

Обзор способов построения темпоральных систем на основе реляционной базы данных

08, август 2012

DOI: 10.7463/0812.0441884

Балдин А. В., Елисеев Д. В., Агаян К. Г.

УДК. 004.652

Россия, МГТУ им. Н.Э. Баумана

d-eli@mail.ru

Введение

Нетемпоральные модели данных и их реализация в системах управления базами данных (СУБД) хранят единственное состояние объектов предметной области. СУБД поддерживают операции изменения, которые переводят базу данных из одного состояния в другое, тем самым, заменяя старые значения новыми. В таких системах считается, что все хранимые в базе данные являются актуальными в момент выполнения соответствующих запросов.

Но существует множество предметных областей, в которых необходимо хранить прошлые состояния базы данных и возможно будущие. Примерами информационных систем, функционирующих в таких предметных областях, являются: система управления предприятием [1, 2], система управления персоналом [3, 4], финансовые приложения, страховые приложения и ряд других, которые обладают следующими свойствами:

1. Информационная система обрабатывает темпоральные (изменяющиеся во времени) данные.
2. Информационная система накапливает историю изменения темпоральных данных.

В настоящее время реляционная модель является наиболее распространённой в базах данных. В разработку методов хранения и обработки темпоральных данных на базе реляционной модели внесли вклад учёные, такие как Р. Снодграс [5, 10], К. Дженсен [5, 11], Дж. Бен-Зви [5], С. Гадия [5], Е. МакКензи [5, 13], А. Стейнер [12]. Исследования в области темпоральных баз данных привели к многообразию соответствующих темпоральных моделей. В статье приведены подходы к моделированию времени в темпоральных моделях, а также подходы к их реализации на базе реляционных СУБД.

1. Темпоральные данные и темпоральные базы данных

СУБД основывается на модели данных, которая определяет конструкции и формализмы, доступные для описания, изменения и доступа к данным. Модель данных

$M=(DS, OP, C)$ состоит из трёх компонент: структуры данных DS , операций OP и ограничений целостности C [8, 9, 12]. Модель данных позволяет описывать объекты предметной области и выполнять над ними соответствующие операции.

Совокупность значений, описывающих объект предметной области в определённый период времени называется **состоянием объекта предметной области**. Совокупность состояний объектов предметной области в определённый период времени называется **состоянием базы данных**.

Современные СУБД обрабатывают постоянные данные, т.е. данные существуют дольше, чем приложения выполняют задачи. Одной из причин, по которой управление изменяющихся во времени данных не рассматривается для большинства приложений, является отсутствие соответствующей поддержки современными СУБД.

Темпоральные данные – произвольные данные, которые связаны с определёнными датами или промежутками времени [6].

В вышеуказанных предметных областях различные состояния базы данных сохраняются как темпоральные данные. Темпоральные модели данных позволяют хранить информацию об эволюции объектов: для любого объекта, который был создан в момент времени t_1 и закончил свое существование в момент времени t_2 , в базе данных будут сохранены все его состояния на временном интервале $[t_1, t_2)$ [7].

Темпоральная СУБД – это СУБД для хранения и обработки темпоральных данных [6].

СУБД и содержащиеся в них данные могут рассматриваться как темпоральные только в том случае, если известно правило интерпретации временных меток и интервалов. В категорию темпоральных СУБД не попадают обычные реляционные СУБД, в которых поддерживаются связанные со временем типы данных, но интерпретацией и связью данных между собой с учетом времени приходится заниматься разработчику приложения [6]. В темпоральной СУБД учитывается изменчивость данных с течением времени.

2. Представление времени в темпоральных моделях данных

Темпоральная модель данных на базе реляционной расширяет конструкции последней, чтобы описывать и обрабатывать темпоральные данные. Принципиальные отличия между темпоральными моделями данных на базе реляционной зависят от интерпретации ключевых понятий, используемых в этих моделях. К таким ключевым понятиям, в частности, относятся: представление времени и тип отметки времени [12].

2.1. Представление времени в темпоральных моделях данных

Время в темпоральных моделях обычно преобразовывается к набору чисел, которые целиком упорядочены относительно предикатов сравнения. Таким образом, время может быть представлено как ось в системе координат. В темпоральных моделях данных время может рассматриваться как непрерывное или дискретное [12].

