

ЗНАНИЕ-ОРИЕНТИРОВАННЫЕ ИНТЕЛЛЕКТУАЛЬНЫЕ МЕТОДЫ ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

Рассмотрены проблемы глобального доступа к данным и знаниям компании. Наличие автоматизированных объектов потребовало реализации интеллектуальных технологий, методов и моделей оценки качества их функционирования.

Предложен метод оптимизации доступа к данным, который ориентирован на интеграцию уже существующих локальных данных и предусматривает независимую поддержку и администрирование каждого из них. Метод ориентирован на детерминированные процессы, что ограничивает его применение.

Сформулированы особенности методов обработки данных, предложен расширенный знание-ориентированный метод нечеткого вывода TSK. Нечеткий метод TSK, который в отличие от существующих, дополнительно включает в систему уравнений функцию полезности и позволяет осуществлять интеграцию знаний в процедурах нечеткого логического вывода с учетом значимости правил. Экспериментом подтверждена эффективность подхода в знание-ориентированных технологиях.

Ключевые слова: федеративная база данных, распределенный запрос, неоднородные системы, локальные данные, база правил.

НОМЕНКЛАТУРА

EAI – enterprise application integration;
ECM – enterprise content management;
EDR – enterprise data replication;
EII – enterprise information integration;
ETL – extract transform loading;
TSK – метод Такаги-Сугэно-Канга;
БД – база данных;
СУБД – система управления базами данных;
ФД – федерализация данных;
 \tilde{F} – нечеткое пространство данных и знаний;
 $R^{(k)}$ – база правил;
 $f(x)$ – полином n -го порядка;
 μ_α – функция полезности;
 μ_{x_i} – функция принадлежности;
 \bar{x}_i – исходные данные;
 \bar{y}_i – параметр логического вывода;
 \bar{y}_k – значения управленческого воздействия;
 \bar{y} – выходной сигнал;
 x' – некоторый известный вектор;
 y' – искомый вектор;
 $\mu(x, y)$ – оператор нечеткого отношения Мамдани;
 y_0 – искомый четкий вектор;
 y_{0_i} – центр масс частного образа;
 S_i – площадь частного образа;
 U – множество узлов;
 S – каналы передачи данных;
 P – выполнение запросов;
 Q – множество запросов;
 W – множество обновлений;
 T_1 – время выполнения запроса;
 q_i – время обновления;
 D – детерминированный процесс;
 D_j – дисковое пространство j -го узла;
 F – нечеткое пространство состояний;
 L_i – объем i -го фрагмента данных;

n – количество фрагментов;
 P_m – вероятностный процесс;
 P – процессы;
 S_n – интеграция знаний;
 V – множество ограничений;
 A – интеграция данных;
 T – максимальное время выполнения запросов.

ВВЕДЕНИЕ

С каждым днем появляются новые технологии, что влечет за собой реорганизацию структур, неоднородных приложений и систем, оптимизацию их применения. Чтобы суметь обеспечить комфортный доступ к данным, требуются применение сложное знание – ориентированных технологий и технологии интеграции информации. Необходимо обеспечить также оптимизированный доступ к этой информации независимо от ее физического местоположения.

Интеграция информации должна происходить во всей распределенной среде. Чаще всего, проводить такого рода оптимизацию затратно для предприятий, поэтому применяют оптимизацию запросов по локальным местам. Сегодня в компаниях, приобретает все большую актуальность задачи с использованием информационных технологий.

Основной информацией, которая используется в небольших компаниях, является документооборот. Данные сохраняются в локальных базах и доступны лишь пользователям данной системы. Это приводит к тому, что доступ к данным и формирование отчетности может проводиться только в локальных системах учета или в бухгалтерском виде.

Работа направлена не только на выявления ошибок в автоматизированной системе микропроцессорной централизации, но также их последующая обработка, классификация и прогнозирование с помощью метода TSK. Целью исследования является интеграция знаний и данных, их совместное использование в системах, что позволит использовать интеллектуальное управление сложных систем и повысить качество управления.

