



ОТКРЫТАЯ НАУКА  
издательство

Международный журнал информационных технологий и энергоэффективности

Сайт журнала:

<http://www.openaccessscience.ru/index.php/ijcse/>



УДК 004.853

## ВЫБОР МЕТОДОВ ИЗВЛЕЧЕНИЯ И ОБРАБОТКИ ЭКСПЕРТНОЙ ИНФОРМАЦИИ ДЛЯ БАЗЫ ЗНАНИЙ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

<sup>1</sup>Балашов О.В., <sup>2</sup>Букачев Д.С.

<sup>1</sup>Смоленский филиал АО «Радиозавод», Смоленск, Россия (214027, г. Смоленск, ул. Котовского д.2); e-mail: smradio@mail.ru

<sup>2</sup>ФГБОУ ВО Смоленский государственный университет, Смоленск, Россия (214000, г. Смоленск, ул. Пржевальского, 4), e-mail: dsbuka@yandex.ru

Существующие методы извлечения и обработки экспертной информации характеризуются значительными временными затратами и потерями извлекаемых знаний, особенно на этапе непосредственного взаимодействия инженера с экспертами. В настоящее время разработано несколько подходов, направленных на автоматизацию данного процесса. Однако применение современных систем извлечения и обработки экспертной информации не решает проблемы потерь извлекаемых знаний. Предлагаемое в данной статье дополнение классического процесса извлечения и обработки экспертной информации этапами «идентификация задачи» и «отбор соответствующих методов извлечения и обработки экспертной информации» позволяет уменьшить потери извлекаемых и обрабатываемых знаний.

Ключевые слова: экспертная информация, извлечение знаний, база знаний, система поддержки принятия решений.

## SELECTION OF METHODS FOR EXTRACTING AND PROCESSING OF EXPERT INFORMATION FOR THE KNOWLEDGE BASE OF DECISION SUPPORT SYSTEM

<sup>1</sup>Balshov O.V., <sup>2</sup>Bukachev D.S.

<sup>1</sup>Smolensk branch of joint-stock company "Radio factory", Smolensk, Russia, (214027, Smolensk, street Kotovskogo, the house 2); e-mail: smradio@mail.ru

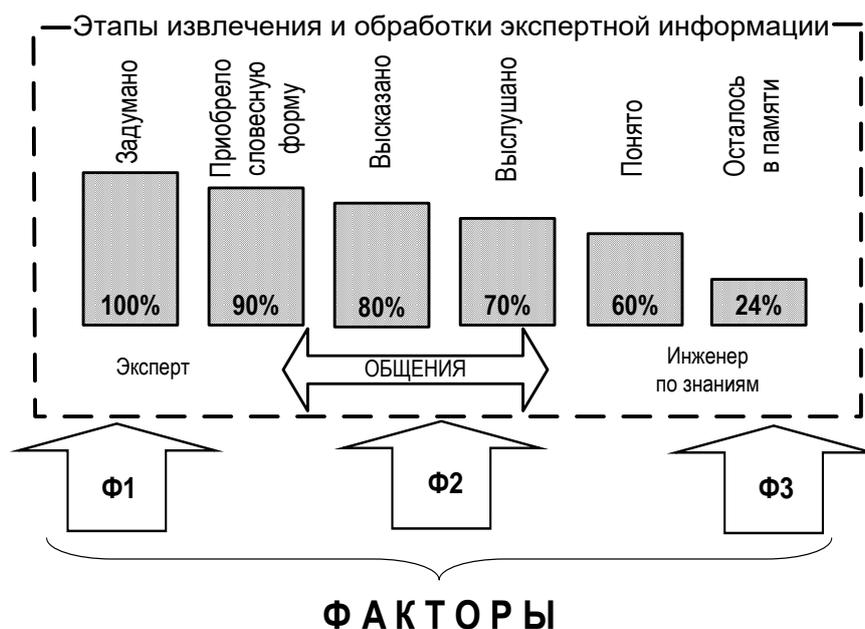
<sup>2</sup>Federal State Educational Institution of Higher Education Smolensk State University, Smolensk, Russia (214000, Smolensk, street Przewalski, 4); e-mail: dsbuka@yandex.ru

Existing methods of extraction and processing of the expert information are characterised by considerable time expenses and losses of taken knowledge, especially at the stage of direct interaction of the engineer with experts. Currently developed several approaches aimed at automating this process. However, the use of modern systems for extracting and processing expert information does not solve the problem of loss of taken knowledge. The proposed addition of the classical process of extracting and processing expert information in this article with the stages "task identification" and "selection of appropriate methods for extracting and processing expert information" reduces the loss of taken and processed knowledge.

