

УДК 004.9

Г. Ч. Набибекова, зав. отд., e-mail: gulnarara58@mail.ru,  
Институт Информационных Технологий НАН Азербайджана, Баку, Азербайджан

## Об одном методе фаззификации атрибутов хранилища данных в системах поддержки принятия решений в сфере внешней политики

*Представлен процесс фаззификации атрибутов хранилища данных системы поддержки принятия решений в сфере внешней политики. С этой целью выбранный алгоритм CLARANS был применен для кластеризации данных хранилища. С помощью медоидов, полученных в результате кластеризации, были выведены значения функции принадлежности. Рассматривается случай, когда число терм-множеств, создаваемых на различных измерениях куба, различно.*

**Ключевые слова:** OLAP-куб, фаззификация, нечеткое множество, лингвистическая переменная, терм-множество, кластеризация, хранилище данных, атрибут, медоид

G. Ch. Habibakova

## A Fuzzification Method of Attributes of Data Warehouse of Decision Support System in the Foreign Policy

*The paper presents a fuzzification process of attributes of data warehouse of decision support system in the foreign policy. To this end, the chosen algorithm CLARANS was applied for clustering of data warehouse. The values of membership function were derived with medoids, which were obtained after clustering. The paper deals with the case, where the number of term-sets generated on different dimensions are various.*

**Keywords:** OLAP-cube, fuzzification, fuzzy-set, linguistic variable, term-set, clustering, data warehouse, attribute, medoid

### Введение

В последнее время руководители учреждений стали интересоваться хранилищем данных (ХД) как важнейшим инструментом управления, входящим в число наиболее значимых элементов инфраструктуры предприятия [1].

Являясь основным элементом систем поддержки принятия решений (СППР), ХД представляет собой единый и достаточно большой репозиторий данных, используемых для принятия решений. Согласно Б. Инмону [2], ХД — это "предметно-ориентированное, привязанное ко времени и неизменяемое собрание данных для поддержки процесса принятия управленческих решений". ХД строят путем консолидации данных из множества гетерогенных источников данных, которые преобразованы в единый формат и содержат данные как в детализированном, так и в агрегированном виде. Агрегирование данных многомерного ХД проводят по атрибутам (измерениям) для облегчения принятия

решений. Детализированные и агрегированные данные хранятся в многомерном OLAP-кубе, что обеспечивает быстрый ответ пользователям, осуществляющим запрос в ХД.

Инструменты OLAP, который является ключевым элементом ХД, обеспечивают различные операции, такие как свертка, развертка, сечение, срез, ранжирование, сводные таблицы. Обычно эти операции проводят над точными данными, и в результате их действия пользователи получают ответ в числовом выражении. Однако в процессе принятия решений можно столкнуться с множеством нечетких задач, вследствие чего в запросах к базам данных (БД), которые пытается формулировать исследователь, часто могут присутствовать неточности и неопределенности. Кроме того, для лиц, принимающих решения, важен и полезен бывает не сам результат запроса, а то, является ли этот результат *хорошим, средним или плохим*, т. е. им важна качественная сторона результата [3, 4]. Все это требует, чтобы СППР, использующая нечеткие и неопределенные

рассуждения, для оказания помощи руководителям в принятии решений поддерживала соответствующую технологию представления знаний [4].

### I. Поликубическая OLAP-модель с нечетким OLAP-кубом

Эта проблема является актуальной и при решении аналогичных задач в сфере внешней политики (ВП).

Отметим, что в основе ХД, спроектированного для сферы ВП, лежит информация о зарубежных поездках сотрудников государственного органа (ГО) [5].

Для более эффективного принятия управленческих решений аналитикам необходимо иметь в своем распоряжении целостную картину развития ГО по различным направлениям. Поэтому разрабатываемая СППР в сфере ВП носит многоцелевой характер, т. е. с ее помощью возможно решение задачи управления в различных аспектах — управление финансовыми, кадровыми, техническими ресурсами и т. д. С этой целью из ХД выделяются витрины данных (ВД), ориентированные на решение подзадач, сопутствующих основной задаче. Вследствие этого, кроме гиперкуба (или вместо гиперкуба), в котором все показатели определяются одним и тем же набором измерений, по каждой подзадаче в соответствии с ее запросами строится свой OLAP-куб. Таким образом, на основании одного ХД можно построить несколько OLAP-кубов. В первом случае OLAP-модель называется гиперкубической, а во втором случае — поликубической [6].