1 ПОСТАНОВКА ЗАДАЧИ

Пусть существует множество процессов, для которых сформированы отчеты:

$$\{Pr_j\}, j \in J, \tag{1}$$

$$\{I_{jk}\}, k \in K. \tag{2}$$

Отчеты (2) характеризуется множеством показателей B , на процессах (1). Процессы (1) формируются в детерминированном, вероятностном, нечетком пространствах состояний. Процессы (1) характеризуются интеграцией динамических данных A и знаний S_n :

$$\forall Pr_j(\tau), Pr_j \in \{Pr_j\} A \cup Sh. \tag{3}$$

Структура предметной области предложена на рис. 1. В работе необходимо:

- выполнить содержательный анализ предметной области;
- рассмотреть существующие методы обработки информации;
- предложить алгоритмы формирования запросов;
- предложить знание-ориентированный расширенный метод нечеткой модели представления знаний на основе подходов TSK [5].

2 ОБЗОР ЛИТЕРАТУРЫ

Существуют различные методы оптимизации и распределения данных в системе [1–4]. Организованный доступ к данным осуществляется несколькими подходами [1]:

- способ консолидации;
- способ федерализации;
- сервисный способ;
- способ распространения данных.

Среди предложенных выше способов более глобальным является способ федерализации. Для реализации данного способа доступа в систему следует оптимизировать запрос [2]. Преимуществом применения такого доступа будет удобный доступ и поиск необходимой информации во всей системе [3, 4].

Знание – ориентированные методы [5–7] на основе правил продукций имеют такие свойства как, снижение уровня неопределенности процессов и объектов, это позволяет повысить скорость в задачах логического вывода [6]. Также существует недостаток это настройка функций принадлежности нечетких процессов, сложность приобретения знаний, интерпретация полученных знаний [7].

В данной работе предложен знание-ориентированный расширенный метод нечеткой модели представления знаний на основе TSK [5], а также сравнительный анализ расширенного метода TSK с методом Мамдани [6].

3 МАТЕРИАЛЫ И МЕТОДЫ

Для реализации доступа к неоднородным данным автоматизированных систем, используют следующие

методы [1]: консолидация, федерализация, распространение данных, сервисный подход.

Консолидация [1] – однонаправленный процесс, то есть данные из нескольких источников сливаются в Хранилище, но не распространяются из него обратно в распределенную систему. Процесс заполнения хранилища состоит из трех фаз – извлечение, преобразование, загрузка ETL (extract transformation loading). Во многих случаях именно ETL понимают под термином «интеграция данных». Технологией консолидации данных является управление содержанием корпорации ECM (enterprise content management). Большинство решений ECM направлены на консолидацию и управление неструктурированными данными, такими как документы, отчеты и web-страницы.

Федерализация: в федеративных БД физического перемещения данных не происходит: данные остаются у владельцев, доступ к ним осуществляется при необходимости (при выполнении запроса). Изначально федеративные БД предполагали создание в каждом из n узлов $n-1$ фрагментов кода, позволяющего обращаться к любому другому узлу [1].

При использовании медиатора создается общая модель данных. Медиатор – посредник, поддерживающий единый пользовательский интерфейс на основе глобального представления данных, содержащихся в источниках, а также поддержку отображения между глобальным и локальным представлениями данных. Пользовательский запрос, сформулированный в терминах единого интерфейса, декомпируется на множество подзапросов, адресованных к нужным локальным источникам данных. На основе результатов их обработки синтезируется полный ответ на запрос. Используются две разновидности архитектуры с посредником – globalas view и localas view.

Интеграция корпоративной информации ЕИ (enterprise information integration) – это пример технологии, которая поддерживает федеративный подход к интеграции данных. Изучение и профилирование первичных данных необходимые для федерализации, несильно отличаются от аналогичных процедур, требуемых для консолидации.

Распространение данных: приложения распространения данных осуществляют копирование данных из одного места в другое. Примерами технологий, поддерживающих распространение данных, являются интеграция корпоративных приложений EAI (enterprise application integration) и тиражирование корпоративных данных EDR (enterprise data replication). От федеративных БД этот способ отличает двустороннее распространение данных [1].

