

ТРЕХМЕРНЫЕ МОДЕЛИ ДАННЫХ

Емельяненко В. И.

Кафедра телекоммуникаций и информационных технологий, Белорусский государственный университет
Минск, Республика Беларусь
E-mail: emelvi@bsu.by

В данной работе рассматриваются вопросы архитектурной организации распределенной среды для практической организации мониторинга сложных систем. В основу положено координатное представление реляционных моделей данных и построение решеток формальных понятий в качестве схем соответствия типов для организации поисковых запросов.

ВВЕДЕНИЕ

В последнее время предметом дискуссии является недостаточная выразительность реляционных СУБД в задачах разработки масштабируемых приложений, что привело к появлению альтернативных систем, таких как объектные базы данных, базы данных NoSQL и т.п. Однако все это присутствует в виде отдельных результатов без достаточно строгого обоснования предлагаемых решений. В работе [1] было показано, что реляционная модель как таковая не накладывает ограничений на формы регистрации данных и доступ к ним при распределенной организации их хранения. Ограничения возникают при выборе отношений между структурными типами. Поэтому для решения прикладных задач требуется наличие обобщенных моделей, предоставляющих возможность создавать гибкие решения. Реляционные СУБД являются моделями такого рода, но для узкой категории задач. Именно здесь и сосредоточены ключевые проблемы.

1. ОСНОВНАЯ ЧАСТЬ

В данной работе делается попытка дальнейшего развития модельных конфигураций [1] на платформе методологии анализа формальных понятий (Formal Concept Analysis – FCA) [2], которая сейчас успешно применяется при моделировании сложных систем. Основу подхода составляет описание поведения системы через посредство ее представления в виде набора локальных координат в конфигурационном пространстве (Рис. 1). На рисунке 1 а) показаны две координатные оси U_k и U_m локального N -мерного пространства $U^R = (u_1^R, \dots, u_m^R, u_k^R, \dots, u_N^R)$, представляющие наблюдаемое состояние $R(u_k^R, u_m^R)$ отдельно объекта. В данном случае регистрация переменных состояния u_i^R объекта осуществляется в пространстве параметров наблюдения D_1, D_2, D_3 в точке $R = (d_1^R, d_2^R, d_3^R)$ его траектории R', R, R'' . В системе с большим числом объектов в пространство наблюдений вводятся составляющие E – событие, T – структурный тип и V – экземпляр объекта заданного структурного типа. В этом случае одновременно регистрируются изменения состояний нескольких объектов. На рисунке 1 б)

такая ситуация представлена для двух событий e_1, e_2 , трех реализаций $I = (i_1, i_2, i_3)$ объектов пяти структурных типов t_1, t_2, t_3, t_4, t_5 .

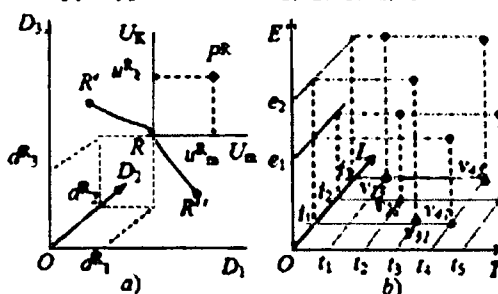


Рис. 1 – Диаграммы траекторий объектов

Собственно объекты $v_{2,1}, v_{2,2}, v_{3,1}, v_{4,1}, v_{4,3}$ и $v_{5,2}$ помечены в нулевом слое I, O, T . Это начальные состояния системы. Далее слои событий идут в порядке их следования. Объекты помечены двойными индексами: первый – номер структурного типа, а второй – индекс реализации. В общем случае модельное пространство имеет следующие составляющие:

$$E, T, V, D_1, \dots, D_M, U_{1,1}, \dots, U_{1,N},$$

$$U_{2,1}, \dots, U_{2,2N}, \dots, U_{L,1}, \dots, U_{L,LN}, \quad (1)$$

где $E = e_1, \dots, e_K$ – события, $T = t_1, \dots, t_L$ – структурные типы, $D = D_1, \dots, D_M$ – измерения, $U = U_{1,1}, \dots, U_{1,N}, U_{2,1}, \dots, U_{2,2N}, \dots, U_{L,1}, \dots, U_{L,LN}$ – атрибуты структурных типов и $V = v_{1,1}, \dots, v_{1,J_1}, v_{2,1}, \dots, v_{2,J_2}, \dots, v_{L,1}, \dots, v_{L,J_L}$ – объекты.