Key words: expert information, knowledge extraction, knowledge base, decision support system.

Развитие вычислительной техники, а также интеллектуальных систем на её основе, связано с извлечением и обработкой экспертной информации. Экспертная информация требуется для последующего наполнения базы знаний системы поддержки принятия решений (СППР). Создание и сопровождение систем поддержки принятия решений требует реализации методов инженерии знаний, определяющих порядок извлечения и обработки экспертной информации [1, 2, 4]. При этом выбор методов извлечения и обработки осуществляется исключительно на основе личных профессиональных качеств инженера по знаниям, его опыта и интуиции.

Существующие методы извлечения и обработки экспертной информации характеризуются значительными временными затратами и потерями извлекаемых знаний до 60%, особенно на этапе непосредственного взаимодействия инженера по знаниям с экспертами (рисунок 1) [2-4].



$$\Phi = \{\Phi 1, \Phi 2, \Phi 3\} = \{\text{гносеологический, психологический, лингвистический}\}$$

Рисунок 1 – Потери информации при извлечении и обработке экспертной информации

Для решения данной проблемы разработано несколько подходов, направленных на автоматизацию этого процесса, на основе которых создавались различные системы приобретения знаний [2-4]. Проводимые исследования по извлечению и обработке экспертной информации направлены на создание инструментальной программной поддержки

деятельности инженера по знаниям и эксперта. Среди систем приобретения знаний можно выделить ряд зарубежных и отечественных разработок: TEIRESIAS, OMIS, OPAL, АРИАДНА, АВТАНТЕСТ, Simer + Mir и т.д. [2-4, 6]. В данных системах, в качестве основных, реализованы следующие методы инженерии знаний: интервью и текстологические методы. При этом априори считается, что эксперт почти всегда выдает, а инженер по знаниям принимает и без серьезных искажений преобразует извлеченную информацию.

Из анализа различных источников [3-6] следует, что работа систем малоэффективна в решении проблемы потерь экспертной информации при ее извлечении и обработке. Кроме того, эти системы характеризуются:

- слабой проработкой методов извлечения и структурирования знаний в части их практической реализации в составе указанных систем;
- слабыми возможностями по выявлению «глубинных» знаний и последовательной реализацией процесса извлечения и обработки знаний;
- отсутствием обоснованного выбора методов извлечения и обработки экспертной информации;
- отсутствием обратной связи и проверки качества знаний, извлеченных (обрабатываемых) на промежуточных этапах исследуемого процесса;
- отсутствием комплексного учета различных методов, используемых при извлечении и последующей обработке знаний.
- жесткими ограничениями на предметную область;
- несовершенством интерфейса, в результате чего неподготовленные эксперты не способны овладеть системой в полной мере;
- отсутствием комплексного учета различных методов, используемых при извлечении и последующей обработке знаний.

Кроме этого, не учитываются особенности задач предметной области, в интересах решений которых используются методы прикладных наук (теории принятия решений, исследования операций, математического моделирования и других).

Исходя из выше сказанного, возникает необходимость в поиске и разработке методических и инструментальных средств, обеспечивающих, с одной стороны, автоматизацию рассматриваемого процесса, с другой – минимизацию отмеченных затрат при его реализации и являющихся общими для решения класса указанных задач.

В рассматриваемой методике извлечения и обработки экспертной информации предлагается использовать соответствие между классами задач и методами извлечения

знаний, а также соответствие между классами задач и методами обработки экспертной информации, и, наконец, методами извлечения данной информации и методами ее обработки.

