

УДК 519.6

UDC 519.6

**МЕТОДИКА АНАЛИЗА
ИНФОРМАЦИОННОЙ СТРУКТУРЫ БАЗЫ
ДАНЫХ АВТОМАТИЗИРОВАННОЙ
СИСТЕМЫ СОСТАВЛЕНИЯ РАСПИСАНИЙ**

**THE METHOD FOR DATABASE
INFORMATIONAL STRUCTURE ANALYSIS IN
TIMETABLE SCHEDULING AUTOMATIC
SYSTEM**

Семенюта Ирина Сергеевна
аспирантка кафедры информатики, специальность
05.13.01 «Системный анализ, управление и
обработка информации»
*Кубанский государственный технологический
университет, Краснодар, Россия*

Semenyuta Irina Sergeevna,
postgraduate of the Informatics Department,
specialism 05.13.01 "System analysis, management
and information computing"
*Kuban State Technological University, Krasnodar,
Russia*

В статье предложена методика проектирования
информационной структуры базы данных
автоматизированной системы составления
расписания вуза. Приведен алгоритм упорядочения
информационных элементов по уровням иерархии

The method for database informational structure
development in timetable scheduling automatic system
of high school is reviewed in this article. The
algorithm for informational elements ordering on
hierarchy levels is described in this work

Ключевые слова: МНОЖЕСТВО, ОБЪЕКТ
АВТОМАТИЗАЦИИ, ГРАФ, ИЕРАРХИЯ,
МАТРИЦА СМЕЖНОСТИ, МАТРИЦА
ДОСТИЖИМОСТИ

Keywords: MULTITUDE, AUTOMATION OBJECT,
EARL, HIERARCHY, RELATED MATRIX,
MATRIX OF REACH ABILITY

Предметная область задачи составления расписания занятий вуза при проектировании базы данных автоматизированной системы включает большое количество информационных элементов и взаимосвязей между ними. На начальном этапе проектирования информационной системы необходимо тщательно проанализировать существующие данные и упорядочить их должным образом. Описанный ниже подход является основанием для последующего проектирования базы данных аудиторного фонда на основе структурированного системного анализа. Модель предметной области может быть представлена в виде совокупности семи множеств [1]:

$$M_{npO} = \langle F, H, P, O, V^{vx}, V^{bix}, R \rangle, \text{ где}$$

$$F = \{f_i \mid i = \overline{1, I}\} \quad \text{множество пользовательских функций (таблица 1);}$$

$$H = \{h_j \mid j = \overline{1, J}\} \quad \text{множество задач обработки данных (таблица 2);}$$

$$O = \{o_m \mid m = \overline{1, M}\} \quad \text{множество объектов автоматизации (таблица 3);}$$

$P = \{p_k \mid k = \overline{1, K}\}$ множество пользователей (таблица 4);

$V^{вх} = \{v_l \mid l \in L_{вх}\}$ множество входных данных;

$V^{вых} = \{v_l \mid l \in L_{вых}\}$ множество выходных данных;

$V = V^{вх} \cup V^{вых}$ полное множество информационных элементов предметной области (таблица 5);

$R = \{r_y \mid y = \overline{1, Y}\}$ множество отношений (взаимосвязей) между компонентами $M_{прО} = \langle F, H, P, O, V^{вх}, V^{вых}, R \rangle$.

Таблица 1 – Множество автоматизируемых пользовательских функций

F_1	Выборка, обновление, добавление, удаление данных по аудиториям
F_2	Выборка, обновление, добавление, удаление данных по группам студентов
F_3	Выборка, обновление, добавление, удаление данных по преподавателям
F_4	Выборка, обновление, добавление, удаление данных учебных занятий
F_5	Выборка, обновление, добавление, удаление данных по учебным корпусам
F_6	Выборка, обновление, добавление, удаление данных по дисциплинам

Таблица 2 – Множество задач обработки данных

H_1	Обновление данных
H_2	Удаление данных
H_3	Добавление данных
H_4	Выборка данных по запросу пользователя