Сервисный подход: сервисно-ориентированная архитектура SOA (service oriented architecture), успешно применяемая при интеграции приложений, применима и при интеграции данных. Данные также остаются у владельцев и даже местонахождение данных неизвестно. При запро-

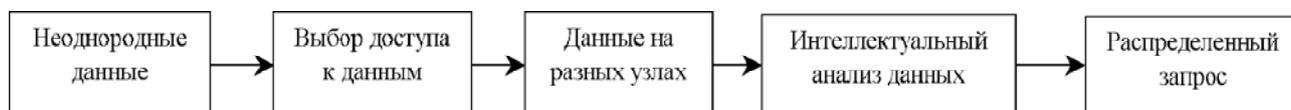


Рисунок 1 – Структура анализа предметной области

се происходит обращение к определенным сервисам, которые связаны с источниками, где находится информация и ее конкретный адрес. Интеграция данных объединяет информацию из нескольких источников таким образом, чтобы ее можно было показать клиенту в виде сервиса.

В базовой модели [2] оптимизация производится путем формирования самого запроса. В нашем случае оптимизация происходит на самих узлах. Для разрешения изменяющихся данных существующих воспользуемся методом федерализации информации [1], путем объединения всех имеющихся данных в единую виртуальную базу данных, хотя фактически местоположение данных не меняется.

Достоинства данного метода в том, что независимость данных приводит к минимальным изменениям, данных. Рассмотрим подходы к расширению существующих методов.

4 ЭКСПЕРИМЕНТЫ

Существующие методы реализации доступа [8, 9] не всегда реализуют требуемые временные ресурсы и оптимизацию требуемых вычислительных ресурсов. В основе данного метода лежит ФД. Под неоднородными системами следует понимать системы, реализованные в реляционной модели данных, но имеющие различные наборы объектов с отличными семантическими связями (рис. 2) [4].

Метод ориентирован на интеграцию уже существующих локальных данных и предусматривает независимую поддержку и администрирование каждого из них. Такая независимость хранилищ позволяет работать с системой даже при выходе из строя отдельных узлов сети или линий связи и не требует наличия знания об операциях, выполняющихся в других узлах сети.

Основой метода является также работа с метаданными каждого из существующих в сети хранилищ данных. Реализация метода сводится к формированию распределенной транзакции к множеству доступных на момент поступления запроса удаленных серверов БД, на основе анализа метаданных узлов сети.

Утверждение 1. Если система реализует функции доступа к данным за минимальное время, то осуществляем оптимизацию поиска методом половинного деления (дихотомии).

Справедливость утверждения 1 очевидна, что и определяет применение методов оптимизации.

В общем виде метод доступа к данным может быть представлен следующим образом:

Этап 1. Формирование неоднородных данных в системе.

Этап 2. Используя метод федерализации, инициируем доступ к неоднородным данным.

Этап 3. Оптимизация данных на основе метода дихотомии, формирование распределенного запроса $\{C_i\}, i \in I$.

Этап 4. Выполнение запросов $\{C_i\}, i \in I$.

Этап 5. Если $C_i = \text{false}$, то выполняются модификация БД и реализация этапов 1–4;

Этап 6. Для этапов 1–5 осуществляем минимизацию временных ресурсов $\{Pr_j(\tau)\} \rightarrow \min$;

Этап 7. Останов.

Так как в решаемой задаче на каждом узле системы хранятся эквивалентные по смыслу данные, то при выполнении транзакции будет происходить запрос данных со всех доступных интегрируемых системах с последующим объединением собранных данных в единое представление.

Учитывая что перспективные решения в задачах управления ориентированы на знание-ориентированные технологии, применим метод Такаги-Сугено-Канга. Для дальнейшего анализа и прогнозирования ошибок в системе следует применить расширение метода нечеткого управления TSK [5].

В рассмотренном модуле нечеткого управления используется база правил $R^{(k)}, k=1,2,\dots,n$. Модель представлена в виде:

$$\text{if } x_i \text{ is } A \text{ then } y = f(x), \quad (4)$$

применяемые в нем правила нечеткие только в части IF, тогда как в THEN неизвестные представлены полиномом:

$$R^{(1)}: \text{IF}(x_1 \text{ is } A_1^N \text{ AND } x_2 \text{ is } A_2^N \dots \text{AND } x_n \text{ is } A_n^N) \text{ THEN } y_N = f^{(1)}(x_1, x_2, \dots, x_n);$$

$$R^{(l)}: \text{IF}(x_1 \text{ is } A_1^N \text{ AND } x_2 \text{ is } A_2^N \dots \text{AND } x_n \text{ is } A_n^N) \text{ THEN } y_N = f^{(l)}(x_1, x_2, \dots, x_n), \quad (5)$$

