

УДК 621.324

С.Ф. Чалый, А.А. Синяков

*Харьковский национальный университет радиоэлектроники, Харьков*

## **РАЗРАБОТКА МЕТОДА ПОСТРОЕНИЯ НЕЧЕТКИХ ЗАПРОСОВ К РЕЛЯЦИОННЫМ БАЗАМ ДАННЫХ ДЛЯ ЗАДАЧ ПОИСКА ПЕРСОНАЛЬНОЙ ИНФОРМАЦИИ**

*В статье рассмотрена проблема построения нечетких запросов к реляционным базам данных для задач поиска персональной информации. Разработан метод построения нечетких запросов к реляционным базам данных для задач поиска персональной информации. Разработанный метод позволяет строить нечеткие запросы к реляционным базам данных, содержащим четкую формализованную информацию.*

**Ключевые слова:** *реляционная база данных, нечеткий запрос, нечеткий лингвистический интерфейс.*

### **Введение**

**Постановка проблемы.** На сегодняшний день значительная часть информационных систем строится с использованием реляционных баз данных. При работе с реляционными базами данных традиционно используется четкая, числовая информация, соответственно, четкими должны быть и параметры запросов к реляционным базам данных.

В то же время существуют прикладные задачи, в которых как экспертам, так и пользователям необходи-

мо оперировать качественными характеристиками и нечеткими методами. К ним относятся задачи поиска персональной информации, в частности в области страхования. В данной области возникает необходимость поиска информации с использованием качественных критериев и нечетко сформулированных условий. Данные, не удовлетворяющие какому-то одному критерию, могут быть выбраны из базы данных, если они имеют соответствующие показатели по другим критериям. При этом использование обычных (четких) запросов к базе данных обычно не дает должных результатов.

Также важным является вопрос эффективного использования накопленных информационных ресурсов. Существующие базы данных могли разрабатываться 5 – 10 лет назад, накопленные ресурсы представляют значительную ценность, но из-за устаревших средств поиска и обработки информации их использование может быть затруднено. Перевод существующих баз данных на новую технико-программную базу зачастую является дорогостоящим проектом, поэтому создание средств поиска информации, позволяющих работать с такими базами данных на уровне современных технологий и не затрагивающих собственно сами базы является экономически оправданной стратегией повышения эффективности использования имеющихся информационных ресурсов.

Таким образом, возникает проблема поиска персональной информации в реляционных базах данных на основе нечетко сформулированных параметров поиска. Исследование данной проблемы предполагает выполнение анализа существующих публикаций в данной области.

**Анализ исследований и публикаций.** Вопросы построения нечетких запросов к реляционным базам данных рассмотрены в ряде работ: для анализа нулевых значений (в частности, случаев «полной неизвестности» атрибута было предложено несколько подходов работы [1 – 3]), в работе [4] рассмотрен случай не полностью известных значений. В статье [6] предложен статистический подход к моделированию неопределенной информации в базе данных. В статье [5] развит метод работы с расплывчатыми запросами (представленных с помощью нечетких множеств), обращенных к точным данным. В то же время вопросы построения нечетких запросов к реляционным базам данных в этих работах рассмотрены недостаточно.

**Постановка задачи.** Исходя из рассмотренной проблемы, актуальной является задача формализации построения нечетких запросов к реляционным базам данных для решения класса задач поиска персональной информации. Исходными данными являются нечеткие переменные для построения запросов к реляционным базам данных. Требуется получить метод построения нечетких запросов к реляционным базам данных для задач поиска персональной информации. Метод должен обеспечивать получение необходимой информации в минимальный срок с минимальными временными затратами и минимальными затратами ресурсов.

### Метод построения нечетких запросов

Метод построения нечетких запросов к реляционным базам данных для задач поиска персональной информации включает в себя следующие этапы:

- формализация понятий пользователя;
- поиск информации;
- анализ результатов поиска.

