

**Список литературы**

1. Кравец А. Г. Моделирование процесса решения задач по физике / А. Г. Кравец, О. В. Титова // Открытое образование. – 2011. – № 2. – С. 76–79.
2. Титова О. В. Результаты исследования потребности в консультировании студентов в сети Интернет / О. В. Титова // Инновационный малый бизнес как основа модернизации региональной экономики : мат-лы межрегион. молодежной науч. конф. (11–12 марта 2011 г.) / ГОУ ВПО «Калмыцкий гос. ун-т» [и др.]. – Элиста, 2011. – С. 51–54.
3. Титова О. В. Основные педагогические концепции консультационного IT-центра / О. В. Титова // Инновационные технологии в обучении и производстве : мат-лы V Всерос. науч.-практ. конф. (г. Камышин, 4–6 декабря 2008 г.) : в 3 т. / КТИ (филиал) ВолГТУ [и др.]. – Камышин, 2008. – Т. 2. – С. 231–235.
4. Титова О. В. Телекоммуникационный подход в обучении студентов-заочников / О. В. Титова // Известия ВолГТУ. – Волгоград, 2010. – Вып. 7 : межвуз. сб. науч. ст. – № 8. – С. 187–189. – (Сер. «Новые образовательные системы и технологии обучения в вузе»).
5. Финкельштейн Р. М. Что делать, когда решить задачу не удастся / Р. М. Финкельштейн. – М. : ИЛЕКСА, 2008. – 74 с.: ил.

**References**

1. Kravets A. G. Modelirovanie protsessa resheniya zadach po fizike / A. G. Kravets, O. V. Titova // Otkrytoe obrazovanie. – 2011. – № 2. – S. 76–79.
2. Titova O. V. Rezultaty issledovaniya potrebnosti v konsultirovanii studentov v seti Internet / O. V. Titova // Innovatsionnyu malu biznes kak osnova modernizatsii regionalnoy ekonomiki: [mater.] mezhregion. molodezhnoy nauch. konf. (11–12 marta 2011 g.) / GOU VPO "Kalmytskiy gos. un-t" [i dr.]. – Elista, 2011. – S. 51–54
3. Titova O. V. Osnovnye pedagogicheskie kontseptsii konsultatsionnogo IT-centra / O. V. Titova // Innovatsionnye tehnologii v obuchenii i proizvodstve : mater. Vseros. nauch.-prakt. konf. (Kamyshin, 4–6 dek. 2008 g.) : v 3 t. / KTI (filial) VolGTU [i dr.]. – Kamyshin, 2008. – T. 2. – S. 231–235.
4. Titova O. V. Telekommunikatsionnyu podhod v obuchenii studentov-zaochnikov / O. V. Titova // Izvestia VolGTU. – Vyp. 7 : mezhvuz. sb. nauch. st. – Volgograd, 2010. – № 8. – S. 187–189. – (Ser. "Novye obrazovatelnye sistemy i tehnologii obucheniya v vuze").
5. Finkelshteyn R. M. Chto delat, kogda reshit zadachu ne udaetsya / R. M. Finkelshteyn. – M. : ILEKSA, 2008. – 74 s.: il.

УДК 519.237.8

**ПРИМЕНЕНИЕ ТЕОРИИ МНОЖЕСТВ ДЛЯ ОРГАНИЗАЦИИ ДАННЫХ  
ИСХОДНОЙ РЕЛЯЦИОННОЙ БАЗЫ ДАННЫХ**

*Меликов Алексей Владимирович, магистр техники и технологии, Пензенский государственный университет, 440026, Россия, г. Пенза, ул. Красная, 40, e-mail: MelikovAlexey@mail.ru.*

*В статье рассматривается возможность организации данных в исходной реляционной базе данных информационной системы анкетирования. Организация данных в реляционной базе данных необходима для осуществления последующей передачи их в систему многомерного анализа данных. Данные следует сформировать в требуемую структуру, которая будет отвечать требованиям хранилища анализируемых данных. Для организации данных в необходимую структуру применяется теория множеств. Данные рассматриваются как отдельные множества, к которым применяются операции над множествами для получения конечного набора данных.*

---

---

## ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

---

---

После получения конечного набора данных из исходной реляционной базы данных рассматривается необходимость организовать данные по измерениям многомерной модели данных. Организация данных по определенному измерению осуществляется в семействе множества, элементами которого являются всевозможные подмножества выбранного множества, которое определяется уровнем детализации. Эта организация нужна для последующего анализа данных, в зависимости от уровня детализации; можно проводить анализ информации в разрезах, которые не были предусмотрены разработчиками системы сбора информации. В дополнении к этому, иерархии в измерениях в конкретно описанной модели данных необходимы как для реализации операций анализа данных, так и для структурирования заголовков пользовательского представления.

