

**КЛАССИФИКАЦИОННАЯ СХЕМА МОДЕЛЕЙ БАЗ ДАННЫХ
ДЛЯ ЛАЗЕРНЫХ ТЕХНОЛОГИЧЕСКИХ КОМПЛЕКСОВ**

Нечипоренко О.В., к.т.н., доцент,
доцент кафедры специализированных компьютерных систем
Миценко С.А., аспирант кафедры специализированных компьютерных систем
Черкасский государственный технологический университет
бул. Шевченко, 460, г. Черкассы, Украина, 18006
kafedra_ckc@mail.ru

***Аннотация.** В статье приведены результаты исследования характеристик основных моделей баз данных, проведена их сравнительная оценка по критериям для определения оптимальной модели баз данных для лазерных технологических комплексов.*

***Ключевые слова:** модель базы данных, сравнительная оценка моделей баз данных, интегрированная модель, оптимальная модель, лазерный технологический комплекс.*

**CLASSIFICATION SCHEME OF MODELS OF DATABASES
FOR LASER TECHNOLOGICAL COMPLEXES**

Nechyporenko O.V., Ph.D. (Engineering),
associate professor of the department of specialized computer systems,
Mitsenko S.A., post-graduate student, department of specialized computer systems
Cherkasy State Technological University
Shevchenko boulevard, 460, Cherkassy, Ukraine, 18006
kafedra_ckc@mail.ru

***Abstract.** The results of research of basic models of databases characteristics have been outlined. Their comparative estimation on criteria for determination of optimal model of database for laser technological complexes has been conducted.*

***Keywords:** model of database, comparative estimation of models of databases, integrated model, optimal model, laser technological complex.*

Введение. Для обеспечения эффективности информационных процессов необходима соответствующая организация данных. Известен ряд различных способов организации данных. Проблематика моделирования связана с таким представлением данных, которое наиболее естественно отражает реальный мир и может поддерживаться компьютерными средствами. Важной проблемой, решаемой при проектировании баз данных (БД), является создание оптимальной модели БД, которая бы обеспечивала минимальное дублирование информации и упрощала процедуры обработки и обновления данных.

Для того чтобы определить наилучшую (с точки зрения конкретного приложения или задачи) организацию данных, необходимо установить, какие характеристики данных важны для выражения сущности их значения. Это позволяет провести оценку и сформировать непротиворечивый набор общих формальных утверждений, характеризующих организацию и обработку данных, который и определяет оптимальную модель БД для лазерных технологических комплексов (ЛТК).

Цель работы – построение классификационной схемы моделей БД и исследование их характеристик, проведение их сравнительной качественной оценки и анализа по критериям для определения оптимальной модели БД ЛТК.

Результаты исследований.

1. **Основные понятия и определения**

Понятие «модель данных» является одним из фундаментальных в области БД, хотя в известной литературе нередко дается его неоднозначная трактовка.

Первоначально понятие *модели данных* употреблялось как синоним *структуры данных в конкретной БД*. В процессе развития термин «модель данных» приобрел новое содержание и в настоящее время в научной литературе трактуется в подавляющем большинстве случаев как инструмент моделирования, т.е. совокупность правил структурирования данных, допустимых операций над ними и видов ограничения целостности, которым они должны удовлетворять [1].

Определение 1. Модель данных – это абстрактное, самодостаточное, логическое определение объектов, операторов и прочих элементов, в совокупности составляющих абстрактную машину доступа к данным, с которой взаимодействует пользователь [2].

Определение 2. Модель данных – это набор принципов, определяющих организацию логической структуры хранения данных в базе [3].

Определение 3. Модель данных – интегрированный набор понятий для описания и обработки данных, связей между ними и ограничений, накладываемых на данные в некоторой организации [4].

2. Основные виды моделей БД и их характеристики

Разнообразие моделей данных соответствует разнообразию областей применения и контингента пользователей.

Различают четыре основные логические модели данных: иерархическая; сетевая; реляционная; многомерная.

С точки зрения структур данных, использующихся для связи между собой внутренних объектов базы, логические модели можно объединить в две группы:

- навигационные модели данных с адресными указателями на данные;
- ссылочные модели данных с именами (идентификаторами) данных.

К навигационным моделям данных относятся иерархическая, сетевая и многомерная логические модели, к ссылочным моделям данных – реляционная и постреляционная модели.

Сначала использовались иерархические логические модели. Простота организации, наличие заранее заданных связей между сущностями, сходство с физическими моделями данных позволяли добиваться приемлемой производительности иерархических систем управления базами данных (СУБД) на медленных ЭВМ с весьма ограниченными объемами памяти. Но, если данные не имели древовидной структуры, то возникала масса сложностей при построении иерархической модели и желании добиться нужной производительности.

Сетевые модели также создавались для малоресурсных ЭВМ. При создании сетевых моделей было разработано множество способов, позволяющих увеличить производительность СУБД, но существенно усложнивших последние.