Этап формализации понятий пользователя позволяет ввести новый поисковый признак и множество его значений. В соответствии с Заде [7], под смыс-

лом некоторого понятия понимается его объем, то есть множество реальных объектов с указанием степени их принадлежности к понятию. Другими словами, задание смысла понятия эквивалентно заданию его функции принадлежности. При этом в качестве универсального множества выступает множество значений некоторого атрибута базы данных (домен).

Существует много различных методов задания функций принадлежности. Выбор того или иного метода зависит от задачи, существующей ситуации (например, наличия экспертов) и других параметров. Отметим, что при построении функций принадлежности мы можем построить оптимальное множество значений признака с точки зрения пользователя – множество, в рамках которого неопределенность при формировании запроса будет минимальна.

Итак, результат этапа формализации понятий пользователя определяет набор значений признака с указанием функций принадлежности, им соответствующих. Эти функции принадлежности определены на доменах соответствующих атрибутов базы данных.

Этап поиска информации по нечетким запросам реализует следующий алгоритм поиска информации по нечетким понятиям пользователя.

1. Выбирается первая запись в базе данных.
2. Берется первый поисковый признак, сформулированный в запросе, и его значение.
3. По признаку выбирается соответствующий атрибут базы данных.
4. Вычисляется значение функции принадлежности, соответствующей значению признака, в точке, соответствующей значению анализируемого атрибута в анализируемой записи базы данных. Полученное значение функции принадлежности запоминается.
5. Берется следующий поисковый признак. Шаги 3 и 4 повторяются до тех пор, пока не окончатся поисковые признаки. В результате мы имеем набор значений функций принадлежности значений атрибутов анализируемой записи нечеткому запросу.
6. На базе полученного в результате выполнения шага 5 набора значений вычисляется обобщенная оценка принадлежности анализируемой записи запросу. В частности, это может быть минимальное значение из указанного набора, произведение элементов набора и т.п. Полученное обобщенное значение запоминается в рабочем поле базы данных.
7. Выбирается следующая запись в базе данных и повторяется шаг 2. Повторение происходит до тех пор, пока не переберутся все записи в базе данных.

Таким образом, результат поиска по нечеткому запросу – упорядочивание записей в базе данных по степени их соответствия данному запросу от 1 (полное соответствие) до 0 (полное несоответствие).

Этап анализа результатов поиска предназначен для уточнения результатов поиска информации. Допуская неопределенность на входе, мы должны предоставить средства ее «снятия» для конкретного пользователя на «выходе». Функции принадлежности, построенные в блоке формализации, отражают либо обобщенное мне-

ние экспертов, либо мнение какого-либо эксперта в зависимости от выбранного метода построения функции принадлежности. Если мнение пользователя и мнение эксперта (экспертов) не совпадают, необходимо предоставить механизм "индивидуализации" функций принадлежности. Если пользователь не удовлетворен результатом поиска, он может сформулировать коррекцию запроса по конкретному признаку в виде направления модификации и модификатора. Направление модификации указывает полюс признака, модификатор – его "силу". По направлению модификации вычисляется направление сдвига функций принадлежности используемых понятий, по модификатору – шаг сдвига. Шаг сдвига может вычисляться разными способами, и этот вопрос требует отдельного исследования. Также шаг зависит от количества объектов, находящихся между существующей точкой и концом универсума. Если таких объектов много – шаг выбирается достаточно большим, если мало – достаточно маленьким. Для полученных в результате модификации функций принадлежности снова повторяется этап поиска информации. Таким образом, пользователь может "уточнять" результаты поиска информации.

Метод построения нечетких запросов к реляционным базам данных для задач поиска персональной информации реализован с помощью нечеткого лингвистического интерфейса (НЛИ). Нечеткий лингвистический интерфейс (НЛИ) имеет 3 блока:

- блок формализации понятий пользователя;
- блок поиска информации;
- блок анализа результатов поиска.