В случае применения теории множеств для организации данных в нужную структуру сокращается интервал времени, который необходим для формирования схемы новой многомерной модели данных.

**Ключевые слова:** информационная система анкетирования, реляционная база данных, многомерный анализ данных, агрегирование данных, уровень детализации, теория множеств, операции над множествами.

### THE USE OF SET THEORY FOR THE DATA MANAGEMENT OF THE INPUT RELATIONAL DATABASE

*Melikov Alexey V., Master in Engineering and Technology, Penza State University, 40 Krasnaya str., Penza, 440026, Russia, e-mail: MelikovAlexey@mail.ru.*

*In this investigation consideration of data organization for the input relational database of information questionnaire system is given. The data organization of the relational database is necessary for its further transferring to the multidimensional data analysis system. The data should be organized into the required structure that would meet the requirements of the analyzed data storage. Set theory is used for the data organization into the required structure. The data is considered as separate sets that are operated in order to acquire the final data set.*

*After final data set acquisition from the input relational data base the data needs to be organized according to the multidimensional data model readings. Data organization for a particular dimension is performed in a collection of sets, whose elements are all possible subsets of the selected set, which is determined by the level of detail. This organization is necessary for the further data analysis, depending on the level of detail; the informational analysis may be performed in the ways that were not provided by the original programmer. In addition, the hierarchies in the dimensions of the data model described are needed to both perform the data analysis operations and structure the headings of the user's view.*

*In case we use the theory of sets to organize the data into the desired structure, the time needed for formation of the scheme of a new multidimensional data model will be reduced.*

**Key words:** *informational questionnaire system, relational data base, multidimensional data analysis, data aggregation, level of detail, set theory, set operation.*

При создании архитектуры приложения, предназначенного для бизнес-аналитики, одним из основных требований является быстрый и гибкий анализ данных. В реляционном хранилище данных содержатся все транзакции детального уровня. Пользователи, как правило, хотят получать итоговую информацию по этим данным, добавлять свои собственные вычисления и анализировать полученные данные с помощью средств для создания произволь-

ных запросов. Программы анализа многомерных данных (или многомерного анализа данных, multidimensional data analysis) – это продукты, занимающие по принципу устройства промежуточное положение между базами данных и электронными таблицами и образующие сегодня особый класс продуктов наряду с последними [4, с. 10]. Исторически они являются развитием электронных таблиц, но в настоящее время больше напоминают базы данных. Если в электронных таблицах наиболее важное свойство – возможность установления связей между ячейками посредством формул, то в программах анализа многомерных данных главное – возможность легко формировать многомерный куб из готовых данных и модифицировать их экранные представления.

В настоящее время распространены программы трех типов:

- 1) добавления к популярным системам электронных таблиц;
- 2) добавления к популярным реляционным системам управления базы данных;
- 3) самостоятельные продукты, способные импортировать данные из различных источников.

В данной работе рассмотрена возможность организации данных исходной реляционной базы данных информационной системы анкетирования для последующей передачи их в систему многомерного анализа данных, которая оптимизирована для быстрого доступа к агрегированным данным. Для организации данных в необходимую структуру применяется теория множеств. Система анализа многомерных данных должна эффективно использовать масштабируемость базы данных и предоставлять быстрый интегрированный доступ к хранилищу данных детального уровня. Однако для пользователей с более сложными требованиями к аналитике система должна подключаться к источникам многомерных данных. Подключение к таким источникам необходимо, чтобы предоставить доступ к расширенным возможностям вычислений, прогнозам, размещениям, запросам к различным измерениям и агрегированным данным. Программа поддерживает извлечение данных из реляционной базы данных информационной системы анкетирования и хранит свои данные в специально оптимизированном многомерном хранилище данных. Но эти данные хранятся вне ядра реляционной базы данных. Обычно для этого служит свой собственный выделенный сервер.