Таблица 3 – Множество объектов автоматизации

O_1	Аудитория
O_2	Группа студентов

O_3	Преподаватель
O_4	Учебное занятие
O_5	Учебный корпус
O_6	Дисциплина

Таблица 4 – Множество пользователей базы данных

P_1	Диспетчер
P_2	Подмножество операторов
P_3	Подмножество факультетов (деканов)
P_4	Подмножество кафедр (заведующих кафедр)
P_5	Подмножество преподавателей

Таблица 5 – Основные информационные элементы

V_1	Номер аудитории
V_2	Вместимость аудитории
V_3	Специализация аудитории
V_4	Принадлежность аудиторному фонду
V_5	Номер группы
V_6	Количество человек в группе
V_7	Специальность
V_8	Кафедра
V_9	Факультет
V_{10}	Фамилия преподавателя
V_{11}	Имя преподавателя
V_{12}	Отчество преподавателя
V_{13}	Ученая степень
V_{14}	Ученое звание
V_{15}	Должность
V_{16}	Специализация преподавателя
V_{17}	День недели
V_{18}	Номер недели
V_{19}	Время начала занятия
V_{20}	Время окончания занятия
V_{21}	Вид занятия
V_{22}	Количество недель проведения занятия
V_{23}	Шифр корпуса
V_{24}	Адрес корпуса
V_{25}	Код дисциплины
V_{26}	Название дисциплины

Следующий этап – выделение отношений (взаимосвязей) между компонентами предметной области:

$r(O,V)$ – отношение «объекты – информационные элементы».

Далее необходимо построить полное множество структурных элементов предметной области D . Оно образуется путем объединения множеств информационных элементов V , объектов автоматизации O и соответствующей индексации полученного множества D :

Таблица 6 – Полное множество структурных элементов

d_1	Аудитория
d_2	Номер аудитории
d_3	Вместимость аудитории
d_4	Специализация аудитории
d_5	Принадлежность аудиторному фонду
d_6	Учебный корпус
d_7	Группа студентов
d_8	Номер группы
d_9	Количество человек в группе
d_{10}	Специальность
d_{11}	Кафедра
d_{12}	Факультет
d_{13}	Преподаватель
d_{14}	Фамилия преподавателя
d_{15}	Имя преподавателя
d_{16}	Отчество преподавателя
d_{17}	Ученая степень
d_{18}	Ученое звание
d_{19}	Должность
d_{20}	Специализация преподавателя
d_{21}	Учебное занятие
d_{22}	День недели
d_{23}	Номер недели
d_{24}	Время начала занятия
d_{25}	Время окончания занятия
d_{26}	Вид занятия
d_{27}	Количество недель проведения занятия
d_{28}	Шифр корпуса
d_{29}	Адрес корпуса
d_{30}	Дисциплина
d_{31}	Код дисциплины
d_{32}	Название дисциплины

Под матрицей семантической смежности $B = \|b_{ij}\|$ будем понимать квадратную бинарную матрицу, проиндексированную по обеим осям множества структурных элементов D_k и содержащую запись $b_{ij} = 1$, если на основании информации пользователей о семантической связности

элементов между структурными элементами d_i и d_j существует отношение R такое, что элемент d_i составляет (расширяет, дополняет и.т. д.) смысловое содержание элемента d_j и $b_{ij}=0$ – в противном случае [2].

Изображение матрицы $B = \|b_{ij}\|$ приведено на рис. 1.



Рисунок 1 – Матрица семантической смежности $B = \|b_{ij}\|$

Матрице $B = \|b_{ij}\|$ ставится в соответствие граф информационной структуры $G(D,U)$, множеством вершин которого являются структурные элементы множества D_k , а дуга (d_i, d_j) соответствует записи $b_{ij}=1$ в матрице B_k . Таким образом, дуги орграфа $G(D,U)$ отражают наличие или отсутствие семантической связности между структурными элементами. Изображение орграфа $G(D,U)$ представлено на рис. 2.