Для обеспечения учета указанных соответствий при практической реализации процесса приобретения знаний в состав классической схемы процесса извлечения и обработки знаний предлагается включить в качестве обеспечивающих еще два этапа (рисунок 2):

– этап идентификации задачи, т. е. соотнесения ее с одним из классов задач предметной области;

– отбор соответствующих методов извлечения и обработки экспертной информации по критерию наибольшего соответствия их классам решаемых задач и друг другу.

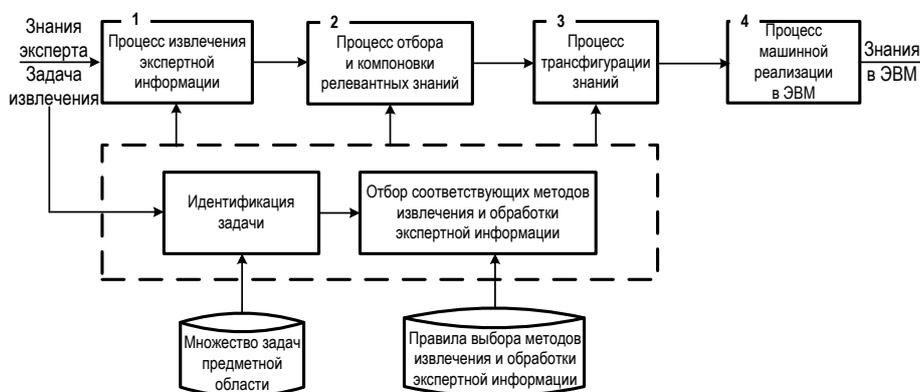


Рисунок 2 – Структурная схема предлагаемой методики извлечения и обработки знаний

**Этап идентификации задачи.** Каждой из рассматриваемых задач присваивается номер, характеризующий ее принадлежность к одному из классов задач, а также ее порядковый номер внутри этого класса. В результате анализа наименования поступившей задачи и перечня наименований задач принимается решение о ее принадлежности к одному из классов задач. Для решения задачи идентификации предлагается использовать номинальную шкалу, так как номинальные шкалы позволяют опознавать, различать, идентифицировать объект. В данном случае рассматривается задача «Определение действий конкурентов» из класса задач «Оценка рыночной обстановки».

**Этап отбора соответствующих методов извлечения и обработки экспертной информации.** После реализации первого этапа осуществляется выбор соответствующих методов извлечения и обработки экспертной информации с помощью нечетких продукционных правил типа «если..., то...» [6].

С целью установления соответствия между методами извлечения знаний и, например, решаемой задачей «Определение вероятных направлений развития рынка» был проведен опрос десяти экспертов по шкале веса важности каждого метода от 0 до 10 баллов применительно к пяти значениям термов лингвистической переменной «Соответствие метода извлечения знаний и класса решаемых задач» = {"не подходит", "мало подходит", "подходит", "хорошо подходит", "отлично подходит"}.

Результаты экспертного опроса о применимости метода анкетирования к извлечению знаний о решении указанной задачи представлены в таблице 1.

Таблица 1 – Результаты опроса экспертов

Терм	Эксперты									
	1	2	3	4	5	6	7	8	9	10
не подходит	1	2	1	3	2	1	3	1	2	1
мало подходит	3	4	4	5	5	5	4	3	4	5
подходит	5	6	7	6	7	7	6	5	6	7
хорошо подходит	9	10	10	8	10	10	9	8	10	9
отлично подходит	7	8	8	7	8	9	8	7	8	8

Из анализа данных таблицы видно, что опрошиваемые эксперты считают, что метод анкетирования «хорошо подходит» для извлечения знаний о решении задачи сбора, обработки и обобщения информации. Аналогично можно провести опрос о применимости других методов для извлечения знаний о порядке решения рассматриваемой задачи, а также и других задач управления. В результате на основе полученных данных была построена логико-лингвистическая шкала (ЛЛШ) соответствия между методами извлечения знаний и решаемой задачей, показанная на рисунке 3, в полной мере соответствующая классическому виду ее представления. Для построения ЛЛШ был использован метод статистической оценки функции принадлежности, реализованный средствами математического пакета Mathcad 13.