– правила (25), (4) характеризуется дополнительно зависимостями [7]:

$$\mu_\alpha \left\{ \text{if } x_i \text{ is } A_i \text{ then } y = f(x_i) \right\}, \alpha \in A. \quad (6)$$

Недостатком существующего метода TSK [5] является отсутствие влияния значимости правил в виде нечет-

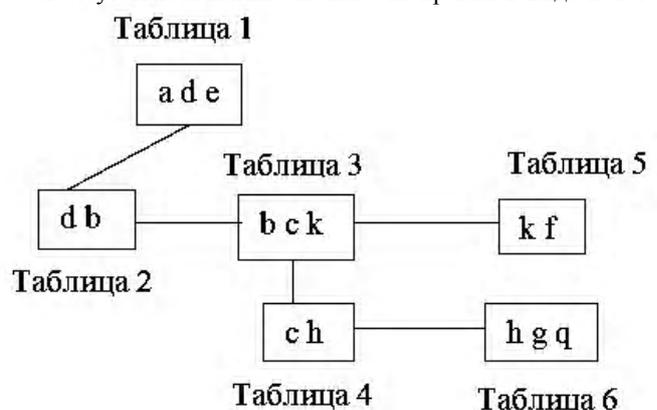
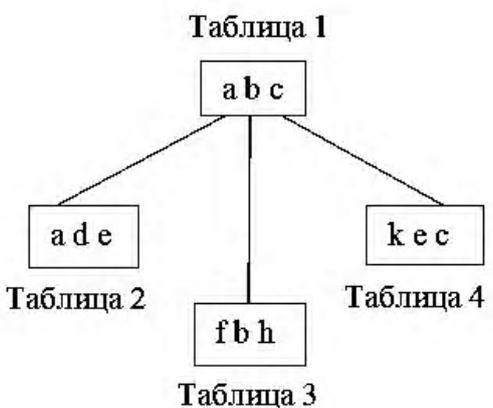


Рисунок 2 – Пример структурной схемы неоднородных БД

кой функции полезности правил (4)–(6) [6]. Предлагается метод, этапы которого представлены:

Этап 1. Формирование правил продукций TSK.

Этап 2. Формирование функции принадлежности μ_{x_i} .

Этап 3. Формирование функций полезности $\{\mu_{\alpha}\}, \mu_{\alpha} = [0,1]$.

Этап 4. Задание исходных данных об системе $\bar{x}_i, i \in I$.

Этап 5. Вычисление параметров логического вывода $\Omega, \bar{y}_j, j \in J$.

Этап 6. Вычисление значений управленческих воздействий $\{\bar{y}_k\}, k \in K$.

Этап 7. Прогон на тестовой выборке соответствия $\bar{x} \rightarrow \bar{y}$.

Этап 8. Если $\bar{x} \rightarrow \bar{y} = \text{true}$, то else $\bar{x} \rightarrow \bar{y} = \text{false}$ и выполнить этапы 1–7;

Этап 9. Останов.

Вычислительная сложность данного метода близка к квадратичной, что определяет перспективность подхода.

Функционирование системы в условиях неопределенности компонент предлагает применение знаний эксперта. Существующие методы и модели [6] Мамдани и TSK обладают достоинствами и недостатками, что ограничивает их реализацию на практике.