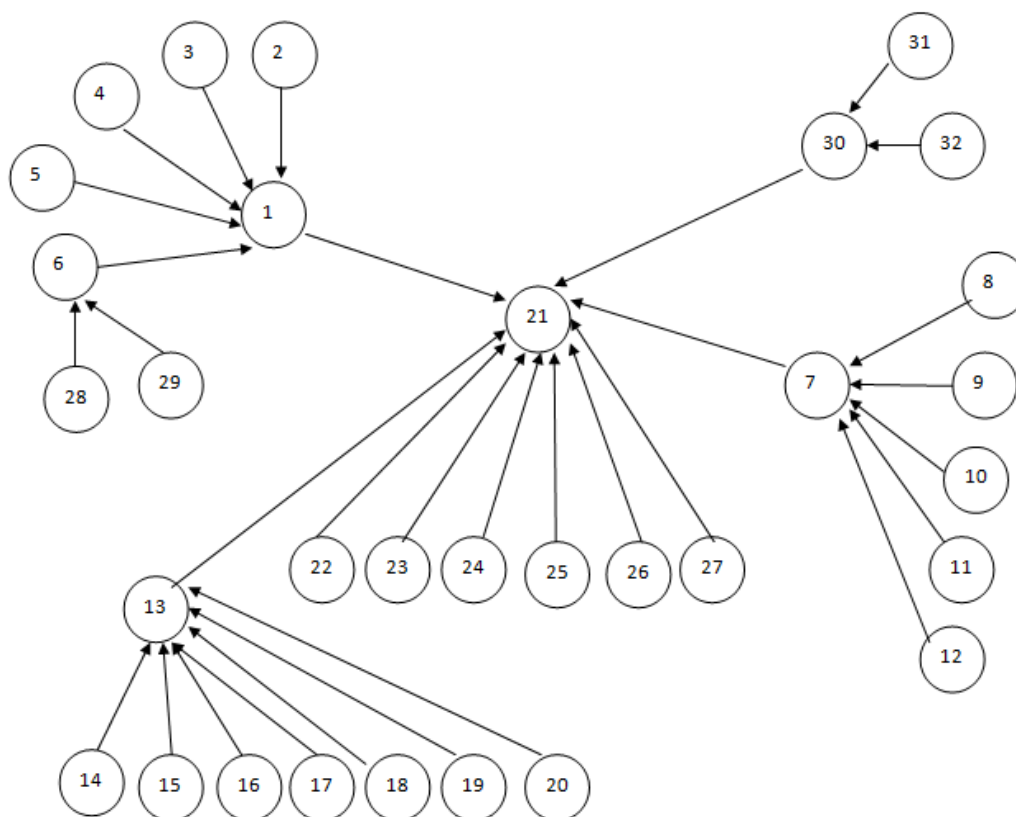


Рисунок 2 – Граф $G(D,U)$ информационной структуры

Под матрицей достижимости $A = \|a_{ij}\|$ будем понимать квадратную бинарную матрицу (рис. 3), проиндексированную одинаковым образом по обеим осям множества структурных элементов D_k . Запись $a_{ij} = 1$ матрицы $A = \|a_{ij}\|$ соответствует наличию или смыслового отношения достижимости R_0 элемента d_j из элемента d_i , $d_i R_0 d_j$. При этом считается, что элемент d_j семантически достижим из элемента d_i , если на графе $G(D,U)$ существует путь от вершины d_i к вершине d_j , имеющий определенное смысловое содержание. При этом отношение достижимости удовлетворяет условию транзитивности, т.е. если $d_i R_0 d_n$ и $d_n R_0 d_j$, то $d_i R_0 d_j$, $i, j, n = 1, 2, \dots, P(D)$.