Непрерывная модель рассматривает время как изоморфное к вещественным числам. Каждое вещественное число соответствует точке на оси времени. Таким образом, в непрерывной модели времени между любыми двумя временными точками существует другая временная точка:

$$\forall t_1, t_2, t_1 < t_2, \exists t_3 : t_1 < t_3 < t_2.$$

Этот подход моделирует время наиболее точно. Так как темпоральные модели используются в базах данных, то невозможно преобразовать эту модель времени без потерь к существующим типам данных.

Дискретная модель преобразует время к целым числам [12]. В дискретной модели можно определить функцию поиска следующего момента времени δ . Между точкой на оси времени и следующей за ней точкой других временных точек не существует.

$$\forall t_1, t_2, \text{ если } t_2 = \delta(t_1), \text{ то не } \exists t_3 : t_1 < t_3 < t_2.$$

2.2 Тип временной отметки

Временная ось может рассматриваться как набор временных точек. Наименьшая неразложимая временная единица на временной оси называется **квантом времени** [12]. Таким образом, отдельный квант времени является подынтервалом фиксированной продолжительности на временной оси. Наименьшая неразложимая временная единица необходима, так как цифровые компьютеры поддерживают только ограниченную степень детализации для вещественных чисел. Для непрерывной модели времени, квант времени охватывает небольшую часть временной оси, тогда как временная точка представляет единственную точку на этой оси. Для дискретной модели кванты времени и временные точки определяются аналогично.

Пусть задан момент времени t , который соответствует точке на оси времени. Тогда событие является мгновенным фактом, который происходит в момент времени t [12]. При рассмотрении более укрупнённой степени детализации временных единиц, соответствующих квантам времени, если событие происходит в квант времени qt , то оно происходит в любой момент времени t в течение кванта времени qt : $t \in qt$.

Темпоральная база данных хранит факты и время, когда они были действительны в предметной области. С теоретической точки зрения это означает, что все моменты времени, когда факт является истинным, должны быть сохранены в темпоральной базе данных. В случае, когда это множество моментов времени образует плотное подмножество, оно может быть записано как интервал времени. Интервалы времени используются, чтобы моделировать период времени, в течение которого факт был истинным в предметной области [12].

Интервал времени представляет период времени, имеющий начальный момент времени t_1 и конечный момент времени t_2 . Моменты времени t_1 и t_2 являются наименьшим и наибольшим значением временного интервала. Нижняя и верхняя границы могут сравниваться с каждым другим моментом времени предикатами сравнения. Интервал времени рассматривается как $[t_1, t_2)$, в который включается нижняя граница t_1 и не включается верхняя граница t_2 [12].

Для интервалов определены набор теоретических операций объединения, пересечения и разности [12]. Это множество операций необходимо, чтобы делать выводы о темпоральных фактах. Однако, рассматриваемые операции не являются замкнутыми относительно интервалов времени. Объединение двух непересекающихся интервалов возвращает набор интервалов. Разность двух интервалов может вернуть пустое множество, один или два интервала.

Пример 1. < Пусть даны три интервала: $I_1=[2000, 2006)$, $I_2=[2008, 2011)$, $I_3=[1995, 2012)$. Тогда

$$\begin{aligned} I_1 \cup I_2 &= \{[2000, 2006), [2008, 2011)\}; \\ I_1 \setminus I_3 &= \emptyset; \\ I_3 \setminus I_2 &= \{[1995, 2008), [2011, 2012)\}. \end{aligned}$$

Также для интервалов можно определить логические операции включения и сравнения, такие как предшествование интервалов, пересечение интервалов, интервал непосредственно следует за другим, интервал содержит в себе другой интервал, равенство интервалов и так далее.

Как правило, реляционная модель может описывать темпоральные данные либо с помощью момента времени, либо с помощью временного интервала. Темпоральные реляционные модели данных, остающиеся в первой нормальной форме (1НФ), ограничены использованием отметок времени, которые можно преобразовать к дополнительным скалярным значениям атрибутов кортежа.

