

МЕТОДЫ ФОРМИРОВАНИЯ, ИСПОЛЬЗОВАНИЯ И АНАЛИЗА БАЗ ЗНАНИЙ*

УДК 004

Василий Михайлович Трембач

к.т.н., доцент, доцент кафедры Прикладной информатики в экономике (ПИЭ), МЭСИ
Тел.: (499)199 84 07,
E-mail: trembach@yandex.ru

В статье представлены методы формирования, использования и анализа баз знаний, содержащих описания предметных областей. Рассматриваемые методы используют интегрированный подход к представлению знаний. Для формирования и анализа описаний предметной области используются сети Петри.

Ключевые слова: знания, база знаний, анализ баз знаний, интегрированный метод представления знаний, сети Петри.

Vasily Mihaylovich Trembach

PhD, the senior lecturer of applied informatics in economy chamber, Moscow State University of Economics, Statistics and Informatics (MESI)
Tel.: (495) 442-8098,
E-mail: trembach@yandex.ru

METHODS OF FORMATION, USAGE AND KNOWLEDGE BASE ANALYSIS

This article considers methods of formation, usage and knowledge base analysis, containing descriptions of object domain. Reviewed methods use integrated approach to knowledge representation. Petri nets are used for formation and analysis of descriptions of object domain.

Keywords: knowledge, knowledge base, knowledge base analysis, integrated approach to knowledge presentation, Petri net.

1. Введение

Современная мировая экономика становится все более динамичной и конкретной. Возникает необходимость разработки стратегии совершенствования процессов формирования и использования знаний, методов их анализа. В технологически развитых странах выделяются огромные ресурсы для перехода на знаниевую экономику. В знаниевой экономике основными источниками знаний являются коллективы людей и базы знаний, входящие в состав различных интеллектуальных информационных систем. Поскольку людские ресурсы для восприятия использования знаний ограничены, а объемы знаний растут по экспоненте, то важным направлением в переходе к знаниевой экономике является расширение области применения компьютерных методов формирования и использования знаний [2,13,14]. В статье рассматриваются методы формирования, использования знаний, а также формализованный подход к анализу и синтезу содержимого баз знаний.

2. Интегрированный подход к представлению знаний

Созданные, к настоящему времени, и используемые методы представления знаний имеют ряд недостатков [1,4,8,11], что влияет на возможности представления предметных областей и сужает множество допустимых операций над описаниями элементов действительности. С учетом расширения множества различных задач, требующих решения с использованием компьютерных технологий, повышаются требования, как к методам представления знаний, так и к методам их эффективного использования.

Для расширения области практического использования систем, основанных на знаниях, за счет совместного использования основных методов, предлагается интегрированный подход, в основе которого заложены семантическое и продукционное представление сущностей реального мира. Описание проблемной области – PAR, как множества представлений сущностей, связанных между собой отношениями, имеет следующий вид:

$$PAR = \{S, R, G\}, \quad (1)$$

где $S = \{Si\}$ – множество представлений сущностей проблемной области – PAR;

$R = \{Ri\}$ – множество взвешенных связей между сущностями Si ;

G – отображение, задающее между сущностями Si связи из заданного набора типов связей R .

Сущности проблемной области Si представляются вершинами-сущностями (в дальнейшем изложение вместо “вершины-сущности”, будет употребляться слово «вершина») сложноорганизованной сети, связываются связями из множества R и задаются атрибутами:

$$Si = (NSi, PRUSi, PSUSi, LNLSi, LNHSi, LNRi, LNAi, SPSi, WSi), \quad (2)$$

где NSi – имя сущности,

$PRUSi$ – предусловия активизации вершины,

$PSUSi$ – постусловия вершины,

$LNLSi$ – список имен вершин, определяющих Si и для рассматриваемой проблемной области связанных отношением “включает”;

$LNHSi$ – список имен вершин, определяемых вершиной Si ;

$LNRi$ – список имен связей для Si с другими вершинами,

WSi – вес/сила вершины Si .

При использовании интегрированного подхода к представлению знаний, где элементы действительности представляются вершинами-сущностями, для эволюции знаний о предметной области необходимо накапливать сведения, позволяющие формировать понятия о новых ситуациях [9,11,15]. С этой целью должны выделяться и запоминаться списки (подмножества) вершин, активизирующихся вместе с рассматриваемой i -й вершиной – LNA_i . Множество созданных списков вершин-ассоциаций LNA предназначается для формирования понятий, соответствующих новым ситуациям.

