

## АНАЛИЗ ФУНКЦИОНАЛЬНОЙ ЗАВИСИМОСТИ АТТРИБУТОВ РАЗЛИЧНЫХ СУЩНОСТЕЙ ДЛЯ ВЫЯВЛЕНИЯ НОВЫХ СВЯЗЕЙ

В. В. Осипова

Национальный исследовательский Томский политехнический университет,  
634050, Томск, Россия

УДК 004.065(075.8)

Предложен новый подход для выявления связей между различными сущностями в концептуальной информационной модели предметной области. Проведен анализ функциональной зависимости атрибутов этих сущностей от направления связи и типов атрибутов.

**Ключевые слова:** функциональная зависимость, концептуальная информационная модель, реляционная модель данных.

In this work a new approach for discovering relationships between different entities in the conceptual information model of the subject domain is proposed. The analysis of functional dependence of the entities attributes depending on relationship directions and attributes types is carried out.

**Key words:** functional dependence, conceptual information model, relational data model.

**Введение.** Эффективность современных информационных систем в сфере управления во многом определяется результатом проектирования базы данных (БД). Наиболее трудно формализуемым этапом проектирования БД, особенно для больших и сложных областей, является этап концептуального проектирования [1], в результате которого собранные требования к данным анализируются и оформляются в виде концептуальной информационной модели предметной области (КИМПО). Поэтому любые подходы, направленные на формализацию создания КИМПО, весьма актуальны.

Традиционные интеграционные методики концептуального проектирования предназначены для инфологического представления предметной области, и процесс создания КИМПО сводится к декларативному описанию последовательности действий на основе интеграции информационных потребностей пользователей. Однако такие методики не предлагают собственно способов формирования модели – формализованного выявления объектов предметной области и связей между ними. Как отметил известный ученый Р. Баркер, процесс построения КИМПО является своего рода искусством [2].

Предложенная в работе [3] альтернативная интеграционная методика проектирования КИМПО позволяет учитывать отмеченную специфику моделирования и на основе анализа информационных потребностей формализованно устанавливать взаимосвязи между объектами – сущностями предметной области. В отличие от известных методик в этой методике, основанной на реляционной модели данных, используется анализ доменов атрибутов для выявления

связей между сущностями, т. е. декларируются принадлежность значений атрибутов определенным доменам и отношения между доменами, а связи автоматически определяются. Таким образом, описание доменов и отношений между ними как часть процесса фиксации базовых знаний о предметной области необходимо для определения наличия связи между сущностями.

В настоящей работе предлагается для выявления связей в КИМПО помимо анализа доменов атрибутов выполнять анализ функциональной зависимости атрибутов различных сущностей.

**1. Элементы реляционной модели данных в КИМПО.** При анализе функциональной зависимости атрибутов сущностей введем следующие обозначения для графического отображения в нотации Бахмана:  $S_i$  – сущность с номером  $i$ ;  $K$  – ключевой атрибут;  $A$  – неключевой атрибут с номером  $i$ ; штрих – атрибут, заимствованный из другой сущности (номер атрибута определяется из заимствованной сущности); " $\rightarrow$ " – связь между двумя сущностями типа 1:M; " $\leftrightarrow$ " – связь между двумя сущностями типа M:M;  $---\blacktriangleright$  – функциональная зависимость атрибута одной сущности от атрибута другой сущности в направлении к зависимому атрибуту.