Рассмотрим пример модуля управления TSK при условии, что $\mu_{\alpha} = \mu_{\beta} = 1$:

$$R^{(1)}: \text{IF}(x_1 \text{ это малое AND } x_2 \text{ это среднее}) \text{ THEN } y_1 = 2 + 3x_1 - x_2,$$

$$R^{(2)}: \text{IF}(x_1 \text{ это малое AND } x_2 \text{ это среднее}) \text{ THEN } y_2 = -x_1 + 4x_2. \quad (7)$$

Определим его выходной сигнал \bar{y} для $\bar{x}_1 = 2$ и $\bar{x}_2 = 3$ при функции принадлежности (рис. 3)

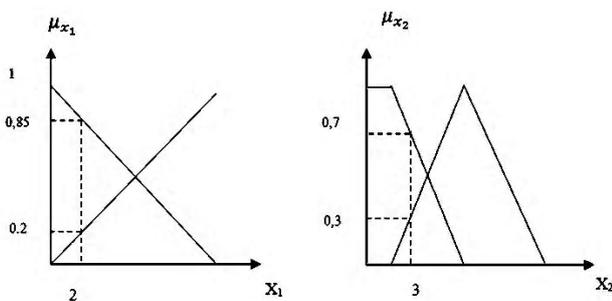
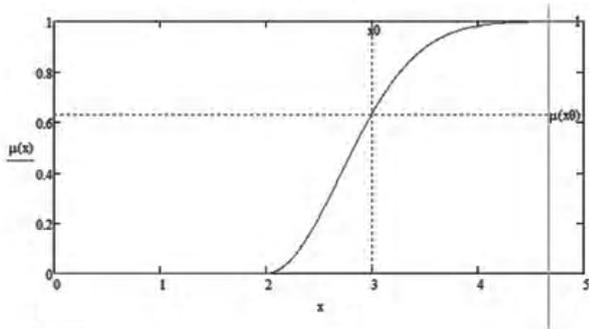
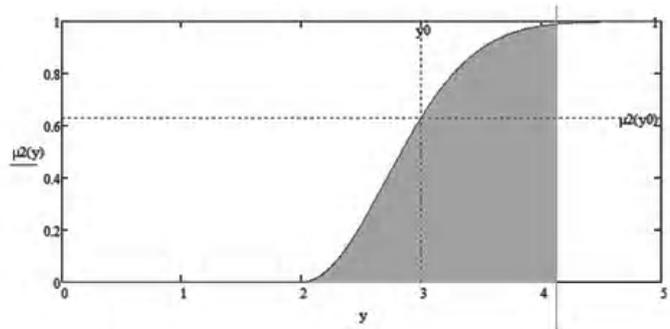


Рисунок 3 – Структура вычислений



а



б

Рисунок 4 – Графический метод: а – метод TSK, б – метод Мамдани

Выполнив вычисления значения \bar{y} на выходе модуля нечеткого управления по формуле (8):

$$\bar{y} = \frac{w^1 \bar{y}_1 + w^2 \bar{y}_2}{w^1 + w^2}, \quad (8)$$

получим $\bar{y} = 7,6$, что является удовлетворительным.

Выполним сравнительный анализ нечетких расширенных методом TSK и методов Мамдани [6].

Тогда метод Мамдани представлен в виде:

$$\left\{ \text{if } x_i \text{ is } \mu(x_i) \text{ then } y \text{ is } \mu(y) \right\}. \quad (9)$$

Решение (9) находим как

$$y' = \vee x' \wedge \mu(x, y), \quad (10)$$

$$\mu(x, y) = \wedge(\mu(x_i), \mu(y)). \quad (11)$$

Построим правила продукций вида (11) для задачи выявления, локализации и устранения рисков при движении мобильного объекта: x – скорость объекта

$$\begin{aligned} \text{if } x_1 \text{ is "малое"} &= \text{then } y \text{ is "малое"}, \\ \text{if } x_1 \text{ is "среднее"} &= \text{then } y \text{ is "среднее"}, \\ \text{if } x_1 \text{ is "большое"} &= \text{then } y \text{ is "большое"}. \end{aligned} \quad (12)$$

Определим функции принадлежности в виде гауссианов:

$$\text{"большое"} - \mu(x) = 1 - e^{-k_3(x-b)^2};$$

$$\text{"малое"} - \mu(x) = 1 - e^{-k_1 x^2};$$

$$\text{"среднее"} - \mu(x) = 1 - e^{-k_2(x-a)^2}.$$

Решим задачу на первом правиле из (12) графическим метод [5] (рис. 4)

Задав x' , находимо решение y' . После дефаззификации [5] на основе центра масс

$$y_0 = \frac{\sum_i y_{i0} \cdot S_i}{\sum_i S_i},$$

$$y_0 = \frac{3,5 \cdot 0,25 + 5,5 \cdot 3}{3,25} = \frac{0,9 + 16}{3,25} = \frac{17}{3,25} = 5,85.$$

Исходя из предметной области, для перемещения мобильного объекта определены искомые величины метода TSK и метода Мамдани (таблица 1).