В процессе функционирования интеллектуальной информационной системы,

* Работа выполнена при финансовой поддержке РФФИ (проект № 10-07-00672).

активизация рассматриваемой вершины может происходить в различных ситуациях (при различных контекстах). Для кластеризации ситуаций, в соответствии с их контекстами, необходимо накапливать описания ситуаций, в которых происходит активизация рассматриваемой вершины. Множество описаний ситуаций, активизирующих i -е вершины $SPSi$, могут также рассматриваться и использоваться как прецеденты.

С учетом зависимостей (1-2) и внесенных дополнений, описание проблемной области PAR имеет следующий вид:

$PAR = (NK, PRU, PSU, LNI, LNO, LNR, LNA, WK)$, (3)

где NK – имя вершины-сущности,

PRU – предусловия активизации вершины,

PSU – постусловия,

LNI – список имен вершин, определяющих рассматриваемую,

LNO – список имен определяемых вершин,

LNR – список имен вершин-связей с другими вершинами,

LNA – список имен вершин, ассоциирующихся с рассматриваемой,

SPSi – множество представлений о ситуациях (прецедентов), активизирующих рассматриваемую вершину,

WS – вес/сила вершины.

3. Основные операции систем управления базами знаний

При формировании знаний о предметной области используются различные источники знаний – эксперты, организации, что требует одинаковой интерпретации формируемых знаний, не зависящей от специфики источника знаний, и используемых методов представления знаний. Для решения такой задачи необходимо использовать эталонные понятия, представляемые вершинами сети описания PAR. Разумно использовать эталоны для элементарных (базовых) понятий, которые позволяют представить любое понятие, сформированное и используемое информационной системой в ходе своего функционирования. Тогда понятие, поступающее из источника с другим форматом представления знаний, будет преобразовано к эталонному базису и комплексировано с существующими знаниями за счет применения определенных операций над знаниями.

Для работы с описаниями сущностей существуют различные операции. Способ выполнения операций часто

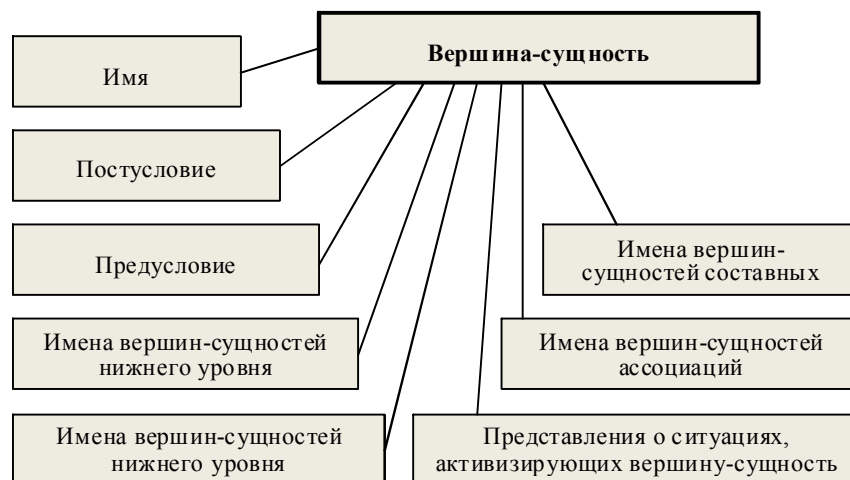


Рис. 1. Структура представления сущности действительности

зависит от используемого языка описания действительности [1,4,5,6,8. Ниже будут показаны способы выполнения операций применительно к интегрированному методу представления знаний [9,11,12].

В качестве основных операций для интегрированного подхода к представлению знаний выделяются: преобразование представления фрагмента знаний к полному виду; поиск фрагмента по образцу; вставка/исключение фрагмента; обобщение; операции сходства/различия, планирование.

Описания проблемных областей PAR могут задаваться множествами вершин, каждая из которых описывается в соответствии с выражением (3) и соответствует определенной сущности действительности. Таксономия понятий предполагает наличие таких элементарных понятий, которыми можно задавать любое понятие в рамках представления рассматриваемой действительности. Фрагмент знаний может представляться:

- в виде сетевой структуры понятий (вершин), где не все вершины являются элементарными;
- в виде набора имен вершин;
- именем вершины.