Для установления зависимостей атрибутов сущностей используем понятие функциональной зависимости (ФЗ) атрибута  $Y$  от атрибута  $X$  в сущности  $S$  (понятие сущности в КИМПО соответствует понятию отношения в реляционной модели данных), означающее, что если известно значение атрибута (детерминанта)  $X$ , то значение атрибута (зависимой части)  $Y$  определено и имеет единственное значение [4]:  $X(S) \rightarrow Y(S)$ . В реляционной модели функциональная зависимость связана прежде всего с понятием ключа: в сущности  $S$  должны быть атрибуты  $K$ , называемые ключом, при этом все остальные, неключевые атрибуты  $A_i$  должны зависеть от ключа:  $K(S) \rightarrow A_i(S)$ . При этом в сущности может появиться транзитивная функциональная зависимость, когда неключевой атрибут  $A_j$  транзитивно зависит от ключа  $K$ , т. е. он функционально зависит не непосредственно от ключа, а от другого неключевого атрибута, зависящего от ключа:  $\exists A_i(S_i): [K(S_i) \rightarrow A_i(S_i)] \cup [A_i(S_i) \rightarrow A_j(S_i)]$  (рис. 1). Данная проблема решается путем приведения сущности  $S_i$  в третью нормальную форму, когда ни один неключевой атрибут не является транзитивно зависимым от ключа сущности. Для этого транзитивно зависимый атрибут  $A_j$  исключается из сущности  $S_i$  и добавляется в новую сущность  $S_j$ , в которую добавляется детерминирующий атрибут  $A_i$ , выполняющий функцию ключевого:  $\exists S_j: [K(S_j) \rightarrow A_j(S_j)] \cup [K(S_j) \leftrightarrow A_i(S_i)]$  (рис. 2).

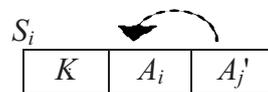


Рис. 1. Транзитивная функциональная зависимость

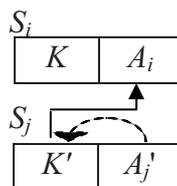


Рис. 2. Разрешение транзитивной функциональной зависимости

Следует отметить, что связь между сущностями  $S_i$  и  $S_j$  обеспечивается за счет сопоставимости ключевого атрибута одной сущности и неключевого атрибута другой сущности  $K(S_j) \Leftrightarrow A_i(S_i)$  на основе пересечения их доменов  $d(K(S_j)) \cap d(A_i(S_i)) \neq \emptyset$ , образующих связь типа 1:M.

**2. Подход, основанный на анализе функциональной зависимости атрибутов сущностей.** Для выявления связей между различными сущностями можно выполнить анализ функциональной зависимости атрибута  $Y$  одной сущности  $S_j$  от атрибута  $X$  другой сущности  $S_i$ :  $X(S_i) \rightarrow Y(S_j)$  на основе описанного выше решения для функциональной зависимости в пределах одной сущности. Для этого требуется наличие такой сущности  $S_k$ , в которую необходимо добавить зависимый атрибут  $Y$  в качестве неключевого, а в качестве ключа использовать детерминирующий атрибут  $X$ :  $\exists S_k: [X(S_k) \rightarrow Y(S_k)] \cup [X(S_k) \in K(S_k)] \cup [Y(S_k) \in A_i(S_k)]$ . При этом подобной сущностью  $S_k$  может служить существующая сущность  $S_i$ , если в ней ключом является атрибут  $X$ . Однако сущность  $S_j$  не может быть сущностью  $S_k$ , так как изначально в сущности  $S_j$  имеется другой ключ, который нельзя заменить атрибутом  $X$ . Таким образом, для исключения функциональной зависимости атрибутов между сущностями необходимо получить функциональную зависимость этих атрибутов внутри одной сущности, в которой детерминирующий атрибут является ключевым, а зависимый – неключевым.

**3. Анализ функциональной зависимости атрибутов различных сущностей.** Проведем анализ функциональной зависимости атрибутов различных сущностей с различными направлениями связи и типами атрибутов, участвующих в зависимостях. Рассмотрим возможные ситуации:

1. Ключевой атрибут одной сущности зависит от ключевого атрибута другой сущности  $K(S_i) \rightarrow K(S_j)$  (рис. 3).

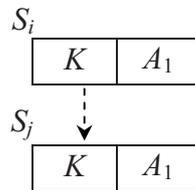


Рис. 3. Зависимость ключевого атрибута одной сущности от ключевого атрибута другой сущности

Предлагаемое решение: добавить зависимый ключевой атрибут  $K(S_j)$  в качестве неключевого  $A_i$  в сущность  $S_i$  с детерминирующим ключевым атрибутом  $\exists A_i(S_i): [K(S_i) \rightarrow A_i(S_i)] \cup [K(S_j) \Leftrightarrow A_i(S_i)]$  (рис. 4).