Сложность практического использования иерархических и сетевых СУБД заставляла искать иные способы представления данных. Э.Ф. Кодд обосновал, что более универсальным решением в области баз данных является ссылочная модель [2]. Однако теоретическая модель Э.Ф. Кодда характеризуется отсутствием в её составе какого-либо механизма поиска данных, за исключением полного перебора данных в базе. При достаточно большом количестве данных в реляционную базу в обязательном порядке включается тот или иной механизм поиска, т.е. дополнительная структура, отсутствие потребности в которой декларировал Э.Ф. Кодд.

Сегодня реляционная модель данных наиболее распространена. Хотя наряду с общепризнанными достоинствами она обладает и рядом недостатков. Классические реляционные СУБД идеально подходят для традиционных приложений, но их применение в системах, основанных на знаниях, является затруднительным. В нетрадиционных приложениях в БД появляется множество таблиц, над которыми постоянно выполняются дорогостоящие операции соединения, необходимые для воссоздания сложных структур данных.

Другим серьезным ограничением реляционных систем являются их относительно слабые возможности по части представления семантики приложения. Самое большее, что обеспечивают реляционные СУБД – это возможность формулирования и поддержки ограничений целостности данных.

Кризис в сфере практического применения реляционных СУБД нарастает. До 80 % созданных корпоративных хранилищ данных не решают полностью поставленных перед ними задач, а 40 % являются проваленными проектами. Около 50 % запросов пользователей являются не предусмотренными в ходе их проектирования. Использование реляционных СУБД неэффективно при количестве заложенных в них отношений (таблиц, информационных объектов) свы-

ше 100–300 [5]. Причина неэффективности СУБД состоит в противоречии между результатами проектирования БД, отражающими статичное состояние предметной области окружающего мира и реальными взаимодействиями пользователей с БД, отражающими состояние предметной области в ее динамике.

Последней логической моделью данных является многомерная модель, отражающая в полной мере естественную структуру данных из любой предметной области, в отличие от иерархической, сетевой и реляционной моделей.

Классическая многомерная модель сочетает в себе первичную структуру хранения информации и механизм поиска. В качестве первичных структур хранения информации используются ячейки или области ячеек многомерного пространства. Поликубический вариант получил наибольшее распространение в силу удобства оперирования отдельными подмножествами многомерной БД, в том числе, при обращении к энергонезависимому носителю информации.

Единственный недостаток классической многомерной модели данных – «взрывной» рост ячеек многомерного пространства при увеличении числа измерений и количества координат в составе измерений. Объем многомерного пространства при этом возрастает по экспоненте.

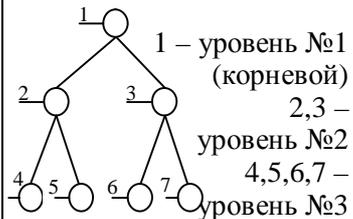
К числу новых моделей данных относится объектно-ориентированная модель. В объектно-ориентированных СУБД ставится задача создания модели данных, позволяющей отражать сложные аспекты объектов и связей предметного мира, для чего создаются не менее сложные структуры их физического хранения, манипулирования и поддержания целостности. Теоретическая модель для объектно-ориентированных СУБД на сегодняшний день полностью не разработана, поэтому крупнейшие разработчики СУБД стали встраивать в свои продукты поддержку объектной ориентации, предлагая смешанный подход – объектно-реляционный.

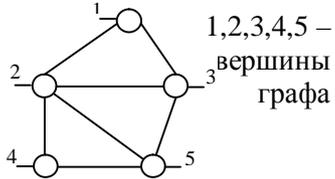
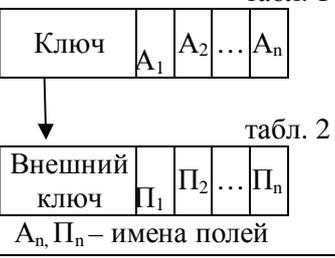
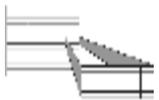
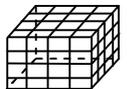
Обобщая вышесказанное, в табл. 1 приведены преимущества и недостатки классических и неклассических моделей данных, относящихся к навигационным и ссылочным моделям.