При описании темпоральных данных с помощью момента времени предполагается, что данные являются истинными только в заданный момент времени. Отношения, содержащие темпоральные данные с моментом времени называют **таблицами событий** [14]. Темпоральная реляционная модель с моментом времени может быть смоделирована посредством расширения схемы с неявным атрибутом T типа «Дата». При этом отношение остаётся в 1НФ. Если $R=(A_1, A_2, \dots, A_n)$ – схема нетемпорального отношения, то таблица событий для этого отношения будет иметь следующую схему: $R_t=(T, A_1, A_2, \dots, A_n)$.

При описании темпоральных данных с помощью интервала времени предполагается, что данные являются истинными в течение заданного интервала времени. Отношения, содержащие темпоральные данные с интервалом времени называют **таблицами состояний** [12]. Темпоральность с временными интервалами может использоваться с отношениями в 1НФ, если нижняя и верхняя граница временного интервала преобразуются в два дополнительных атрибута T_1 и T_2 типа «Дата». Если $R=(A_1, A_2, \dots, A_n)$ – схема нетемпорального отношения, то таблица состояний для этого отношения будет иметь следующую схему: $R_t=(T_1, T_2, A_1, A_2, \dots, A_n)$.

Записи в отношении, которые являются истинными в течение нескольких неперекрывающихся периодов времени, хранятся отдельно для каждого периода времени. Так как интервалы времени не замкнуты относительно ряда теоретико-множественных операций, таких как разность и объединение, то история объектов предметной области может распространяться на несколько записей как при хранении отношений, так и при вычислении результата запроса. Пример таблицы состояний представлен на рис. 1, на котором приведена часть отношения «Работник».

Id	Имя	Зарплата	Подразделение	T1	T2
1	Иван	20000	Отдел 1	01.01.2000	30.11.2005
1	Иван	25000	Отдел 2	01.12.2005	31.08.2010
1	Иван	30000	Отдел 1	01.09.2010	
2	Пётр	20000	Отдел 2	02.03.2004	06.07.2008
2	Пётр	20000	Отдел 2	07.07.2008	31.12.2011

Рис. 1. Отношение «Работник» с временным интервалом

В отношении «Работник» хранятся данные о двух работниках, действительные в разные неперекрывающиеся временные интервалы.

3. Способы реализации темпоральных баз данных

Изменение или расширение модели данных для работы с темпоральными объектами, означает, что реализация модели, т.е. СУБД, также должна быть изменена. Так как большинство СУБД рассматриваются как чёрные ящики, то соответствующее изменение программного обеспечения может быть выполнено только разработчиками СУБД. Другой подход обработки темпоральных данных заключается во встраивании темпоральной функциональности в приложение или использование расширяемости, присущей некоторым СУБД. Фактически, можно выделить два различных пути реализации темпоральных реляционных баз данных:

1. Реализация темпоральной поддержки на уровне приложения.
2. Расширение нетемпоральной модели данных до темпоральной [5].

Первый подход не подразумевает каких-либо изменений СУБД. Система берётся как есть, а работа с темпоральными данными реализуется программистами приложения. Последний подход связан с доработкой реляционной СУБД.

3.1. Реализация темпоральной поддержки на уровне приложения

В этом подходе единственным инструментом является использование типа «Дата», поддерживаемого непосредственно СУБД, с помощью которого описываются временные метки хранящихся данных. Любая темпоральная семантика должна быть встроена в приложение программистом самостоятельно. Темпоральные операции OP^T и темпоральные ограничения целостности C^T также должны поддерживаться в самом приложении [5, 12]. Рассматриваемый подход представлен на рис. 2.

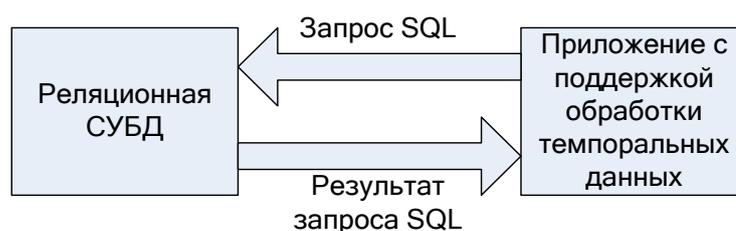


Рис. 2. Схема реализация работы с темпоральными данными на уровне приложения.