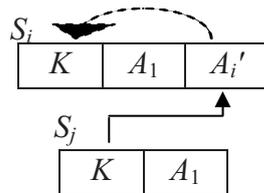


Рис. 4. Разрешение зависимости ключевого атрибута одной сущности от ключевого атрибута другой сущности

2. Имеет место зависимость между ключевым и неключевым атрибутами сущностей:

2.1. Ключевой атрибут одной сущности зависит от неключевого атрибута другой сущности  $A_i(S_i) \rightarrow K(S_j)$  (рис. 5).

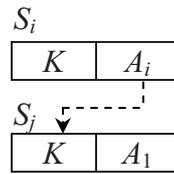


Рис. 5. Зависимость ключевого атрибута одной сущности от неключевого атрибута другой сущности

Предлагаемое решение: так как детерминирующий атрибут  $A_i(S_i)$  является неключевым в своей сущности  $S_i$ , необходимо создать новую сущность  $S_k$ , в которой он становится ключевым  $K(S_k)$ , и связать обе сущности  $S_i$  и  $S_k$  отношением 1:M. Далее зависимый ключевой атрибут  $K(S_j)$  добавляется в качестве неключевого атрибута в новую сущность  $S_k$ , связанную также с сущностью  $S_j$  с зависимым атрибутом отношением 1:M  $\exists S_k: [K(S_k) \rightarrow A_i(S_i)] \cup [K(S_k) \leftrightarrow A_i(S_i)] \cup [K(S_j) \leftrightarrow A_i(S_k)]$  (рис. 6).

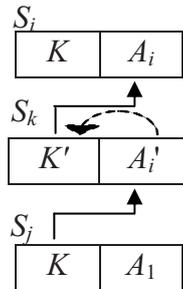


Рис. 6. Разрешение зависимости ключевого атрибута одной сущности от неключевого атрибута другой сущности

2.2. Неключевой атрибут одной сущности зависит от ключевого атрибута другой сущности:  $K(S_i) \rightarrow A_i(S_j)$  (рис. 7).

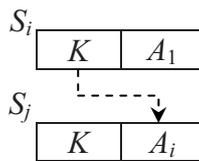


Рис. 7. Зависимость неключевого атрибута одной сущности от ключевого атрибута другой сущности

Предлагаемое решение: добавить зависимый неключевой атрибут  $A_i(S_j)$  в сущность  $S_i$  с детерминирующим ключевым атрибутом (рис. 8).

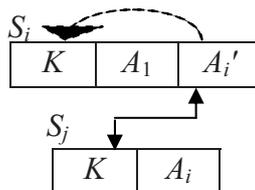


Рис. 8. Исключение зависимости неключевого атрибута одной сущности от ключевого атрибута другой сущности

Так как обе сущности  $S_i$  и  $S_j$  связаны по зависимому атрибуту  $A_i$  отношением М:М, для избавления от этой связи добавляется новая сущность  $S_k$ , в которой зависимый атрибут  $A_i$  становится ключевым  $K(S_k)$  и которая связана с сущностями  $S_i$  и  $S_j$  по этому атрибуту отношением  $1:M \exists A_i(S_i) \exists S_k: [K(S_i) \rightarrow A_i(S_i)] \cup [K(S_k) \leftrightarrow A_i(S_i)] \cup [K(S_k) \leftrightarrow A_i(S_j)]$  (рис. 9):

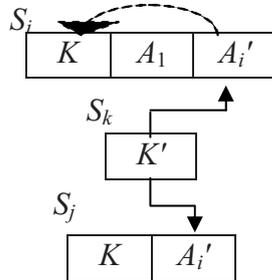


Рис. 9. Разрешение зависимости неключевого атрибута одной сущности от ключевого атрибута другой сущности

3. Неключевой атрибут одной сущности зависит от неключевого атрибута другой сущности:  $A_i(S_i) \rightarrow A_i(S_j)$  (рис. 10).