В пространстве E, T, V траектории объектов находятся на вертикальных линиях. Участие объекта в событии отмечается точкой, что означает регистрацию его переменных состояний. Если траектория просто пересекает слой события, то это означает, что объект в этом событии не участвует. Поля E, T, V гарантируют отсутствие повторяющихся записей при регистрации событий и являются ключевыми идентификаторами структуры (1). Значения полей D_1, \dots, D_M задают параметры условий, при которых реализуются наблюдаемые состояния объектов. В полях $U_{1,1}, \dots, U_{1,N}, U_{2,1}, \dots, U_{2,2N}, \dots, U_{L,1}, \dots, U_{L,LN}$ представляются зарегистрированные значения переменных состояния каж-

дого из объектов. Это реляционная модель данных. Но, пока данная модель представляет последовательный набор, в общем-то, не связанных между собой записей определенного формата. Как уже отмечалось, в задачах обработки данных распространение получили методологии, основу которых составляет возможность создавать иерархии классов путем задания соответствий между объектами предметной области и наборами свойств, которыми они обладают. Для этого на множестве объектов V и множестве атрибутов U задается отношение $I \in V \times U$, такое, что vIu ($v \in V, u \in U$) тогда и только тогда, когда u есть атрибут объекта v . В этом случае образуются взаимно однозначно связанные пары (G, P) семейств, где $G \subseteq V$ и $P \subseteq U$. Такие пары (G_j, P_j) называются концептами (формальными понятиями), а тройка $K = (U, V, I)$ называется формальным контекстом. Формальный контекст может быть представлен в виде матрицы, строки которой помечены именами объектов, а столбцы – значениями атрибутов. Множество формальных понятий (P, G) , частично упорядочено отношением R (можно назвать его, например, «менее общий чем или равен»): $(P_j, G_j) \leq (P_q, G_q)$, если $P_j \leq P_q$ или $G_q \leq G_j$ (что эквивалентно) – и образует концептуальную решетку контекста K . Рассмотрим эти положения на примере, где в качестве атрибутов выступает свойство объектов участвовать в том или ином системном событии. На Рис. 1 б) представлена ситуация с двумя событиями, в которых участвует шесть объектов. В событии e_1 зарегистрированы объекты $v_{2,2}$, $v_{4,1}$ и $v_{5,2}$. В событии e_2 зарегистрированы объекты $v_{1,3}$, $v_{3,1}$ и $v_{4,3}$. Всего в системе может происходить конечное число типов событий с разными составами участников. Так объекты типов t_2 и t_5 участвуют только в событии e_1 , а объекты типов t_1 и t_3 участвуют только в событии e_2 . В то же время объекты типа t_4 принимают участие и в e_1 и в e_2 . В данном случае каждое из событий e_1 и e_2 также может рассматриваться как самостоятельный тип события, поскольку в них принимают участие разные составы объектов. Для типов событий может быть построена бинарная матрица формального контекста. На рисунке 2 а) представлен вариант такой матрицы для следующего набора событий предметной области. В деятельности предприятия принимают участие: t_1 – руководители, t_2 – заказчики, t_3 – менеджеры, t_4 – документы (договора, задания), t_5 – сотрудники, t_6 – документы (акты приемки). При этом имеют место следующие события: e_1 – подготовка и заключение договоров, e_2 – выдача заданий, e_3 – выполнение работ, e_4 – промежуточные испытания, e_5 – приемка работы. Таким образом, средствами ФСА задача построения схемы иерархии классов разрешается в виде концептуальной ре-

шетки по таблице, представленной на Рис. 2 а). Для рассмотренного случая эта решетка имеет вид, представленный на Рис. 2 б).

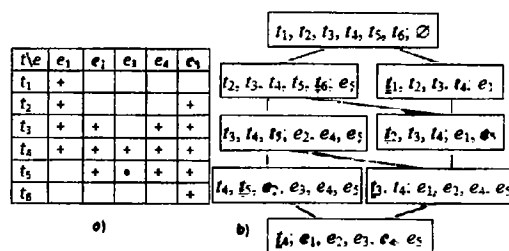


Рис. 2 – Пример концептуальной решетки.

Эта схема отличается от того, что представляет нормализованная схема базы данных SQL. Если по аналогии с традиционными базами данных рассматривать каждый узел данной решетки как тип таблицы, то можно заметить, что здесь в одной таблице присутствуют и наборы объектов и наборы событий. При этом, избыточность устраняется, если маркировать узлы тем типом, для которого в этом узле зафиксировано наибольшее число событий. На диаграмме такие типы помечены двойным подчеркиванием. В результате для каждого типа будет определен один единственный узел со всем перечнем связанных с ним событий. Такая база данных будет объектно-ориентированной. При этом с каждым объектом будет связан набор его атрибутов-свойств, которые изменяются в процессе его функционирования в контексте системы, а также условия и события, при которых происходит изменение значений атрибутов. Таким образом, получена модель организации данных, которая сопоставима с типовой схемой нормализованных баз данных, но она ее не заменяет, а имеет свою область применения в задачах представления объектных взаимодействий. Следует отметить, что в данной работе представлен только один из вариантов использования методологии ФСА для отношения «объект – событие». В то же время ФСА позволяет получать достаточно широкую гамму схем, отвечающих различным планам исследования. Это позволяет типовым образом строить на одном и том же фактическом материале различные классификации. Каждая такая классификация обслуживает свой круг задач, но работа с ними может быть построена на единой методологической, а, следовательно, и инструментальной платформе.

II. СПИСОК ЛИТЕРАТУРЫ

1. Емельяненко, В. И. Архитектура распределенной системы мониторинга Сетевые компьютерные технологии /В. И. Емельяненко // (NTECH'05): Материалы II Международной научной конференции (Минск, 21-23 сентября 2005 г.), Стр 102-106.
2. Ganter, B. Formal Concept Analysis: Foundations and Applications /B. Ganter, G. Stumme, R. Wille //Springer, 2005.