Таблица 1 – Результаты сравнения методов

Метод	Результат
Метод TSK	7,6
Метод Мамдани	5,85

Искомое решение, по данным экспертных оценок, находится в интервале $5,9 \leq y_0 \leq 8,2$. Как следует из табл. 1 и данных эксперимента, метод TSK является более точным, позволяет получать устойчивые решения. Это и определяет преимущества метода.

5 РЕЗУЛЬТАТЫ

Рассматриваемая БД представляет собой множество фрагментов данных в системе

$$F = \{f_i, i = 1 \dots n\}, \quad (13)$$

которые распределены на множестве узлов

$$U = \{u_j, j = 1 \dots m\}, \quad (14)$$

соединенных каналами передачи данных

$$S = \{s_k, k = 1 \dots h\}. \quad (15)$$

На следующем этапе функционирует приложение, которые инициируют выполнение запросов

$$P = \{p_s, s = 1 \dots r\}. \quad (16)$$

В процессе функционирования БД порождается множество запросов

$$Q = \{q_d, d = 1 \dots n_d\}, \quad (17)$$

и множество обновлений

$$W = \{w_e, e = 1 \dots n_w\}, \quad (18)$$

где $e \in w$ для элементов, которых определены функции

$$T_1(q_d, A), \quad (19)$$

– время выполнения запроса

$$q_i(d = 1 \dots n_d) \text{ и } T_2(w_e, A), \quad (20)$$

– время обновления

$$w_e(e = 1 \dots n_w). \quad (21)$$

Основной задачей оптимизации БД является необходимость найти систему распределения данных с учетом (13)–(21), при которой суммарное среднее время выполнения запросов и распространения обновлений, порожденных функционированием системы, минимально. Таким образом, оптимизация запросов реализуется в виде

$$\langle F, Q, W, T \rangle \xrightarrow{V} \min. \quad (22)$$

При вычислении критерия оптимальности необходимо учитывать следующие ограничения:

– в БД должны присутствовать хотя бы одна копия данных:

$$\sum_{j=1}^n A_{ij} \geq 1, (i = 1, n), \quad (23)$$

– суммарный объем данных, хранящихся на узле, не должен превышать общее дисковое пространство данного узла:

$$\sum_{j=1}^n L_i \times A_{ij} \leq D_j, (j = 1, m), \quad (24)$$

– максимальное время выполнения запросов не должно превышать заданного предельного значения $T \leq T^*$. Тогда

$$V = \sum_{i=1}^m A_{ij} \cup \sum_{j=1}^n L_i \times A_j. \quad (25)$$

Оптимизация запросов в работе, реализована с использованием расширения метода дихотомии (половинного деления) [10]. Применение метода позволило сократить время реализации запросов до 10%, что является удовлетворительным.

6 ОБСУЖДЕНИЕ

Усовершенствованный метод доступа является динамичным, но СУБД не учитывает особенности реализации распределенной обработки [3]. При таком подходе проблема метода оптимизации распределения данных рассматривается как комплексная проблема.

Структурная схема осуществления такого запроса представлена на (рис. 5). На основе поступающего запроса проводится анализ доступных на данный момент серверов и формируется запрос к каждому активному серверу. После чего осуществляется объединение полученных наборов данных в результирующее множество.

Особенностью формирования распределенного запроса (рис. 5) является требование доступности активных



Рисунок 5 – Схема формирования распределенного запроса



Рисунок 6 – Схема обработки распределенного запроса

серверов, что в ряде случаев вызывает сложности. Экспериментом подтверждено, что вычислительная сложность процессов (рис. 5) близка к линейной – $O(n)$. Конвертирование названий атрибутов и условия выборки происходит в соответствии с таблицами метаданных, расположенными на используемом сервере сети, и может быть представлена в виде (рис. 6).