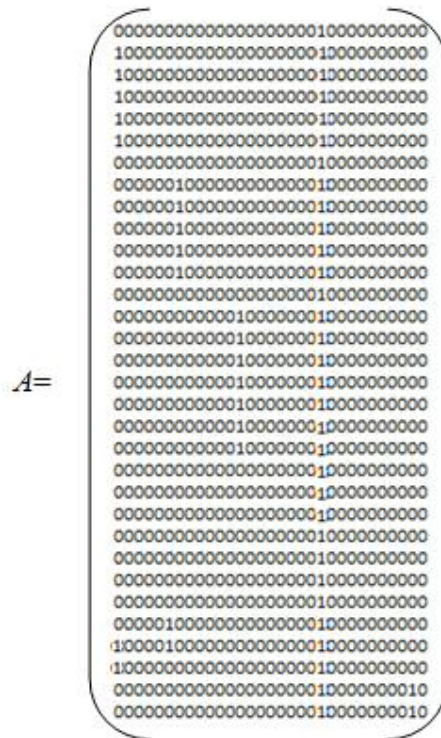


Рисунок 3 – Матрица достижимости $A = \|a_{ij}\|$

Матрица $A = \|a_{ij}\|$ дает возможность определения множества предшествования $C(d_i)$ и достижимости $F(d_i) \forall d_i \in D$. Множество $C(d_i)$ формируется из элементов, соответствующих единичным записям в i -м столбце, а множество $F(d_i)$ – из элементов, соответствующих единичным записям в i -й строке матрицы $A = \|a_{ij}\|$. Анализ множества $C(d_i)$ позволяет выделить базовые типы структурных элементов – информационные элементы и группы. Информационным элементам соответствуют те структуры, для которых $C(d_i) = 0$. На графе $G(D, U)$ им соответствуют висячие вершины. Для определения информационных элементов необходимо просуммировать элементы каждого столбца j матрицы $A = \|a_{ij}\|$.

Если $\sum_{i=1}^{P(D)} a_{ij} = 0$ то j -й элемент структурного множества является информационным. В противном случае структурный элемент является групповым элементом (группой). В данном случае имеем:

$$\begin{aligned} \sum_{i=1}^{32} a_{i2} = \mathbf{K} &= \sum_{i=1}^{32} a_{i5} = \sum_{i=1}^{32} a_{i8} = \dots = \sum_{i=1}^{32} a_{i12} = \sum_{i=1}^{32} a_{i14} = \dots = \sum_{i=1}^{32} a_{i20} = \\ &= \sum_{i=1}^{32} a_{i22} = \dots = \sum_{i=1}^{32} a_{i29} = \sum_{i=1}^{32} a_{i31} = \sum_{i=1}^{32} a_{i32} = 0 \end{aligned}$$

$$\sum_{i=1}^{32} a_{i1} > 0, \sum_{i=1}^{32} a_{i6} > 0, \sum_{i=1}^{32} a_{i7} > 0, \sum_{i=1}^{32} a_{i13} > 0, \sum_{i=1}^{32} a_{i21} > 0, \sum_{i=1}^{32} a_{i30} > 0.$$

Множество информационных элементов обозначим через D^δ :

$$D^\delta = \{d_2 - d_5, d_8 - d_{12}, d_{14} - d_{20}, d_{22} - d_{29}, d_{31}, d_{32}\}.$$

Множество групп D^ϵ определяется из выражения:

$$D^\epsilon = D \setminus D^\delta = \{d_1, d_6, d_7, d_{13}, d_{21}, d_{30}\}.$$

С целью упорядочения групп по уровням иерархии в матрице A выделяется подматрица (удалением индексов элементов множества D^δ)

$A^\epsilon = \|a_{ij}^\epsilon\|$ (рис. 4), где $a_{ij}^\epsilon = 1$ при наличии связи между группами $d_j^\epsilon \in D^\epsilon$.

$$A^\epsilon = \begin{pmatrix} & \mathbf{1} & \mathbf{6} & \mathbf{7} & \mathbf{13} & \mathbf{21} & \mathbf{30} \\ \mathbf{1} & 0 & 0 & 0 & 0 & 1 & 0 \\ \mathbf{6} & 1 & 0 & 0 & 0 & 1 & 0 \\ \mathbf{7} & 0 & 0 & 0 & 0 & 1 & 0 \\ \mathbf{13} & 0 & 0 & 0 & 0 & 1 & 0 \\ \mathbf{21} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{30} & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Рисунок 4 – Подматрица $A^\epsilon = \|a_{ij}^\epsilon\|$

Матрице $A^\epsilon = \|a_{ij}^\epsilon\|$ соответствует подграф графа $G(D,U)$. Изображение подграфа G представлено на рис. 5.

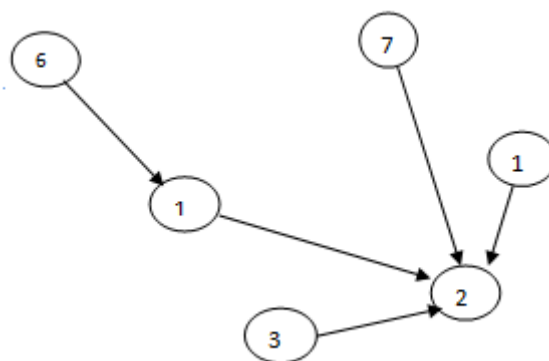


Рисунок 5 – Изображение подграфа G

Определим множества предшествования $C(d_i)$ и достижимости $F(d_i) \forall d_i \in D$, которые представлены в табл. 7.