Данный подход означает, что дополнительная нагрузка ложится на программиста, так как отсутствует специальная поддержка со стороны СУБД. Этот подход является ненадёжным и требует много времени для дополнительной реализации темпоральной оболочки, используемой приложением. Эта часть кода обычно заново изобретается каждой компанией, которая имеет дело с темпоральными данными. В этом случае отсутствует единый стандартизированный подход к реализации обработки темпоральных данных, что приводит к неэффективным, сложным в администрировании приложениям и проблемам при их доработке.

3.2. *Расширение нетемпоральной модели данных до темпоральной*

Расширение нетемпоральной модели данных до темпоральной означает, что понятия, уже поддерживаемые нетемпоральными моделями, используются для определения темпорального расширения. Новые понятия вводятся, когда трудно или невозможно выразить их в нетемпоральной модели [5, 12].

Например, темпоральные данные можно описать, расширив нетемпоральную схему или типы специальными временными атрибутами. Языки запросов и алгебры расширяют дополнительными операциями для обработки темпоральных данных. В результате для темпоральных моделей данных определяют темпоральную алгебру, которая ссылается на специальные временные атрибуты для темпоральных вычислений. Этот подход расширения схем и языков запросов можно применять для реляционной СУБД (рис. 3).

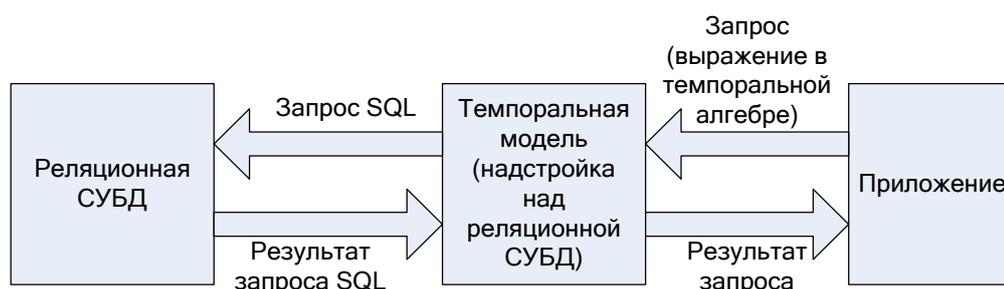


Рис. 3. Схема расширения реляционной модели данных до темпоральной

В схеме на рис. 3 темпоральность реляционных структур данных достигается за счёт связывания их с соответствующими отметками времени. При этом, темпоральная модель выступает как некоторая надстройка над реляционной, которая использует понятия и конструкции последней.

Фактически, этот подход является самым широко используемым на практике для реализации темпоральных моделей данных. Преимуществом является то, что только часть нетемпоральной модели должна быть изменена, например, язык запросов и ограничения целостности. Относительно существующей СУБД, это означает, что часть её также должна быть изменена. Методы доступа и структуры хранения данных, например, не затрагиваются.

К недостаткам этого подхода относится то, что повторное использование существующих понятий нетемпоральной модели данных, приводит к тому, что темпоральная модель автоматически наследует ограничения исходной модели. Добавляя дополнительные атрибуты к отношениям, только специальные структуры данных, например, кортежи, могут быть темпоральными. Так как конструкции нетемпоральной

модели данных могут также изменяться со временем, например набор отношений и схем отношений, то хранение и обработку истории развития каждой из этих конструкций необходимо выполнять в темпоральных системах, что при рассматриваемом подходе реализовать невозможно.

Заключение

В статье проведён анализ подходов к моделированию времени и способов описания времени на основе реляционных баз данных. Так как темпоральные базы данных, в отличие от нетемпоральных, сохраняют историю изменения состояний объектов предметной области, с которой связано выполнение темпоральных запросов, то рассмотрены подходы и способы реализации таких темпоральных систем. Наиболее используемым на практике является подход расширения реляционной модели для обработки темпоральных данных. Существенным достоинством данного подхода является то, что используются хорошо проработанные механизмы современных СУБД, которые не надо заново реализовывать для темпоральной СУБД.

Таким образом, рассмотренные подходы к моделированию времени и способы описания темпоральных данных на основе реляционной модели позволяют использовать современные реляционные СУБД при разработке темпоральных систем.