#### *Организация данных реляционной базы данных*

Одним из основных исходных понятий математики является понятие множества и его элементов. Основатель теории множеств Кантор дал такую трактовку: «Под множеством понимают объединение в одно общее объектов, хорошо различимых нашей интуицией или нашей мыслью» [2].

Понятие «множество», как и любое другое исходное понятие, не имеет строгого математически точного описания. В данной работе используется следующее понимание этого определения: «Множество – это совокупность определенных различаемых объектов, причем таких, что для каждого можно установить, принадлежит этот объект данному множеству или нет».

Множества могут быть заданы списком, порождающей процедурой, арифметическими операциями, описанием свойств элементов или графическим представлением [1, с. 27]. В нашем случае интересно задание множеств списком, которое предполагает перечисление элементов. Например, множество  $A$  состоит из букв  $a, b, c, d$ :  $A = \{a, b, c, d\}$  или множество  $N$  включает цифры  $0, 2, 3, 4$ :  $N = \{0, 2, 3, 4\}$ . Таким образом, данные, которые необходимо извлечь из исходной реляционной базы данных информационной системы анкетирования, можно записать в виде следующих множеств:

$R_1$  = прохождение теста № (дата, № теста, № вопроса, № ответа, № группы);

$R_2$  = тесты (№ теста, название теста, описание теста);

$R_3$  = вопрос (№ вопроса, текст вопроса);

$R_4$  = группа (№ группы, № кафедры, название группы);

## ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

$R_5$  = кафедра (№ кафедры, № факультета, название кафедры);

$R_6$  = факультет (№ факультета, название факультета).

Существуют такие операции над множествами, как объединение, пересечение, разность, симметрическая разность и дополнение. В данной работе внимание будет уделено объединению множеств.

Определение 1. Объединением множеств **A** и **B** ( $A \cup B$ ) называется множество, состоящее из всех тех элементов, которые принадлежат хотя бы одному из множеств **A** или **B** [2, с. 53]. Математически данная операция записывается так:  $A \cup B = \{x | x \in A \text{ или } x \in B\}$ . В общем случае операция объединения может быть использована для нескольких множеств:  $A \cup B \cup C \cup D$  или  $S = \{A_1, A_2, \dots, A_k\}$ . Последнее можно представить в следующем

виде:  $S = \bigcup_{i=1}^k A_i$ , где  $k$  – количество объединенных множеств.

**Пример 1.** Необходимо организовать исходные данные реляционной базы данных, представленные множествами  $R_1, R_2, R_3, R_4, R_5, R_6$ , в такую структуру, которая была бы удобна для последующей передачи в многомерную модель данных.

$R_1 \cup R_2 = R_{1,2}$  {дата, № теста, название теста, описание теста, № вопроса, № ответа, № группы};

$R_{1,2} \cup R_3 = R_{1,2,3}$  {дата, № теста, название теста, описание теста, № вопроса, текст вопроса, № ответа, № группы};

$R_{1,2,3} \cup R_4 = R_{1,2,3,4}$  {дата, № теста, название теста, описание теста, № вопроса, текст вопроса, № ответа, № группы, № кафедры, название группы};

$R_{1,2,3,4} \cup R_5 = R_{1,2,3,4,5}$  {дата, № теста, название теста, описание теста, № вопроса, текст вопроса, № ответа, № группы, название группы, № кафедры, название кафедры, № факультета};

$R_{1,2,3,4,5} \cup R_6 = R_{1,2,3,4,5,6}$  {дата, № теста, название теста, описание теста, № вопроса, текст вопроса, № ответа, № группы, название группы, № кафедры, название кафедры, № факультета, название факультета}.

Объединение происходит вокруг множества  $R_1$  (так как в нем содержатся базовые данные), путем добавления информации к ключевым полям в множестве. Порядок объединения с множеством  $R_1$  определен следующими иерархиями (см. рис.).



Рис. Иерархии в исходной базе данных

Самой важной информацией в многомерной базе данных является ее структура. Она описывает многомерное и иерархическое устройство базы данных и содержит подробные сведения об измерениях, членах измерений, иерархиях, показателях, производных показателях и ключевых параметрах эффективности, которые хранятся в ней.