Недостатком метода TSK [5–7] является его ориентация на непрерывные процессы и в меньшей мере учитывает неоднородность входных данных и знаний. В связи с этим, усовершенствование метода TSK на основе функции полезности является перспективным, однако требует дополнительных исследований. Экспериментом подтверждено (табл. 1), что реализация метода отличается большей универсальностью по сравнению с методом Мамдани [8, 9].

ВЫВОДЫ

Разработка знание-ориентированных интеллектуальных методов и моделей анализа сложных объектов является важной составляющей технологических процессов в производственных системах, которые функционируют в условиях неопределенности. Знание-ориентированные методы направлены на моделирование и обработку детерминированных, вероятностных и нечетких знаний, как фактор повышения качества систем.

В результате проведенных исследований получены научные и практически значимые результаты, которые

позволяют сократить время их реализации при одновременном повышении достоверности в прятках решений до 12%, что является удовлетворительным.

1. Выполнен содержательный анализ предметной области, определены преимущества и недостатки существующих решений.

2. Определены особенности методов обработки информации на основе оптимизации запросов, сделан вывод об их применимости в предметных областях, что определяет практическую значимость полученных результатов.

3. Получил дальнейшее развитие нечеткий метод Такаги-Сугено-Канга, который в отличие от существующих, дополнительно включает в систему уравнений функцию полезности на интервале лингвистических термов $[0, 1]$, что позволяет осуществлять интеграцию знаний в процедурах нечеткого логического вывода Такаги-Сугено-Канга с учетом значимости правил и повышает достоверность решений.

4. Перспективой дальнейших исследований является адаптация исследований в предметных областях проектирования и внедрение результатов в производственных системах.

БЛАГОДАРНОСТИ

Работа выполнена в рамках государственной научно-исследовательской темы Харьковского национального университета радиотехники «Нейро-фаззи системы для текущей кластеризации и классификации последо-

вательностей данных в условиях их искривления отсутствующими и аномальными» (номер гос. регистрации 0113U000361), где авторами предложены новые методы на основе развития метода Такаги-Сугено-Канга для оптимизации запросов и принятия решений на знаниях.

СПИСОК ЛІТЕРАТУРИ

1. Архипенков С. Аналитические системы на базе Oracle Express OLAP / С. Архипенков. – М. : Диалог-МИФИ, 2000. – 287 с.
2. Data integration in the distributed information systems // International Journal of Advanced Studies. – 2012. – №1. [Электронный ресурс]. – Режим доступа <http://ijournal-as.com/issues/2012/1/bisterfeld.pdf>
3. Кузьмина С. П. Базы данных [Текст] / С. П. Кузьмина. – С.Пб. :СПбГИЭУ, 2006. –189 с.
4. Райордан Р. Основы реляционных баз данных / Р. Райордан. – М. : Издательско-торговый дом «Русская редакция», 2001. – 384 с.
5. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилинский, Л. Рутковский. – М. : Горячая линия – Телеком, 2006. – 452 с.

6. Интеллектуальное управление технологическими процессами [Текст]: монография научное издание / [Бодянский Е. В., Кучеренко Е. И., Михальов А. И. и др.]. – Днепропетровск : Национальная металлургическая академия Украины, 2013. – 213 с.
7. Tsoukalas L. H. Fuzzy and Neural Approaches in Engineering / L. H. Tsoukalas, R. E. Uhrig. – New York : John Wiley & Sons, Inc, 1997. – 587 p.
8. Ходак М. В. Информационная технология автоматизированного управления мобильными объектами / М. В. Ходак, Е. И. Кучеренко // Праці VII міжнародної школи-семинару «Теорія прийняття рішень». – Ужгород, УжНУ, 2014. – С. 261–262
9. Ходак М. В. Применение нечеткой логики в моделях процессов сложных систем / М. В. Ходак, Е. И. Кучеренко // Материалы 19-го Международного молодежного форума «Радиоэлектроника и молодежь в XXI веке» Том 6. – Харьков, ХНУ-РЭ, 2015. – С 62–63.
10. Метод дихотомии или метод половинного деления [Электронный ресурс]. – Режим доступа – <http://bpascal.ru/download/desc/319.php>

Статья поступила в редакцию 27.07.2015
После доработки 11.08.2015.