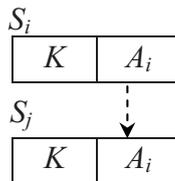


Рис. 10. Зависимость неключевого атрибута одной сущности от неключевого атрибута другой сущности

Предлагаемое решение: выполнить п. 2.1 для детерминирующего атрибута  $A_i(S_i)$  и п. 2.2 для новой добавленной сущности  $S_k$ , добавив в нее зависимый атрибут  $A_i(S_j)$ . Поскольку детерминирующий атрибут  $A_i(S_i)$  является неключевым в своей сущности  $S_i$ , необходимо создать новую сущность  $S_k$ , в которой он становится ключевым  $K(S_k)$ , и связать ее с сущностью  $S_i$  с детерминирующим атрибутом отношением 1:М. Далее зависимый неключевой атрибут  $A_i(S_j)$  добавляется в качестве неключевого атрибута в новую сущность  $S_k$ , связанную с сущностью  $S_j$  с зависимым атрибутом отношением М:М (рис. 11).

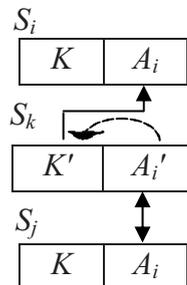


Рис. 11. Исключение зависимости неключевого атрибута одной сущности от неключевого атрибута другой сущности

Так как новая сущность  $S_k$  и сущность  $S_j$  с зависимым атрибутом связаны по зависимому атрибуту  $A_i$  отношением М:М, для разрешения этой связи нужно добавить вторую новую сущ-

ность  $S_i$ , в которой зависимый атрибут становится ключевым  $K(S_i)$  и которая связана с сущностями  $S_j$  и  $S_k$  по этому атрибуту отношением 1:M:  $\exists S_k \exists S_j: [K(S_k) \rightarrow A_i(S_k)] \cup [K(S_k) \leftrightarrow A_i(S_i)] \cup [K(S_i) \leftrightarrow A_i(S_k)] \cup [K(S_i) \leftrightarrow A_i(S_j)]$  (рис. 12).

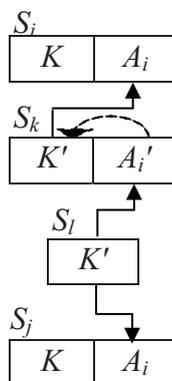


Рис. 12. Разрешение зависимости неключевого атрибута одной сущности от неключевого атрибута другой сущности

**Заключение.** Таким образом, для выявления новых связей между различными сущностями в КИМПО выполнен анализ функциональной зависимости атрибутов этих сущностей. В зависимости от направления связи и типов атрибутов, участвующих в зависимостях, рассмотрены возможные типы функциональных зависимостей атрибутов для выявления связей между сущностями, которые удовлетворяют реляционной модели данных.

#### Список литературы

1. NAVATHE S. B. Evolution of data modeling for databases // Commun. ACM. 1992. V. 35, iss. 9. P. 112–123.
2. BARKER R. Case\*method – entity relationship modelling. Wokingham: Addison Wesley Prof., 1990. P. 13–14.
3. ЧУДИНОВ И. Л., ИСАЕВ И. В. Интеграционный подход к концептуальному проектированию информационной базы единой информационной среды // Теоретические и прикладные вопросы современных информационных технологий: Материалы 6-й Всерос. науч.-техн. конф., Улан-Удэ, 25–30 июля 2005 г. Улан-Удэ: ВСГТУ, 2005.
4. ДЕЙТ К. Дж. Введение в системы баз данных: Пер. с англ. 6-е изд. Киев; М.; СПб.: Вильямс, 2000.

*Осипова Виктория Викторовна – ассист.  
Томского политехнического университета;  
тел.: (382-2) 56-33-94; e-mail: vikosi@tpu.ru*

Дата поступления – 16.09.12 г.