**Организация полученных данных по измерениям**

После получения данных из исходной реляционной базы данных необходимо организовать данные по измерениям многомерной модели данных. Эта организация нужна для последующего анализа данных в зависимости от уровня детализации; можно проводить анализ информации в разрезах, которые не были предусмотрены разработчиками системы сбора информации.

Определение 2. Мощностью конечного множества  $M$  называется число его элементов. Обозначается  $|M|$ . Принято считать, что пустое множество  $\emptyset$  является подмножеством любого множества. Множество может обладать иерархической структурой. В этом случае говорят о семействе множества, или булеане. Семейством множества  $M$ , или булеаном ( $\beta(M)$ ), является множество, элементами которого являются всевозможные подмножества множества  $M$ . Например,

$$M = \{a, b, c\}, \beta(M) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}. |M| = 3, |\beta(M)| = 8.$$

В общем случае мощность булеана  $|\beta(M)| = 2^{|M|}$  [3, с. 65].

**Пример 2.** Организовать полученные данные из исходной реляционной базы данных по измерениям многомерной модели данных в зависимости от уровня детализации (в данном случае: иерархия «Группа-Кафедра-Факультет»).

Сначала стоит определить семейство множества  $F$  (факультет), в котором существуют как выпускающие, так и не выпускающие кафедры (семейство множества  $K$ ). Таким образом,  $F \ni K_i (i=1,2,\dots,5)$ , то есть  $F = \{K_1, K_2, K_3, K_4, K_5\}$  (см. табл. 1).

Таблица 1

**Выпускающие/не выпускающие кафедры факультета**

Кафедра	Выпускающая	Не выпускающая
$K_1$		
$K_2$		
$K_3$		
$K_4$		
$K_5$		

$$\beta(F) = \{\emptyset, \{K_1\}, \{K_2\}, \{K_3\}, \{K_4\}, \{K_5\}, \{K_1, K_4\}, \{K_2, K_3, K_5\}, \{K_1, K_2, K_3, K_4, K_5\}\}.$$

$|F| = 5, |\beta(F)| = 9$ . Понятным становится тот факт, что базисом является следующая последовательность кортежей:

$$B = \{\emptyset, \{K_1\}, \{K_2\}, \{K_3\}, \{K_4\}, \{K_5\}\}.$$

Некоторые кафедры выпускают группы учащихся по нескольким специальностям. В таблице 2 показан пример соответствия специальностей ( $S_j$ , где  $j = 1,2,3$ ) выпускающим кафедрам ( $K_i$ ), входящим в то же семейство множеств, но обозначенных  $F^*$ ; где  $G_1$  будет обозначена группа учащихся.

Таблица 2

Выпускающие кафедры по специальностям			
Специальность / Кафедра	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>
K <sub>1</sub>			
K <sub>2</sub>	G <sub>1</sub>		
K <sub>3</sub>	G <sub>2</sub>		
K <sub>4</sub>			
K <sub>5</sub>		G <sub>3</sub>	G <sub>4</sub>

$$\beta(F^*) = \{\emptyset, \{K_1\}, \{K_2\}, \{K_3\}, \{K_4\}, \{K_5\}, \{K_2, K_3\}, \{K_1, K_2, K_3, K_4, K_5\}\}.$$

$|F^*| = 5$ ,  $|\beta(F^*)| = 8$ . Из приведенных таблиц визуально определяется, что количество групп учащихся на факультете  $F = 4$ , то есть  $l = 1, 2, 3, 4$ . Таким образом,  $G_1, G_2, G_3, G_4$  являются точками множеств [5, с. 101], комбинирование которых позволяет получать семейства множеств различных уровней детализации данных. Зная, что  $S_1 = \{G_1, G_2\}$ ,  $S_2 = \{G_3\}$ ,  $S_3 = \{G_4\}$ , и принимая во внимание, что  $K_1 = \emptyset$ ,  $K_2 = \{G_1\}$ ,  $K_3 = \{G_2\}$ ,  $K_4 = \{\emptyset\}$ ,  $K_5 = \{G_3, G_4\}$ , получаем:  $S_1 = K_2 \cap K_3 = \{G_1, G_2\}$ ,  $K_5 = S_2 \cap S_3 = \{G_3, G_4\}$ . Применительно к семейству множества  $F$  мы имеем, что  $F = K_1 \cap K_2 \cap K_3 \cap K_4 \cap K_5 = \{\emptyset, \{G_1\}, \{G_2\}, \{G_3\}, \{G_4\}\}$ . Из вышесказанного следует, что:

$$\begin{aligned} G_1 &\in S_1, G_1 \in K_2, G_1 \in F; \\ G_2 &\in S_1, G_2 \in K_3, G_2 \in F; \\ G_3 &\in S_2, G_3 \in K_5, G_3 \in F; \\ G_4 &\in S_3, G_4 \in K_5, G_4 \in F. \end{aligned}$$