Таблица 7 – Элементный состав множеств предшествования и достижимости для множества групп

Множество предшествования	Индекс	Множество достижимости	Индекс	$F \cap C$
$C(d_1)$	6	$F(d_1)$	1,21	
$C(d_6)$	6	$F(d_6)$	1,21	
$C(d_7)$	7	$F(d_7)$	21	
$C(d_{13})$	13	$F(d_{13})$	21	
$C(d_{21})$	1,6,13,21,30	$F(d_{21})$	21	21
$C(d_{30})$	30	$F(d_{30})$	21	

Группа $d_i^c \in D_k^c$ принадлежит множеству групп верхнего уровня p_1 , если $F(d_i^c) \cap C(d_i^c) = F(d_i^c)$. Принадлежность остальных групп уровням иерархии $p_m, m \geq 2$ определяется итеративным образом из соотношения:

$$p_m = \{d_i^c \in D_k^c \setminus p_1 \setminus \dots \setminus p_{m-1} / F_{m-1}(d_i^c) \cap C_{m-1}(d_i^c) = F_{m-1}(d_i^c)\}$$

Из данных таблицы 7 можно сделать вывод о том, что в данной предметной области верхний уровень иерархии представлен группой d_{21}

«Учебное занятие». Аналогично определяются второй и последующие уровни иерархии. Упорядоченный по иерархии граф представлен на рис. 6:

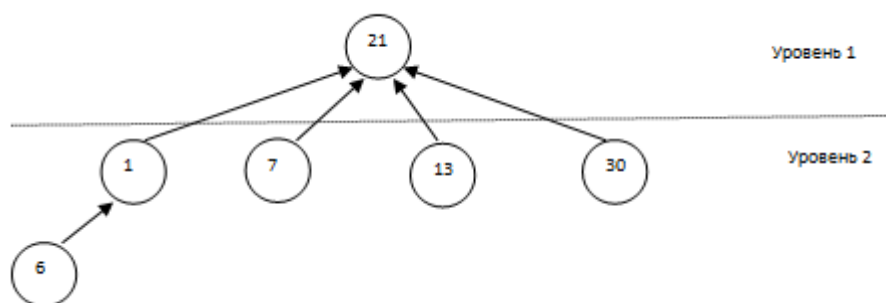


Рисунок 6 – Упорядоченный по иерархии граф G

Таким образом, предлагаемый подход позволяет упорядочить группы информационных элементов по уровням иерархии, выделить группы, являющиеся корневыми группами структуры, и группы, занимающие промежуточное положение. Выявленная в процессе исследования корневая группа «Учебное занятие» определяет возможные точки входа в информационную структуру (информационные группы «Преподаватель», «Аудитория», «Группа студентов», «Дисциплина»), а промежуточные («Учебный корпус») расширяют сведения об информации, помещенной в корневых и вышележащих группах. Предложенная методика может быть в дальнейшем использована для нормализации информационных структур в задаче составления автоматизированного расписания занятий [3] в целях обеспечения минимальной избыточности и дублируемости данных и связей между элементами базы данных.

Список литературы

1. Атрощенко В.А., Семенюта И.С. К вопросу построения автоматизированных систем составления расписаний // Инновационные процессы в высшей школе: Труды XV Юбилейной Всероссийской научно-практической конференции; Изд-во ЮГ, 2009. С. 6–10.
2. Теоретические основы проектирования оптимальных структур распределенных баз данных / В.В. Кульба, С.С. Ковалевский, С.А. Косяченко и др. М.: СИНТЕГ, 1999. С. 114–116.
3. Атрощенко В.А., Чигликова Н.Д., Коновалов Д.П., Семенюта И.С. К вопросу построения оптимальных структур баз данных автоматизированной системы составления расписания // Технические и технологические системы: Материалы международной научной конференции. Краснодар: КубГАУ, 2009. С.350–353.