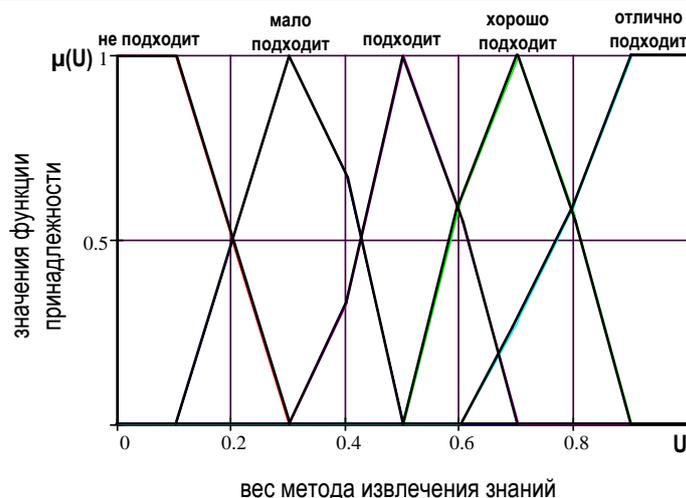


Рисунок 3 – Логико-лингвистическая шкала соответствия между методами извлечения знаний и решаемой задачей

Структурно указанное соответствие представлено на рисунке 4 на примере двух классов задач: сбора, обработки и обобщения информации и оценки обстановки. Аналогично можно выделить соответствие между задачами других классов и существующими методами извлечения знаний об их решении.

Наличие такого соответствия в распоряжении инженера по знаниям способствует повышению эффективности реализации процесса извлечения экспертной информации в целом, так как при этом упрощается выбор метода или методов, с помощью которого можно извлекать знания о решении конкретной задачи с наименьшими потерями знаний и временными затратами.

Учитывая неопределенный (нечёткий) характер выделенных связей, характеризующих соответствие классов задач и самих задач предметной области методам извлечения знаний об их решении, для формализованного представления указанных соответствий предполагается использовать математический аппарат нечётких множеств и нечёткий ситуационный логический вывод.

Ниже приводится пример формализованного описания в общем виде указанных соответствий и решения задачи выбора подходящего метода извлечения знаний о конкретной задаче на основе нечетких продукционных правил типа "если ..., то ..." [5, 6]:

П: если  $x$  есть  $Z$ , то ( $y$  есть  $M1$  и  $y$  есть  $M2$ ),

где  $x$  – входная переменная,  $y$  – переменная вывода,  $Z$  – задача оценки спроса на товар,  $M1$  – метод анкетирования,  $M2$  – метод анализа документов.

