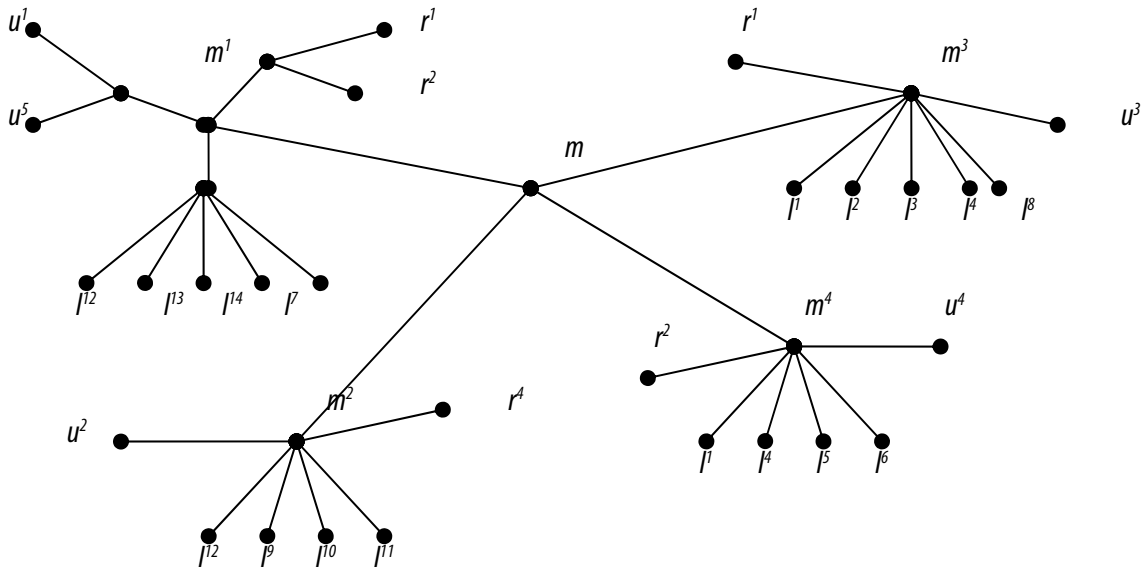


Рис. 4. Логическая сеть, связывающая предметные переменные r, u, l, m

Таким образом, разработанная математическая модель отношений между областью интеллектуальных знаний СУ и достоверными глубинными знаниями, представленными в документе, позволяет извлекать новые понятия и отношения, связывающие данные понятия. Т.е. позволяет извлекать пополнять новыми знаниями базы знаний МИСУ. Графическая интерпретация модели позволяет в дальнейшем реализовать ее аппаратно, что полностью решит проблему сокращения времени обработки документов, поступающих в ИСУ.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1. Показано использование методов теории интеллекта для моделирования процедуры извлечения информации для интеллектуальных активов компании. Разработана модель представления знаний в МИСУ, лингвистические ресурсы которой

имеют открытую архитектуру. Модель позволяет реализовать отношения между областью знаний СУ и достоверными глубинными знаниями, представленными в поступающих на обработку документах.

2. Предложены алгебраические принципы построения логической сети для формального описания локальной области деятельности СУ. Построена математическая модель локализации области исследования СУ, характеризующаяся системой бинарных отношений P_l, P_r, P_u , отображаемых двудольными графами.

3. Разработанная математическая модель отношений между областью интеллектуальных знаний СУ и достоверными глубинными знаниями, представленными в документе, позволяет извлекать новые понятия и отношения, связывающие данные понятия, извлекать и пополнять новыми знаниями базы знаний ИСУ.



ЛИТЕРАТУРА

1. V.S. Belov, Informacionno-analiticheskie sistemy: Osnovy proektirovanija i primenenija, uch.-prakt. posobie, izd. 2-oe, per. i dop., M. MGUJeSl, 2005
2. V.V. Buhtojarov "Jevoljucionnyj metod formirovanija obshhego reshenija v kollektivah nejronnyh setej", Zhurnal «Iskusstvennyj intellekt i prinjatje reshenij» pod gl. redakciej akademika S.V. Emel'janova, nomer 2010 / 03
3. Morozov M.N, Kurs lekcij po discipline "Sistemy iskusstvennogo intellekta", Laboratorija sistem mul'timedia, Marijskij gosudarstvennyj tehničeskij universitet.
4. Ju.F. Tel'nov, Intellektual'nye informacionnye sistemy v jekonomike. (Uchebnoe posobie) /M.: SINTEG, 2002. – 118 str.
5. Timashova L.A. Zadachi monitoringa i upravljenja biznes-processami virtual'nogo predprijatija / L.A. Timashova, L.P. Tur L.P, V.A. Muzaleva, V.A. Leshhenko, L.A. Janenko // Zb. nauk. pr. SNUJaEtaP. – Sevastopol': SNUJaEtaP, 2007. – Vip.4(24). – s. 244 – 253.

Рецензент: д.т.н., проф. Шерстюк В.Г.
Херсонский национальный технический университет