На рис. 1 представлена поликубическая OLAP-модель, в которой из ХД выделены три ВД.

В качестве примера создания нечеткого OLAP-куба рассмотрим подзадачу управления кадровыми ресурсами, включающую такие вопросы, как анализ кадровых ресурсов, их размещение, обучение, планирование продвижения сотрудников и т. д. При решении такой подзадачи для лиц, принимающих решение, при формировании запроса к БД часто бывает важен не конкретно возраст сотрудника, а то, является ли он *молодым, средним* или *пожилым*, либо, например, не конкретно стаж работы, а является он *малым, средним* или *продолжительным*. Кроме того, его может интересовать не точная дата мероприятия, в котором сотрудник принимал участие, а то, когда произошло это событие — *давно* или *недавно*. В этом случае ХД должно обеспечивать ответы на запросы, использующие лингвистические термины, которые являются терм-множествами соответствующей лингвистической переменной. Здесь следует отметить, что главное преимущество ХД перед остальными типами источников данных — это наличие семантического слоя, который дает возможность пользователю оперировать терминами предметной об-

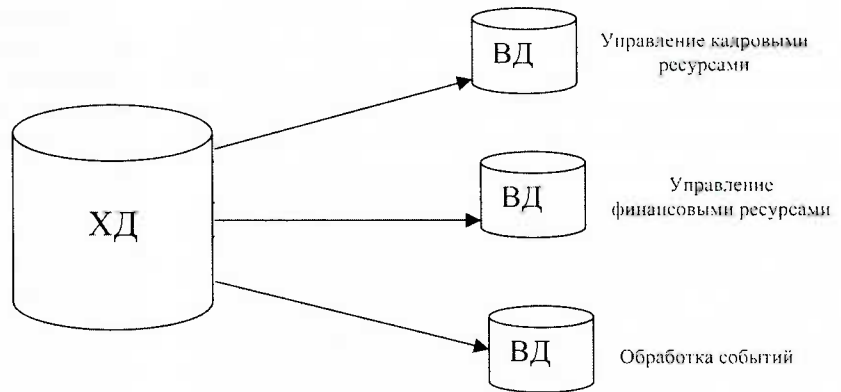


Рис. 1. Поликубическая OLAP-модель

ласти для формирования аналитических запросов к хранилищу [7].

В работе поставлена цель фазифицировать измерения OLAP-куба в целях разработки нечеткого OLAP-куба в поликубической модели OLAP на основе информации ХД в СППР в сфере ВП.

Для достижения поставленной цели были решены следующие задачи.

- Исследованы алгоритмы кластеризации PAM (Partitioning Around Medoids), CLARA (Clustering Large Applications) и CLARANS (Clustering Large Applications based upon Randomized Search), сделан их сравнительный анализ, на основании которого в качестве используемого был выбран алгоритм CLARANS.
- В ХД был применен алгоритм кластеризации CLARANS с указанием требуемого числа кластеров.
- В целях фазификации атрибутов AGE (возраст), EXPERIENCE (стаж работы), DATE (дата) и получения нечеткого OLAP-куба для них с помощью медоидов (медоид — центр кластера, являющийся одной из его точек) полученных кластеров были выведены значения функции принадлежности.

### II. Фазификация измерения OLAP-куба

Для решения данной проблемы обратимся к получившему в настоящее время широкое распространение аппарату теории нечетких множеств в задачах поиска информации [8].

Нечеткое множество, включая функцию принадлежности, являющуюся его характеристикой, может быть задано экспертом. Но в некоторых случаях функцию принадлежности удобнее задать аналитической формулой и для наглядности изобразить графически. Для задания функций принадлежности существуют различные типовые формы кривых, среди которых для решения нашей задачи выберем треугольную и трапециевидальную кусочно-линейные формы. В общем случае треугольная функция принадлежности определяется тремя числами, а трапециевидальная — четырьмя, хотя их число может и

меняться в зависимости от условий задачи. Для определения чисел, с помощью которых будет задана функция принадлежности, будем основываться на методе, описанном в работе [3], адаптируя его к нашей задаче, в которой по условию число термножеств на различных измерениях различно.