Список литературы

1. Информационная управляющая система МГТУ им. Н.Э. Баумана «Электронный университет»: концепция и реализация / Т.И. Агеева, А.В. Балдин, В.А. Барышников и др.; [под ред. И.Б. Фёдорова, В.М. Черненко]. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2009. – 376 с.
2. Информационные технологии в инженерном образовании / под ред. С.В. Коршунова, В.Н. Гузнецова. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2007. – 432 с.
3. Куприянова С.В. «1С: Зарплата и Кадры Бюджетного Учреждения 8» - мощный инструмент по автоматизации расчёта заработной платы в учебных заведениях // Новые информационные технологии в образовании: Сб. науч. тр. десятой международной научно-практической конференции: Москва. - 2010. – Часть 1. - С 311 – 315.
4. «Персонал – Кадры» – адаптивная функционально-развиваемая система информационной поддержки управления кадрами [Электронный ресурс] - Режим доступа: <http://www.personal.bravosoft.ru/doc/doc002.htm>, (Дата обращения: 09.04.2012)
5. Исследование и разработка темпоральной модели данных в рамках МИС Интерин PROMIS [Электронный ресурс]/ Базаркин А. Н., 2009 - Режим доступа: https://edu.botik.ru/proceedings/psta2009/v2/037-Bazarkin.Research_and.pdf, (Дата обращения: 09.04.2012)
6. История и актуальные проблемы темпоральных баз данных [Электронный ресурс] / Костенко Б.Б., Кузнецов С.Д., 2007 – Режим доступа: <http://citforum.ru/database/articles/temporal/4.shtml>, (Дата обращения: 14.08.2011)
7. Варламов О.О. Эволюционные базы данных и знаний для адаптивного синтеза интеллектуальных систем. Миварное информационное пространство. - М: Радио и связь, 2002.-286 с.
8. Дейт К. Введение в системы баз данных. - 8-е изд. – М.: Вильямс, 2006. – 1328 с.
9. Гарсиа-Молина, Гектор, Ульман, Джеффри, Д., Уидом, Дженнифер Г Системы баз данных. Полный курс. : Пер. с англ. — М.: Издательский дом "Вильямс", 2003. — 1088 с.

10. Developing Time-Oriented Database Applications in SQL, Richard T. Snodgrass, Morgan Kaufmann Publishers, Inc., San Francisco, July, 1999, 504 pages.
11. Jensen C. S. and R. T. Snodgrass, Semantics of Time-Varying Information // Information Systems, Vol. 21, No. 4, 1996, pp. 311–352. [Электронный ресурс], Режим доступа: <http://www.cs.arizona.edu/~rts/pubs/ISJune96.pdf>, свободный, (Дата обращения: 09.04.2012)
12. Andreas Steiner. A Generalisation Approach to Temporal Data Models and their Implementations [Электронный ресурс], 1998 - Режим доступа: <http://www.timeconsult.com/Publications/diss.pdf>, (Дата обращения: 09.04.2012)
13. McKenzie E., Snodgrass R.T. An evaluation of relational algebras incorporating the time dimension in databases [Электронный ресурс] // ACM computing surveys. - 1991. - 23(4). - 501–543 p. - Режим доступа: <http://www.cs.arizona.edu/~rts/pubs/CompSurv91.pdf> (Дата обращения: 09.04.2012).

Overview of methods of creating systems based on temporal relational database

08, August 2012

DOI: 10.7463/0812.0441884

Baldin A.V., Eliseev D., V., Agayan K.G.

Russia, Bauman Moscow State Technical University
d-eli@mail.ru

The article presents an analysis of approaches to implementation of temporal systems based on relational database. In this paper the authors give definitions of temporal data and temporal databases. The authors present approaches to time modeling on the basis of which the types of time stamps used in temporal databases are described. The paper identifies two main ways to implement temporal systems using current relational database: implementation of temporal support at the application level and expansion of the non-temporal relational model to the temporal one. As a result, for each approach the authors identified respective advantages and disadvantages.