Кучеренко С. І.¹, Ходак М. В.²

¹Д-р технічних наук, професор Харківського національного університету радіоелектроніки, Харків, Україна

²Аспірантка Харківського національного університету радіоелектроніки, Харків, Україна

ЗНАННЯ-ОРІЄНТОВАНІ ІНТЕЛЕКТУАЛЬНІ МЕТОДИ ПРИЙНЯТТЯ РІШЕНЬ ЗА УМОВ НЕВИЗНАЧЕНОСТІ

Розглянуто проблеми глобального доступу до даних і знань компанії. Наявність автоматизованих об'єктів зажадало реалізації інтелектуальних технологій, методів і моделей оцінки якості їх функціонування.

Запропоновано метод оптимізації доступу до даних, який орієнтований на інтеграцію вже існуючих локальних даних і передбачає незалежну підтримку та адміністрування кожного з них. Метод орієнтований на детерміновані процеси, що обмежує його застосування.

Сформульовано особливості методів обробки даних, запропонований розширений знання-орієнтований метод нечіткого виводу TSK. Нечіткий метод TSK, який на відміну від існуючих, додатково включає в систему рівнянь функцію корисності і дозволяє здійснювати інтеграцію знань в процедурах нечіткого логічного висновку з урахуванням значущості правил. Експериментом доведено ефективність підходу в знання-орієнтованих технологіях.

Ключові слова: федеративна база даних, розподілений запит, неоднорідні системи, локальні дані, база правил.

Kucherenko Ye. I.¹, Khodak M. V.²

¹Dr. Sc., prof., Professor of Department of Artificial Intelligence, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

²Postgraduate student of Department of Artificial Intelligence, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

KNOWLEDGE-ORIENTED INTELLECTUAL METHODS DECISION UNDER UNCERTAINTY

The problems of global access to data and knowledge of the company. Availability of automated facilities require the implementation of intelligent technologies, methods and models for assessing the quality of their functioning.

A method for optimizing data access, which is focused on the integration of existing local data, and provides independent support and administration of each of them. The method is focused on a deterministic process, which limits its application.

Formulated especially data processing techniques, to offer enhanced knowledge-oriented method of fuzzy inference TSK. Fuzzy method TSK, which, unlike the existing ones, further includes a system of equations of the utility function and enables the integration of knowledge in the procedures of fuzzy inference based on the importance of rules. The experiment confirmed the effectiveness of the approach to knowledge-oriented technologies.

Keywords: federated database, distributed query, inhomogeneous systems, local data, rule base.

REFERENCES

1. Arhipenkov S. Analytical systems based on Oracle Express OLAP. Moscow, Dialogue-MIFI, 2000, 287 p.
2. Data integration in the distributed information systems [electronic resource]. Access mode <http://ijournal-as.com/issues/2012/1/bisterfeld.pdf>
3. Kuzmina S. P. Databases. Sankt-Peterburg, SPbGIEY, 2006, 189 p.
4. Riordan R. Relational database systems. Moscow, Publishing trade house «Russian edition», 2001, 384p.
5. Rutkovska D., Rutkovski L., Pilinski M. Neural network, genetic algorithms and fuzzy systems. Moscow, Hotline-Telecom, 2006, 452 p.
6. Bodyanskiy E. V., Kucherenko E. I., Mihalev A. I., Filatov V. A., Gasik M. M., Kucin V. S. Intelligent management technological

- process monograph-scientific publication. Dnipropetrovsk, National metallurgical academy of Ukrain, 2013, 213 p.
7. Tsoukalas L. H., Uhrig R. E. Fuzzy and Neural Approaches in Engineering. New York, John Wiley & Sons, Inc, 1997, 587 p.
8. Khodak M. V., Kucherenko E. I. Information technology computer-aided management of mobile objects, *Proceedings of the VII international school-seminar «Decision theory»*. Uzhgorod, Uzhgorod National University, 2014, pp. 261–262.
9. Khodak M. V., Kucherenko E. I. The use of fuzzy logic in the process models of complex systems, *Proceedings of the 19th International Youth Forum «Radioelectronics and Youth in XXI century»*, Volume 6. Kharkiv, KNURE, 2015, pp. 62–63
10. The method of dichotomy or bisection method [electronic resource]. Access mode <http://bpascal.ru/download/desc/319.php>