**А. Выбор алгоритма кластеризации.** Как известно, основная цель кластеризации заключается в разделении исследуемого множества объектов на группы "похожих" объектов, называемых кластерами. Классификация алгоритмов кластеризации дана в работе [9]. Две основные категории — это иерархические алгоритмы и неиерархические алгоритмы. В целях выбора эффективного алгоритма были рассмотрены неиерархические алгоритмы, основанные на методе разбиения, а именно на методе *k-medoids*, так как медоид — это центр кластера, который принадлежит кластеру, что является для нас главным условием. Кроме того, методы *k-medoids* являются устойчивыми к наличию выбросов и могут работать достаточно эффективно с большими наборами данных.

Алгоритмы, основанные на методе разбиений, проходят два основных этапа [10]:

- начальный шаг, на котором  $k$  объектов выбраны в качестве медоидов;
- оценочный шаг, на котором происходит попытка минимизировать целевую функцию, обычно основанную на сумме общего расстояния между невыбранными объектами (значениями) и их медоидами, т. е.

$$D(r, s) = \sum_{j=1}^n d(r_i, s_j),$$

где  $s_j \in S$  ( $S$  — множество объектов (значений) для кластеризации) и  $d(r_i, s_j) < d(r_c, s_j)$ ,  $\forall r_i, r_c \in R$  ( $R$  — множество объектов (значений) из  $S$ , выбранных в качестве медоидов),  $r_i \neq r_c$ .

Чем меньше сумма расстояний между медоидом и всеми другими объектами соответствующего ему кластера, тем лучше кластеризация.

Были рассмотрены три известных алгоритма, основанных на методе *k-medoids*: PAM, CLARA и CLARANS.

PAM является одним из первых алгоритмов, основанных на методе *k-medoids*. В работе [11] представлен основной принцип процесса кластеризации PAM, заключающийся в переборе всех объектов, которые в данный момент не являются медоидами, для вычисления расстояния от них до выбранных медоидов. Алгоритм PAM приводит к высокому качеству кластеров, но поскольку он пробует все возможные комбинации, он эффективен для небольших наборов данных. И в связи с его вычислительной сложностью его применение непрактично для кластеризации больших наборов данных.

Вычислительная сложность алгоритма PAM явилась мотивацией для разработки алгоритма CLARA — алгоритма кластеризации на основе выборки. Его принцип работы также можно найти в работе [11]. CLARA выделяет из множества данных несколько выборок, применяет PAM на каждой выборке и находит ее медоиды. Поскольку выборка проведена случайным образом, то медоиды выборки можно считать медоидами всего набора данных. Чтобы выбрать лучшие приближения, CLARA создаст несколько выборок и на выходе выдает лучшую кластеризацию. В работе [11] также экспериментально показано, что пять выборок размером  $(40 + 2k)$  дают удовлетворительные результаты.

CLARANS был разработан в рамках пространственного анализа данных. При поиске лучшего медоида на шаге оценки CLARANS случайным образом выбирает объекты из  $(n - k)$  объектов ( $n$  — число объектов в множестве;  $k$  — заданное число кластеров, или число медоидов). Число объектов, перебираемых на этом этапе, ограничено параметром *maxneighbor*, заданным пользователем. Если после *maxneighbor* попыток не будет найдено лучшее решение, то локальный оптимум считается достигнутым. Процедура продолжается до тех пор, пока не найдены *numloc* локальных оптимумов [12]. Алгоритм CLARANS, использующий стратегию случайного поиска, улучшает алгоритмы PAM и CLARA с точки зрения эффективности (сложность вычислений или времени) и эффекта кластеризации (среднее искажение на расстояниях) соответственно.