НЛИ позволяет:

- вводить пользователю понятия, не содержащиеся в описаниях объектов базы данных;
- определять "смысл" введенных понятий как нечетких подмножеств доменов атрибутов базы данных;
- осуществлять поиск информации по введенным понятиям;
- уточнять результаты поиска информации за счет указания направления модификации и модификатора смысла используемых понятий.

При применении НЛИ пользователь получает "лингвистическую оболочку" базы данных. Наличие

различных классов пользователей, имеющих различные функции принадлежности, позволяет создавать и хранить несколько таких оболочек одной и той же базе данных, или несколько различных "взглядов" на одну и ту же информацию. Этот подход позволит значительно повысить эффективность использования существующих баз данных.

## Выводы

Получен метод построения нечетких запросов к реляционным базам данных для задач поиска персональной информации. Разработанный метод включает в себя этапы: формализации понятий пользователя; поиска информации; анализа результатов поиска.

Метод позволяет производить поиск персональной информации в реляционных базах данных с использованием качественных критериев и нечетко сформулированных условий. Разработанный метод реализован с помощью НЛИ.

## Список литературы

1. Biskup J. A formal approach to null values in database relations / J. Biskup // Workshop: Formal bases for data bases. CERT-DERI. Toulouse. Proc. Edited by H. Gallaire et J.M. Nicolas. – Dec. 12-14, 1979.
2. Codd E.F. Extending the database relational model to capture more meaning / E.F. Codd // ACM Trans. Database Syst. – 1979. – № 4. – P. 397-434.
3. Grant J. Null values in relational data base / J. Grant // Information Processing Letters. – 1977. – № 5. – P. 156-157.
4. Grant J. Partial values in a tabular database / J. Grant // Information Processing Letters. – 1979. – № 2. – P. 97-99.
5. Tahani V. A conceptual framework of fuzzy query processing – A step forward very intelligent database systems / V. Tahani // Information – Processing & Management. – 1977. – № 13. – P. 289-303.
6. Леоленков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH / А.В. Леоленков. – СПб., 2003. – 340 с.
7. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений / Л. Заде. – М.: Мир, 1976. – 260 с.

Поступила в редколлегию 11.11.2008

**Рецензент:** д-р техн. наук, проф. Э.И. Кучеренко, Харьковский национальный университет радиоэлектроники, Харьков.

## РОЗРОБКА МЕТОДУ ПОБУДОВИ НЕЧІТКИХ ЗАПИТІВ ДО РЕЛЯЦІЙНИХ БАЗ ДАНИХ ДЛЯ ЗАДАЧ ПОШУКУ ПЕРСОНАЛЬНОЇ ІНФОРМАЦІЇ

С.Ф. Чалий, О.А. Синяков

*У статті розглянута проблема побудови нечітких запитів до реляційних баз даних для завдань пошуку персональної інформації. Розроблений метод побудови нечітких запитів до реляційних баз даних для завдань пошуку персональної інформації. Даний метод дозволяє будувати нечіткі запити до реляційних баз даних, що містять чітку формалізовану інформацію.*

**Ключові слова:** реляційна база даних, нечіткий запит, нечіткий лингвістичний інтерфейс.

## DEVELOPMENT OF METHOD OF CONSTRUCTION OF FUZZY QUERIES TO RELATIONAL DATABASES FOR TASKS OF PERSONAL INFORMATION RETRIEVAL

S.F. Chaliy, A.A. Sinyakov

*In the article the problem of construction of fuzzy queries is considered to the relational databases for the tasks of personal information retrieval. The analysis of existent researches and publications is conducted in area of construction of fuzzy queries to the relational databases. The method of construction of fuzzy queries is developed to the relational databases for the tasks of personal information retrieval. The developed method allows to build fuzzy queries to the relational databases, to containing clear formalized information.*

**Keywords:** relational database, fuzzy query, fuzzy linguistic interface.