Представление фрагмента знаний является полным, если все входные вершины фрагмента являются элементарными. Для фрагментов, у которых не все входные вершины являются элементарными, необходимо выполнить операцию преобразования представления фрагмента к полному виду. Данная операция выполняется следующим образом. Рассматриваются описания всех входных вершин, которые не являются элементарными. В рассматриваемом описании из предусловий [9,11] выби-

раются имена признаков, образующих условие. Если соответствующая признаку вершина является входной, то дальнейшее рассмотрение описания этой вершины не проводится, а рассматриваются описания тех вершин, которые не являются элементарными. Операция преобразования представления фрагмента к полному виду завершается при условии, что все входные вершины описания фрагмента знаний являются элементарными.

Поиск фрагмента по образцу может выполняться различными способами. Если фрагмент сводится к одной вершине с именем (эта вершина является единственной выходной вершиной фрагмента знаний), то происходит поиск вершины с таким именем. Если такой вершины нет, то поиск прекращается, так как нет искомой вершины. Если вершина найдена, то происходит анализ найденной вершины на соответствие искомой.

При задании фрагмента несколькими вершинами поиск происходит через операцию преобразования представления фрагмента к полному виду.

Операции вставки/исключения основываются на машинном обучении методом взвешивания связей. Операции обобщения могут выполняться по тому же принципу, что и предыдущие. Отличием операций обобщения от операций вставки/удаления является содержание обучающих примеров. В качестве примеров для операции обобщения используются вершины, имена которых указаны в списке имен вершин-ассоциаций. Активизация этих вершин, вместе с рассматриваемой, является условием активизации новой, обобщающей вершины.

Операции сходства/различия бази-

руются на методе интерпретации рассматриваемого фрагмента с последующей оценкой полученных результатов. Важной операцией является операция планирования, включающая два этапа [9,10] – собственно формирование плана решения и реализацию этого плана.

4. Формирование описания предметной области

Формирование предметной области является сложной и ответственной задачей, поскольку неверно сформированный фрагмент описания проблемной области может оказать негативное воздействие на результаты работы всей базы знаний. Для снижения вероятности ошибок в описании предметной области целесообразно использовать вспомогательные инструменты, как для создания новых описаний элементов действительности, так и для анализа существующих. Рассматривается использование сетей Петри, которые предназначены для моделирования систем, состоящих из множества взаимодействующих друг с другом компонент [3,7]. В качестве компонент рассматриваются вершины-сущности [9]. Вершинам-сущностям, как и другим компонентам реальных систем присущ параллелизм.

В одном из подходов к проектированию и анализу различных систем, сети Петри используются, как вспомогательный инструмент анализа существующих систем.

При создании нового описания предметной области используется другой подход, при котором предполагается построение модели предметной области сразу в виде сети Петри, а затем сеть Петри преобразуется в описание предметной области в формате интегрированного метода представления знаний. Сеть Петри, моделирующая фрагмент описания предметной области, представлена на рис. 2. Построенная сеть Петри теперь может использоваться для анализа моделируемой составной связи.

5. Анализ баз знаний с помощью сетей Петри

Основными подходами к анализу сетей Петри [3,7] являются методы, основанные на использовании дерева достижимости, подход, основанный на матричном представлении сетей Петри, решении матричных уравнений и методах преобразования сетей.

На основе дерева достижимости могут выполняться: анализ безопасности и ограниченности, анализ сохранения, покрываемости, живости. Другой