**В. Использование медоидов для вывода функции принадлежности.** Понятия "математическая теория нечетких множеств" (*fuzzy sets*) и "нечеткая логика" (*fuzzy logic*) впервые были предложены в работе [13]. Основной причиной появления новой теории стало наличие нечетких и приближенных рассуждений при описании человеком процессов, систем, объектов. В работах [14, 15] также дана информация о теории нечетких множеств, описание нечетких операторов, различные комбинации которых используются в операциях на нечетком OLAP-кубе.

Напомним основные определения. Характеристикой нечеткого множества является функция принадлежности (*membership function*). Обозначим через  $\mu(x)$  степень принадлежности элемента  $x$  к нечеткому множеству, представляющую собой обобщение понятия характеристической функции обычного множества. Тогда нечетким множеством  $S$  называется множество упорядоченных пар вида  $S = \{\mu(x)/x\}$ , при этом  $\mu(x)$  может принимать любые значения в интервале  $[0, 1]$ ,  $x \in X$ . При этом значение  $\mu(x) = 0$  означает отсутствие принадлежности к множеству  $X$ ;  $\mu(x) = 1$  — полную принадлежность к множеству  $X$ .

Под средой поиска информации будем понимать пару запрос—база данных. С точки зрения четко-

**Возможные варианты нечеткости в среде поиска информации**

№ варианта	Запрос	База данных
1	Четкий	Четкая
2	Четкий	Нечеткая
3	Нечеткий	Четкая
4	Нечеткий	Нечеткая

сти—нечеткости среды поиска информации могут быть четыре варианта (см. таблицу) [8].

Как видно из таблицы, наиболее общим является вариант 4.

Рассмотрим ХД СППР сферы внешней политики, в котором выделена ВД по подзадаче управления кадровыми ресурсами.

Пусть в целях создания нечеткого OLAP-куба на основании этой ВД из всего множества атрибутов надо фаззифицировать три атрибута — AGE (возраст), EXPERIENCE (стаж работы) и DATE (дата). При этом другие атрибуты, например LOCATION (место проведения мероприятия), могут оставаться четкими.

*Первый шаг.* Применим алгоритм CLARANS в БД, имеющей  $n$  атрибутов, задав число кластеров  $k = 3$ . В результате будут получены три кластера с центрами (медоидами) в точках  $m_1, m_2, m_3$ .

*Второй шаг.* Определим числа, на основании которых будет задана функция принадлежности для двух атрибутов — AGE (возраст) и EXPERIENCE (стаж работы). С этой целью каждый медоид представим как вектор  $m_i = \{a_{ij}, \dots, a_{ij}\}$ , где  $i = \overline{1, 3}$ ,  $j = \overline{1, 2}$ .

Для нашего случая имеем:

$$m_1 = \{a_{11}, a_{12}\}; m_2 = \{a_{21}, a_{22}\}; m_3 = \{a_{31}, a_{32}\}.$$

Множество  $\{a_{1j}, a_{2j}, a_{3j}\}$ ,  $j = \overline{1, 2}$ , представляет значения трех медоидов  $j$ -го атрибута, который предстоит фаззифицировать.

*Третий шаг.* На этом шаге будет проводиться фаззификация атрибута. В результате будут получены значения принадлежности всех значений  $j$ -го атрибута в каждом нечетком множестве.

Пусть  $J$  — множество значений  $j$ -го атрибута;  $F_{ij}(x)$  — функция принадлежности для получения значений принадлежности всех значений  $j$ -го атрибута в  $i$ -м нечетком множестве.

Для всех  $x \in J$  выполняем:

1) для первого нечеткого множества (т. е. при  $i = 1$ )

$$F_{1j}(x) = 1.0 \quad \text{if } x \leq a_{1j}$$

$$F_{1j}(x) = (x - a_{2j}) / (a_{1j} - a_{2j}) \quad \text{if } a_{1j} < x < a_{2j}$$

$$F_{1j}(x) = 0 \quad \text{if } x \geq a_{2j}$$

2) для второго нечеткого множества (т. е. при  $i = 2$ ):

$$F_{2j}(x) = 0 \quad \text{if } x \leq a_{1j}$$

$$F_{2j}(x) = (x - a_{1j}) / (a_{2j} - a_{1j}) \quad \text{if } a_{1j} < x < a_{2j}$$