Publications with keywords: [temporal database](#), [relational data model](#), [timestamping data](#), [modeling time](#), [time instant](#), [time interval](#)

Publications with words: [temporal database](#), [relational data model](#), [timestamping data](#), [modeling time](#), [time instant](#), [time interval](#)

References

1. Ageeva T.I., Baldin A.V., Baryshnikov V.A., et. al. *Informatsionnaia upravliaiushchaia sistema MGTU im. N.E. Baumana «Elektronnyi universitet»: kontseptsii i realizatsiia* [Information management system of the Bauman MSTU "Electronic University": the concept and the realization of]. Moscow, Bauman MSTU Publ., 2009. 376 p.
2. Korshunov S.V., Guznenkov V.N. *Informatsionnye tekhnologii v inzhenernom obrazovanii* [Informational technology in engineering education]. Moscow, Bauman MSTU Publ., 2007. 432 p.
3. Kupriianova S.V. «1S: Zarplata i Kadry Biudzhetnogo Uchrezhdeniia 8» - moshchnyi instrument po avtomatizatsii rascheta zarabotnoi platy v uchebnykh zavedeniiaxh [«1C: Wages and Personnel of the Budget Institution 8» - a powerful tool for the automation of the calculation of wages in the educational institutions]. *Novye informatsionnye tekhnologii v*

obrazovanii: Sb. nauch. tr. desiatoi mezhdunarodnoi nauchno-prakticheskoi konferentsii [New information technologies in education: collection of scientific works of the 10th international scientific-practical conference]. Moscow, 2010, part 1, pp. 311– 315.

4. «Personal – Kadry» – *adaptivnaia funktsional'no-razvivaemaia sistema informatsionnoi podderzhki upravleniia kadrami* [«Cadres - Personnel» - adaptive functional-developed system of information support of personnel management]. Available at: <http://www.personal.bravosoft.ru/doc/doc002.htm> , accessed 09.04.2012.

5. Bazarkin A.N. *Issledovanie i razrabotka temporal'noi modeli dannykh v ramkakh MIS Interin PROMIS* [Research and development of temporal data model in the IIA Interin PROMIS]. 2009. Available at: https://edu.botik.ru/proceedings/psta2009/v2/037-Bazarkin.Research_and.pdf , accessed 09.04.2012.

6. Kostenko B.B., Kuznetsov S.D. *Istoriia i aktual'nye problemy temporal'nykh baz dannykh* [History and current problems of temporal databases]. 2007. Available at: <http://citforum.ru/database/articles/temporal/4.shtml> , accessed 14.08.2011.

7. Varlamov O.O. *Evoliutsionnye bazy dannykh i znaniia dlia adaptivnogo sinteza intellektual'nykh sistem. Mivarnoe informatsionnoe prostranstvo* [Evolutionary knowledge and information base for adaptive synthesis of intelligent systems. Developing multi-dimensional information space]. Moscow, Radio i sviaz', 2002. 286 p.

8. Date C.J. *An introduction to database systems*. 8th ed. Addison-Wesley, 2003. (Russ. ed.: Deit K.Dzh. *Vvedenie v sistemy baz dannykh*. 8th ed. Moscow, Vil'iams, 2006. 1328 p.).

9. García-Molina H., Ullman J.D., Widom J. *Database Systems: The Complete Book*. New Jersey, Pearson Prentice Hall, 2002. (Russ. ed.: Garsia-Molina G., Ul'man Dzh., Uidom Dzh. *Sistemy baz dannykh. Polnyi kurs*. Moscow, Izdatel'skii dom "Vil'iams", 2003. 1088 p.).

10. Snodgrass R. *Developing Time-Oriented Database Applications in SQL*. Morgan Kaufmann Publishers, Inc., San Francisco, July, 1999. 504 p.

11. Jensen C.S., Snodgrass R.T. Semantics of Time-Varying Information. *Information Systems*, 1996, vol. 21, no. 4, pp. 311–352. Available at: <http://www.cs.arizona.edu/~rts/pubs/ISJune96.pdf> , accessed 09.04.2012.

12. Steiner A. *A Generalisation Approach to Temporal Data Models and their Implementations*. 1998. Available at: <http://www.timeconsult.com/Publications/diss.pdf> , accessed 09.04.2012.

13. McKenzie E., Snodgrass R.T. An evaluation of relational algebras incorporating the time dimension in databases. *ACM Computing Surveys*, 1991, vol. 23, no. 4, pp. 501–543. Available at: <http://www.cs.arizona.edu/~rts/pubs/CompSurv91.pdf> , accessed 09.04.2012.