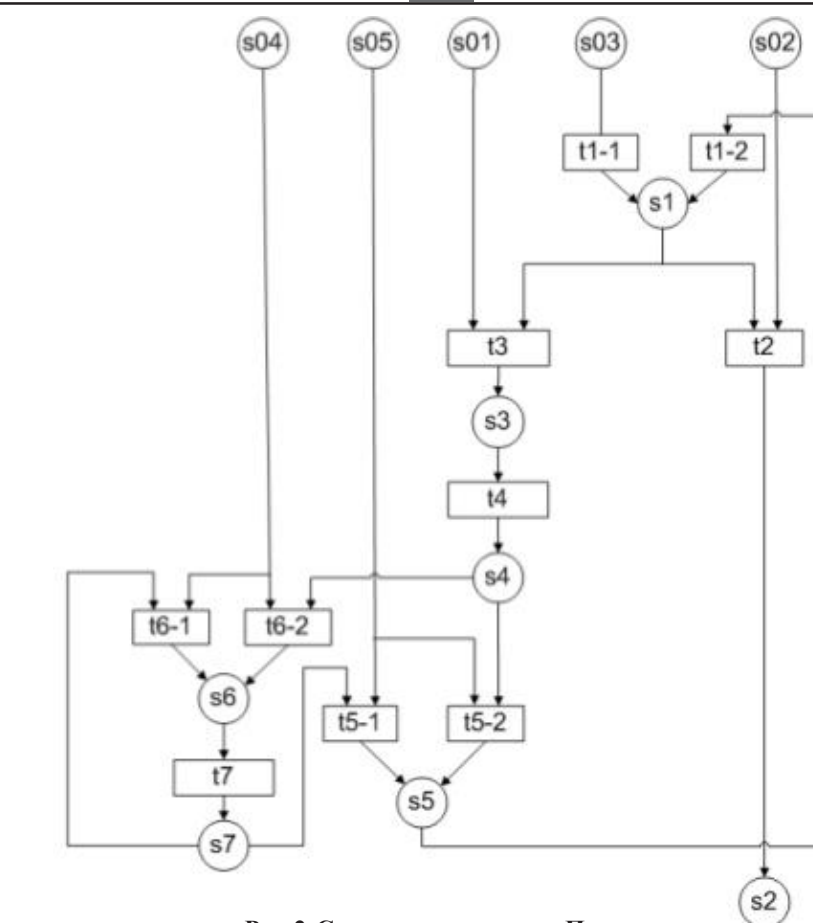


Рис. 2. Синтезированная сеть Петри

подход, к использованию сетей Петри, заключается в представлении элементов действительности в виде сети Петри, исследовании полученной сети Петри на наличие ошибок и в последующем преобразовании сети Петри в описание проблемной области на основе интегрированного подхода к представлению знаний.

Основными методами анализа свойств достижимости являются [3,7]:

- построение дерева достижимости;
- метод матричных уравнений.

Для синтезированной сети Петри (рис. 2) дерево достижимости представлено на рис.3. Из основных свойств сетей Петри, для составных связей, могут использоваться следующие: живость – отсутствие тупиковых состояний, исключая выходное (целевое) состояние составной связи; безопасность – метки вершин включают только «0» и «1», что означает отсутствие зацикливаний; правильность – сеть безопасна и живая; пассивность переходов – переходу не соответствует ни одной дуги графа.

Совокупность представленных подходов к формированию, использованию и анализу знаний о предметной области позволяет избежать многих

ошибок в описании предметной области.

6. Заключение

Стремительное развитие элементной базы, технологий производства вычислительных средств, разработка новых программно-аппаратных платформ позволяет создавать приложения для решения задач реальной сложности. В полной мере эти возможности относятся и к интеллектуальным системам и, в частности, к системам, основанным на знаниях. Размер баз знаний современных интеллектуальных информационных систем постоянно увеличивается и, соответственно, цена ошибки возрастает. Привлечение экспертов для выявления и исправления ошибок становится все более затратным процессом. Рассмотренные методы ориентированы на использование формальных методов. С их помощью возможно добавление новых сложных связей с использованием имеющихся, что позволяет описывать предметные области для задач реальной сложности. Анализ новых сложных связей с применением сетей Петри позволяет исключать ошибки, влияющие на формирование решений в динамической интеллектуальной

системе.

Перспективным направлением развития исследований является формализация процессов создания корректно построенных баз знаний и реализации этих процессов в вычислительных средах: локальных, распределенных. Создание корректно построенных баз знаний может выполняться, как с помощью классических технологий, так и с применением новых, например, облачных вычислений.

Литература

1. Гаврилова Т.А., Хорошевский В.Ф., Базы знаний интеллектуальных систем - СПб: Питер, 2000. 384 с..

2. Гаврилова Т.А., Муромцев Д.И. Интеллектуальные технологии в менеджменте: инструменты и системы. Учебное пособие. / Т.А.Гаврилова, Д.И.Муромцев; Высшая школа менеджмента СПбГУ. - СПб: Изд-во «Высшая школа менеджмента»; Издат.дом С.-Петербур. - гос. ун-та, 2008. - 488 с. 2-е издание, исправленное и дополненное.