$$F_{2j}(x) = 1.0 \quad \text{if } x = a_{2j}$$

$$F_{2j}(x) = (x - a_{3j}) / (a_{2j} - a_{3j}) \quad \text{if } a_{2j} < x < a_{3j}$$

$$F_{2j}(x) = 0 \quad \text{if } x \geq a_{3j}$$

3) для третьего нечеткого множества (т. е. при  $i = 3$ ):

$$F_{3j}(x) = 0 \quad \text{if } x \leq a_{2j}$$

$$F_{3j}(x) = (x - a_{2j}) / (a_{3j} - a_{2j}) \quad \text{if } a_{2j} < x < a_{3j}$$

$$F_{3j}(x) = 1 \quad \text{if } x \geq a_{3j}$$

Отметим, что после фаззификации полученные три кластера станут тремя терм-множествами. Для атрибута EXPERIENCE это {малый, средний, продолжительный}, а для атрибута AGE это {молодой, средний, пожилой}. При этом атрибуты EXPERIENCE и AGE становятся лингвистическими переменными.

На рис. 2 и рис. 3 дано графическое изображение терм-множеств соответствующих атрибутов. Отметим, что номер  $j$  следует закреплять за соответствующим измерением. В данном случае  $j = 1$  — измерение AGE,  $j = 2$  — измерение EXPERIENCE.

Атрибут DATE будет разбит на два терм-множества (*недавно* и *давно*), поэтому для его фаззификации на первом шаге применим в БД алгоритм CLARANS, задав число кластеров  $k = 2$ . В результате будут получены два кластера с центрами (медоидами) в точках  $m_1, m_2$ . На втором шаге определяются числа, на основании которых будет задана функция принадлежности для этого атрибута. В этом случае имеем  $m_1 = \{a_1\}$ ,  $m_2 = \{a_2\}$ , т. е. две точки на

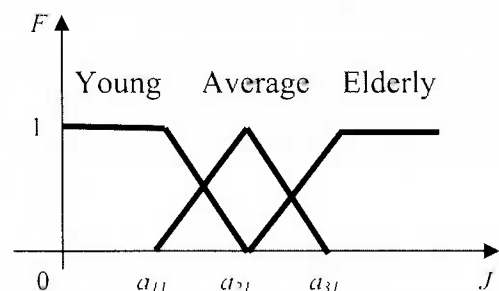


Рис. 2. Графическое изображение лингвистической переменной AGE

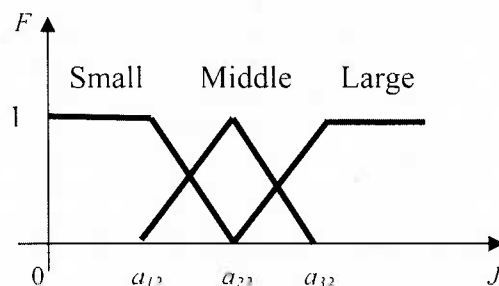


Рис. 3. Графическое изображение лингвистической переменной EXPERIENCE

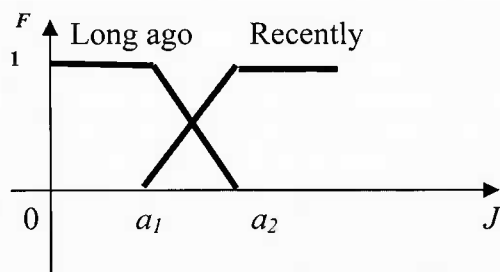


Рис. 4. Графическое изображение лингвистической переменной DATE

одном измерении для задания функции принадлежности.

На третьем шаге проводится соответствующая фаззификация, в результате которой будут получены два терм-множества. На рис. 4 показано графическое изображение лингвистической переменной DATE.

В результате фаззификации измерений куба рассматриваемой СППР можно выполнять нечеткие запросы, используя термины *молодой*, *зрелый*, *пожилой* (касательно возраста), *малый*, *средний*, *продолжительный* (касательно стажа работы), *недавно* и *давно* (касательно даты мероприятия) в сочетании с четкими значениями базы данных, таких как место проведения мероприятия, название мероприятия и др.