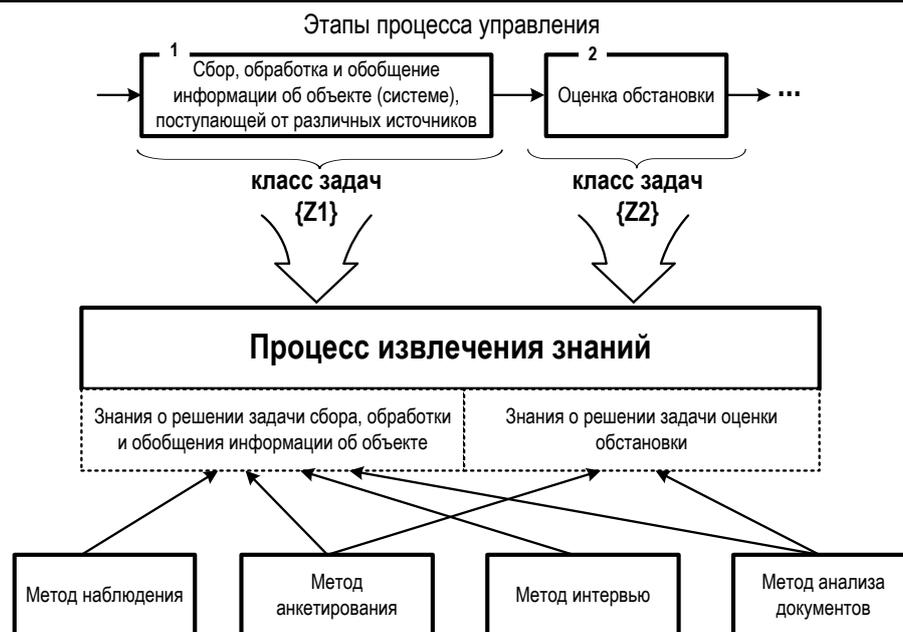


Рисунок 4 – Соответствие между задачами и существующими методами извлечения знаний об их решении

Следовательно, для извлечения знаний о порядке решения задачи управления «Оценки спроса на товар» следует использовать следующие методы извлечения экспертной информации: метод анкетирования (М1) и метод анализа документов (М2).

Аналогично было установлено соответствие между методами обработки знаний и решаемыми задачами, а также между методами извлечения знания и методами их обработки. В качестве таких методов были выбраны статистический и параметрический методы.

Таким образом, практическая реализация предложений по структурному дополнению классического процесса извлечения и обработки экспертной информации этапами «идентификация задачи» и «отбор соответствующих методов извлечения и обработки экспертной информации» позволит уменьшить потери извлекаемых и обрабатываемых знаний.

### Список литературы

1. Балашов О.В., Букачев Д.С. Подход к анализу эффективности работы банка с использованием нечетких когнитивных карт // Международный журнал информационных технологий и энергоэффективности. – 2017. – Т. 2, №4 (6). – С. 47-53.
2. Гаврилова Т. А., Хорошевский В. Ф. Базы знаний интеллектуальных систем. – СПб: Питер, 2000. – 383 с.
3. Круглов В. В., Длин М. И. Интеллектуальные информационные системы: компьютерная поддержка систем нечеткой логики и нечеткого вывода. – М: Изд-во Физико-математической литературы, 2002. – 256 с.

Балашов О.В., Букачев Д.С. Выбор методов извлечения и обработки экспертной информации для базы знаний систем поддержки принятия решений // Международный журнал информационных технологий и энергоэффективности. – 2018. – Т.3 №4(10) с. 28-35

---

4. Люггер Дж. Ф. Искусственный интеллект: стратегии и методы решения сложных проблем. М., Изд. дом «Вильямс», 2003. – 864 с.
5. Рыков В.В. Семиотическое моделирование процессов понимания // Информационные ресурсы России. – 2010. – № 6. С. 31 – 35.
6. Червинская К.Р., Щелкова О.Ю. Медицинская диагностика и инженерия знаний. – СПб: Ювента, 2002. – 624 с.

#### References

1. Balashov O.V., Bukachev D.S. Podhod k analizu ehffektivnosti raboty banka s ispol'zovaniem nechetkih kognitivnyh kart // Mezhdunarodnyj zhurnal informacionnyh tekhnologij i ehnergoehffektivnosti. – 2017. – Т. 2, №4 (6). – S. 47-53.
  2. Gavrilova T. A., Horoshevskij V. F. Bazy znaniy intellektual'nyh sistem. – SPb: Piter, 2000. – 383 s.
  3. Kruglov V. V., Dli M. I. Intellektual'nye informacionnye sistemy: komp'yuternaya podderzh-ka sistem nechetkoj logiki i nechetkogo vyvoda. – M: Izd-vo Fiziko-matematicheskoy literatury, 2002. – 256 s.
  4. Lyugger Dzh. F. Iskusstvennyj intellekt: strategii i metody resheniya slozhnyh problem. M., Izd. dom «Vil'yams», 2003. – 864 s.
  5. Rykov V.V. Semioticheskoe modelirovanie processov ponimaniya // Informacionnye resursy Rossii. – 2010. – № 6. S. 31 – 35.
  6. CHervinskaya K.R., SHCHelkova O.YU. Medicinskaya diagnostika i inzheneriya znaniy. – SPb: YUventa, 2002. – 624 s.
-