3. Котов В.Е. «Сети Петри». - М.: Наука. Главная редакция физико-математической литературы, 1984. - 160 с.

4. Осипов Г.С. Лекции по искусственному интеллекту. - М.: КРАСАНД, 2009. - 272 с.

5. Осипов Г.С., Поспелов Д.А. Введение в прикладную семиотику Глава 5. Операции в семиотических базах знаний// Новости искусственного интеллекта. 2002. - № 6, стр. 28-35.

6. Осипов Г.С. Динамические интеллектуальные системы//Искусственный интеллект и принятие решений. - 2008.- №1. С. 47-54.

7. Питерсон Дж. Теория сетей Петри и моделирования систем: Пер. с англ. -М.: Мир, 1984. -264 с., ил.

8. Рассел, Стюарт, Норвиг, Питер. Искусственный интеллект: современный подход, 2-е изд.: Пер. с англ. - М.: Издательский дом «Вильямс», 2007. - 1408 с.

9. Трембач В.М. Методы представления эволюционирующих знаний, обеспечения и оценки их соответствия действительности. // В кн. КИИ - 2008. Одиннадцатая национальная конференция по искусственному интеллекту с международным участием. Труды конференции. Т.3., М.: ЛЕНАНД, 2008, с. 315-322.

10. Трембач В.М. Приобретение знаний из опыта и внешних источников. // В кн. КИИ - 2006. Десятая национальная конференция по искусственно-

му интеллекту с международным участием. Сборник научных трудов в 3-х томах. Т.3., М.: Физматлит, 2006, с. 1100-1108.

11. В.М. Трембач, Формирование и использование моделей компетенций обучающихся на основе эволюционирующих знаний// Научно-практический журнал «Открытое образование», МЭСИ, №6 (77), 2009, с. 12-26.

12. Трембач В.М.Применение интеллектуальных технологий к формированию компетенций обучающихся//Искусственный интеллект и принятие решений. - 2008.- №2. С. 34-54.

13. Тельнов Ю.Ф. Разработка государственного образовательного стандарта высшего профессионального образования 3-го поколения по направлению «Прикладная информатика» на основе компетентного подхода. // В кн. 2-я Российская научно-методическая конференция «Совершенствование подготовки IT-специалистов по направлению «Прикладная информатика» на основе инновационных технологий и Е-

LEARNING»: Сборник научных трудов / Моск. госуд. ун-т экономики, статистики и информатики – М., 2006.

14. Тельнов Ю.Ф. Реализация компетентного подхода к обучению на основе управления знаниями. В кн.: Научная сессия МИФИ - 2007. Сборник научных трудов. Т.3. М.: МИФИ, 2007, с. 40-42.

15. Тельнов Ю.Ф., Трембач В.М. Интеллектуальные информационные системы. Учебное пособие - М.: МЭСИ, 2009. - стр. 202.

Bibliography

1. Gavrilova T.A., Horoshevskij V.F., Bazy znaniy intellektual'nyh sistem - SPb: Piter, 2000. 384 s..

2. Gavrilova T.A., Muromcev D.I. Intellektual'nye tehnologii v menedzhmente: instrumenty i sistemy. Uchebnoe posobie. / T.A.Gavrilova, D.I.Muromcev; Vysshaja shkola menedzhmenta SPbGU. - SPb: Izd-vo «Vysshaja shkola menedzhmenta»; Izdat.dom S.-Peterb.gos. un-ta, 2008. - 488

```

μ₀=(1,0,1,0,1,0,0,0,0,0,0,0)
↓ t1-1
μ₁=(1,0,0,0,1,1,0,0,0,0,0,0)
↓ t3
μ₂=(1,0,0,0,1,0,0,1,0,0,0,0)
↓ t4
μ₃=(1,0,0,0,1,0,0,0,1,0,0,0)
↓ t5-2
μ₄=(1,0,0,0,1,0,0,0,0,1,0,0)
↓ t1-2
μ₅=(1,0,0,0,1,1,0,0,0,0,0,0)
↓ t3
μ₆=(1,0,0,0,1,0,0,1,0,0,0,0)
↓ t4
μ₇=(1,0,0,1,0,0,0,0,0,1,0,0)
↓ t6-2
μ₈=(1,0,0,1,0,0,0,0,0,0,1,0)
↓ t7
μ₉=(1,0,0,0,1,0,0,0,0,0,0,1)
↓ t5-1
μ₁₀=(1,0,0,0,1,0,0,0,0,1,0,0)
↓ t1-2
μ₁₁=(1,0,0,0,1,1,0,0,0,0,0,0)
↓ t3
μ₁₂=(1,0,0,0,1,0,0,1,0,0,0,0)
↓ t4
μ₁₃=(0,1,0,0,1,0,0,0,1,0,0,0)
↓ t9
μ₁₄=(0,1,0,0,1,0,0,0,0,1,0,0)
↓ t1-2
μ₁₅=(0,1,0,0,1,1,0,0,0,0,0,0)
↓ t2
μ₁₆=(0,1,0,0,1,0,1,0,0,0,0,0)