Приведем примеры типовых запросов, которые могут сыграть не последнюю роль в управлении кадровыми ресурсами.

1. Сколько *молодых* сотрудников посетили *недавно* Турцию?
2. Выдать список сотрудников, имеющих *продолжительный* стаж работы, которые посетили *недавно* Турцию в связи с мероприятиями в медицинской сфере.
3. В каких странах *молодые* сотрудники прошли *недавно* тренинги в военной сфере?
4. Выдать список *молодых* и *зрелых* сотрудников, прошедших тренинги в США в сфере ИКТ и т. д.

### III. Практическое применение

Следует отметить, что включение разрабатываемой СППР в программу электронного правительства способно обеспечить информационную поддержку проводимой в стране внешней политики. В настоящее время эту систему, включая описанный выше подход разработки нечеткого OLAP-куба, планируется реализовать в Управлении международных связей (УМС) Национальной академии наук Азербайджана (НАНА) на основе Распоряжения Президиума НАНА об информатизации Центрального Аппарата Президиума НАНА, куда как структурное подразделение включено и УМС. Отметим, что на данный момент вся документация, связанная с зарубежными поездками сотрудников НАНА, хра-

нится в бумажном или электронном виде в архивах различных отделов — отделе кадров, бухгалтерии, научно-организационном отделе и т. д., т. е. хранение этой информации носит неорганизованный, бессистемный характер. В целях внедрения разрабатываемой системы будет использован минимум 2,2 ГГц на платформе Microsoft Windows 7. Для анализа OLAP будет использована Microsoft Analysis Server, 2008. Алгоритм CLARANS и задание функции принадлежности, а также Front-end (интерфейс пользователя) планируется реализовать в Object Pascal (в среде Delphi). В качестве базы данных будет использован SQL-сервер, язык запросов — T—SQL. Набор данных будет содержать информацию за последние 12 лет, что приблизительно составляет 16 000—20 000 записей. При таких условиях время отклика на запрос пользователя к данной системе ожидается очень коротким — не более 2 с. Измерениями могут быть *Data*, *Event* и *Persons*, таблицей фактов — *costs* (финансовые затраты на мероприятие). Измерение *Event* состоит из атрибутов: ID; название мероприятия; дата проведения мероприятия; сфера, к которой оно относится; город; страна; численность населения; текущая информация о стране; документы, содержащие международные (в которые включена данная страна) или двусторонние контракты, договоры, соглашения и т. п. или информация об этих документах (т. е. их название, дата заключения, срок действия и т. д.) и др. Измерение *Data* в качестве атрибутов содержит: ID; день (может отсутствовать); название месяца и год поездки (не путать с датой проведения мероприятия). Измерение *Persons* в качестве атрибутов содержит подробную информацию о командированном сотруднике: ID; имя, фамилию, год рождения, место работы, должность, стаж работы и т. д.

### Заключение

Обычное ХД, объединяющее данные из различных источников, обеспечивает лицам, принимающим решения, возможность делать анализ точных данных. Однако в процессе принятия решений можно столкнуться с множеством нечетких задач, вследствие чего в запросах к базам данных, которые пытается формулировать исследователь, часто могут присутствовать неточности и неопределенности. В этих случаях необходимо, чтобы СППР для оказания помощи руководителям в принятии решений поддерживала соответствующую технологию представления знаний. В статье показан метод фаззифицирования измерений OLAP-куба в целях разработки нечеткого OLAP-куба в поликубической OLAP-модели на основе информации ХД в СППР в сфере внешней политики. В статье рассматривается случай, когда число создаваемых на измерениях куба терм-множеств разное.