В данной работе была описана возможность организации данных в исходной реляционной базе данных информационной системы анкетирования для последующей передачи их в систему многомерного анализа данных. Применяя теорию множеств для организации данных в необходимую структуру, сокращается интервал времени, необходимый для формирования схемы новой многомерной модели данных. Кроме того, иерархии в измерениях на примере описанной выше модели данных необходимы как для реализации операций анализа данных, так и для структурирования заголовков пользовательского представления. Формирование иерархий осуществляется таким образом, чтобы представление самой модели в виде двумерной таблицы было удобным для работы.

#### Список литературы

1. Бурбаки Н. Теория множеств : пер. с фр. / Н. Бурбаки // Начала математики. – 2010. – 2-е изд. – С. 27–28, 51–57, 65–69.
2. Князьков В. С. Введение в теорию множеств и комбинаторику / В. С. Князьков, Т. В. Волченская. – М. : Интернет-университет информационных технологий – ИНТУИТ.ру, 2004.
3. Левзнер Л. Д. Математические основы теории систем / Л. Д. Левзнер, Е. П. Чураков. – М. : Высшая школа, 2009. – С. 100–101.
4. Bingham E. Multidimensional data analysis / E. Bingham. – Dept of Computer Science, University of Helsinki, 2008. – P. 9–13.

**References**

1. Burbaki N. Teoriya mnozhestv : per. s fr. / N. Burbaki // Nachala matematiki. – 2010. – Izd. 2. – S. 27–28, 51–57, 65–69.
2. Knyaz'kov V. S. Vvedenie v teoriyu mnozhestv i kombinatoriku / V. S. Knyaz'kov, T. V. Volchenskaya. – M. : Internet-universitet informacionnyh tehnologii – INTUIT.ru, 2004.
3. Levzner L. D. Matematicheskie osnovy teorii sistem / L. D. Levzner, E. P. Churakov. – M. : Vyssh. shk., 2009. – S. 100–101.
4. Bingham E. Multidimensional data analysis / E. Bingham. – Dept of Computer Science, University of Helsinki, 2008. – P. 9–13.

УДК 004.031

**АЛГОРИТМ ФОРМИРОВАНИЯ СТРУКТУРЫ  
КОРПОРАТИВНОГО ИНТРАНЕТ-ПОРТАЛА**

*Сафоненко Надежда Викторовна, ассистент, Национальный исследовательский ядерный университет «МИФИ», 115409, Россия, г. Москва, Каширское ш., 31, e-mail: NadinSaf@mail.ru.*

*В условиях постоянно меняющегося рынка крайне важной является возможность быстро и точно создавать корпоративный портал в соответствии с требованиями заказчика. При создании такого портала необходимо предусмотреть возможность масштабирования, а также изменения самого портала. Как показывает опыт, ранее созданный портал со временем требует серьезной модификации – изменяются корпоративные стандарты, регламенты, бизнес-процессы, а также потребности бизнеса в информационном обеспечении. Ниже предложен алгоритм, описывающий процесс разработки нового корпоративного портала, по которому процесс разработки нового корпоративного портала дополняется процессами модификации. Таким образом, разработка, внедрение и сопровождение интранет-портала объединяются в единый процесс, целью которого является получение продукта, востребованного пользователями и отвечающего современным требованиям к информационным системам в целом и требованиям данного заказчика в частности. При разработке алгоритма использовались основные рекомендации RUP, примененные для задания основных шагов алгоритма: итерационное формирование структуры интранет-портала, применение графического проектирования, в особенности для визуализации функциональных требований заказчика и требований к интерфейсу. Данная работа касается продуктов, взаимодействующих на верхних уровнях сетевой модели OSI, в том числе сеансовом, прикладном и уровне представления.*

**Ключевые слова:** портал, интранет, структура, экспертная оценка, группа экспертов, проектирование, процесс разработки, функциональные требования, техническое задание, интерфейс.