```

Рис.3. Дерево достижимости для синтезированной сети Петри

s. 2-e izdanie, ispravlennoe i dopolnennoe.

3. Kotov V.E. "Seti Petri". - M.: Nauka. Glavnaja redakcija fiziko-matematicheskoy literatury, 1984. - 160 s.

4. Osipov G.S. Lekcii po iskusstvennomu intellektu. - M.: KRASAND, 2009. - 272 s.

5. Osipov G.S., Pospelov D.A. Vvedenie v prikladnuju semiotiku Glava 5. Operacii v semioticheskikh bazah znaniy / Novosti iskusstvennogo intellekta. 2002. - № 6, str. 28-35.

6. Osipov G.S. Dinamicheskie intellektual'nye siste-my // Iskusstvennyj intellekt i prinjatие reshenij. - 2008. - № 1. S. 47-54.

7. Piterson Dzh. Teorija setej Petri i modelirovaniya sistem: Per. s angl. - M.: Mir, 1984. - 264 s., il.

8. Rassel, Stjuart, Norvig, Piter. Iskusstvennyj intellekt: so-vremennyj podhod, 2-e izd.: Per. s angl. - M.: Izdatel'skij dom "Vil'-jams", 2007. - 1408 s.

9. Trembach V.M. Metody predstavlenija jevoljucionirujuvix znaniy,

obespechenija i ocenki ih sootvetstviya dejstvitel'nosti. // V kn. KII - 2008. Odinnadcataja nacional'naja konferencija po iskusstvennomu intellektu s mezhdunarodnym uchastiem. Trudy konferencii. T.3., M.: LENAND, 2008, s. 315-322.

10. Trembach V.M. Priobretenie znaniy iz opyta i vneshnih istochnikov. // V kn. KII - 2006. Desjataja nacional'naja konferencija po iskusstvennomu intellektu s mezhdunarodnym uchastiem. Sbornik nauchnyh trudov v 3-h tomah. T.3., M.: Fizmatlit, 2006, s. 1100-1108.

11. V.M. Trembach, Formirovanie i ispol'zovanie modelej kompetencij obuchajuvihsja na osnove jevoljucionirujuvix znaniy // Nauchno-prakticheskij zhurnal «Otkrytoe obrazovanie», MESI, №6(77), 2009, s. 12-26.

12. Trembach V.M. Primenenie intellektual'nyh tehnologij k formirovaniju kompetencij obuchajuvihsja // Iskusstvennyj intellekt i prinjatие reshenij. - 2008. - №2. S. 34-54.

13. Tel'nov Ju.F. Razrabotka gosudarstvennogo obrazovatel'nogo standarta vys-shego professional'nogo obrazovaniya 3-go pokoleniya po napravleniju "Prikladnaja in-formatika" na osnove kompetentnostnogo podhoda. // V kn. 2-ja Rossijskaja nauchno-metodicheskaja konferencija "Sovershenstvovanie podgotovki IT-specialistov po napravleniju "Prikladnaja informatika" na osnove innovacionnyh tehnologij i E-LEARNING": Sbornik nauchnyh trudov / Mosk. gosud. un-t jekonomiki, statistiki i informatiki - M., 2006.

14. Tel'nov Ju.F. Realizacija kompetentnostnogo podhoda k obucheniju na osnove upravlenija znaniyami. V kn.: Nauchnaja sessija MIFI - 2007. Sbornik nauchnyh trudov. T.3. M.: MIFI, 2007, s. 40-42.

Tel'nov Ju.F., Trembach V.M. Intellektual'nye informacionnye sistemy. Uchebnoe posobie - M.: MESI, 2009. - str. 202.