## Список литературы

1. **Арсентьев А.** Хранилища данных становятся инфраструктурным компонентом № 1 // CNews аналитика. 2010. [Электронный ресурс]. URL: <http://retail.cnews.ru/reviews/free/BI2010/articles/articles6.shtml>
2. **Immon W.** Building the Data Warehouse, Second eds. New York: John Wiley & sons Inc., 1996.
3. **Kumarl K. N. V. N. P., Radha K. P., De S. K.** Fuzzy OLAP Cube for Qualitative Analysis. Intelligent Sensing and Information Processing, 2005 // Proc. of 2005 International Conference on. 2005. P. 290—295. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1529464&tag=1](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1529464&tag=1)
4. **Wang S.** Application of Decision Support System in E-government // Proc. of International Conference "Management and Service Science", 2009. MASS'09. URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5302532>
5. **Nabibakova G.** Development of Data Warehouse Architecture in the Informational Analytical Systems of Decision-Making Support in the Field of Foreign Policy // International Journal of Ubiquitous Computing and Internationalization. December 2011. Vol. 3, N 2. P. 15—18.
6. **Каширин И. Ю., Семченков С. Ю.** Интерактивная аналитическая обработка данных в современных OLAP-системах // Бизнес-информатика. 2009. № 2(08).
7. **Паклин Н. Б., Орешков В. И.** Бизнес-аналитика: от данных к знаниям. СПб.: ПИТЕР, 2010.
8. **Рыжов А. П.** Модели поиска информации в нечеткой среде. М.: ЦПИ МГУ, 2004.
9. **Чубукова И. А.** Data Mining: учеб. пособие. М.: БИНОМ. Лаборатория знаний, 2006. 382 с.
10. **Barioni M. C. N., Razente H. L., Traina A. J. M., Jr C. T.** An efficient approach to scale up k-medoid based algorithms in large databases. XXI Brazilian Symposium on Databases. 2006. <http://www.lbd.dcc.ufmg.br:8080/colcooes/sbbd/2006/018.pdf>
11. **Kaufman L., Rousseeuw P. J.** Finding Groups in Data: an Introduction to Cluster Analysis. John Wiley & Sons, 2005.
12. **Ng R. T., Han J.** Clarans: A method for clustering objects for spatial data mining // IEEE Transactions on Knowledge and Data Engineering (TKDE). 2002. 14(5). P. 1003—1016.
13. **Zadeh L.** Fuzzy sets // Information and Control. 1965. N 8. P. 338—353.
14. **Klir G. J., Yuan B.** Fuzzy sets and fuzzy logic: Foundations and applications. New Jersey: Prentice Hall, 1995. 256 p. URL: <http://www.bookadda.com/books/fuzzy-set-theory-foundations-applications-g-j-0133410587-9780133410587>
15. **Рыжов А. П.** Элементы теории нечетких множеств и измерения нечеткости. М.: Диалог-МГУ, 1998.

УДК 004.652.5

**Н. Т. Фисун**, д-р техн. наук, проф., зав. каф., e-mail: [ntfis@kma.mk.ua](mailto:ntfis@kma.mk.ua)

**Г. В. Горбань**, аспирант,

Черноморский государственный университет им. П. Могилы, г. Николаев, Украина

## Модели и методы построения системы OLAP для объектно-ориентированных баз данных

*Рассмотрены аспекты построения системы OLAP в объектно-ориентированных СУБД, предложена схема связи фактов с измерениями в объектно-ориентированной модели данных и диаграмма метаклассов для реализации OLAP при использовании метаданных.*

**Ключевые слова:** базы данных, OLAP, реляционные БД, объектно-ориентированные БД, таблица фактов, измерение, мера, многомерный массив, класс, экземпляр класса, свойство, ссылка, связь, метаданные, нотация IDEF1

**N. T. Fisun, G. V. Gorbanj**

## Models and Methods for Building OLAP System for Object-Oriented Databases

*Aspects of building OLAP systems in an object-oriented DBMS are discussed in the article. A scheme of relationship of facts with dimensions in the object-oriented data model and metaclass diagram for OLAP implementation using metadata are proposed.*

**Keywords:** database OLAP, relational database, object-oriented database, fact table, dimension, measure, multidimensional array, class, instance of the class, property, reference, relationship, metadata, notation IDEF1

### Введение

В настоящее время в системах поддержки принятия решений (СППР) важное место занимают технологии оперативного (OLAP) и интеллектуаль-

ного (Data Mining) анализа данных [1, 2]. Данные технологии уже реализованы в некоторых коммерческих системах управления базами данных (СУБД), таких как Microsoft SQL Server, Oracle и других, которые представляют собой реляционные СУБД.