Таблица 1

Сравнительная качественная оценка моделей базы данных

Модели базы данных		Преимущества	Недостатки
1		2	3
Классические	Ранние		
	Основанная на инвертированных списках	<ul style="list-style-type: none"> – развитые средства управления данными во внешней памяти на низком уровне – возможность построения вручную эффективных прикладных систем 	<ul style="list-style-type: none"> – сложность в использовании – необходимы знания о физической организации – зависимость прикладных систем
	Прототипная	<ul style="list-style-type: none"> – возможность экономии памяти за счет разделения подобъектов – активно использовались в течение многих лет 	<ul style="list-style-type: none"> – логика перегружена деталями организации доступа к БД – невозможность использования совместно с современными моделями
	Ассоциативная	<ul style="list-style-type: none"> – не основывались на каких-либо абстрактных моделях – доступ к БД производился на уровне записей 	<ul style="list-style-type: none"> – пользователь сам производит всю оптимизацию доступа к БД
	Иерархическая связи между данными представляются с помощью упорядоченного графа.	<ul style="list-style-type: none"> – простота – иерархический принцип соподчиненности – эффективное использование памяти – быстродействие при выполнении основных операций над данными 	<ul style="list-style-type: none"> – большой объем для обработки информации с сложными связями; – сложность понимания для обычного пользователя – невозможность реализовать много взаимосвязей данных – допустимость только навигационного принципа доступа к данным.



	1	2	3
Неклассические	<p><i>Сетевая</i> разнообразные взаимосвязи между данными представляются в виде произвольного графа</p>  <p>1,2,3,4,5 – вершины графа</p>	<ul style="list-style-type: none"> – эффективная реализация по показателям затрат памяти и оперативности. – универсальность – возможности модели данных наиболее широкие в сравнении с остальными моделями. – возможность доступа к данным через значения нескольких отношений 	<ul style="list-style-type: none"> – высокая сложность схемы базы данных – сложность для понимания и выполнения обработки информации обычным пользователем – слабый контроль целостности связей – допустимость только навигационного принципа доступа к данным.
	<p><i>Реляционная</i> представление взаимосвязей в виде таблиц и операции над таблицами табл. 1</p>  <p>табл. 2 An, Пn – имена полей</p>	<ul style="list-style-type: none"> – простота модели, понятность для пользователя – удобство физической реализации – независимость данных и программ – наличие теоретически обоснованных методов нормализации отношений – действия над данными сводятся к операциям реляционной алгебры 	<ul style="list-style-type: none"> – большой объем памяти для представления базы данных – сложность описания связей – низкая скорость при выполнении операции соединения – отсутствие стандартных средств идентификации отдельных записей
	<p><i>Постреляционная</i> расширенная реляционная модель, снимающая ограничение неделимости данных в записях таблиц</p> 	<ul style="list-style-type: none"> – высокая наглядность представления информации – высокая эффективность обработки данных – возможность формирования совокупности связанных реляционных таблиц через одну постреляционную. 	<ul style="list-style-type: none"> – сложность решения проблемы обеспечения целостности и непротиворечивости хранимых данных.
	<p><i>Многомерная</i> узкоспециализированная, предназначенная для интерактивной аналитической обработки</p> 	<ul style="list-style-type: none"> – удобство и эффективность аналитической обработки больших объемов данных, связанных со временем – многомерная организация данных имеет высокую наглядность и информативность. 	<ul style="list-style-type: none"> – большой объем памяти для простейших задач обычной оперативной обработки информации.
	<p><i>Объектно-ориентированная</i> представление данных позволяет идентифицировать отдельные записи</p>	<ul style="list-style-type: none"> – возможность отображения информации о сложных взаимосвязях объектов – позволяет идентифицировать отдельную запись БД – возможность описывать в рамках единого информационного поля объекты, имеющие разнородную внутреннюю структуру и состав элементов. 	<ul style="list-style-type: none"> – необходимость большого времени для освоения обработки данных – сложность понимания для обычного пользователя – неудобство обработки данных при построении БД – низкая скорость выполнения запросов.

Становится интересным рассмотрение вопросов, посвященных разработке модели БД с учетом предметной области, когда сама структура БД не зависит от рассматриваемой предметной области.

Проблема создания стандартной модели данных обсуждалась на международных симпозиумах, начиная с 1988 года, однако и сегодня эта проблема остается актуальной. Некоторые придерживаются мнения о возможности существования такой супермодели данных, которая, если бы ее удалось построить, разрешила бы все проблемы моделирования.

В работе [6] предлагается подход к созданию БД на основе использования универсальной модели данных, позволяющей представлять данные любой предметной области в универсальном виде и дающей возможность осуществлять простой переход к компьютерной реализации.

Универсальная (стандартная) модель данных основывается на семантической модели данных «объект-событие», теории множеств и логическом исчислении.

3. Оценка моделей данных и выбор оптимальной модели для БД ЛТК

Оценка модели данных определяется ее полезностью с точки зрения формирования взглядов на организацию и использование данных. Полезность модели зависит от степени ее адекватности моделируемой предметной области.

Все модели данных равносильны, т.е. все, выразимое в одной из них, выразимо в остальных. Различие составляет то, насколько удобно использовать модель для работы с реальными задачами, и то, насколько эффективно можно реализовать работу с моделью на ЭВМ.

Часто выбор модели данных определяется успехом СУБД, однако модель данных не всегда залог успеха СУБД и не относится к первичным факторам оценки. Модель данных «приходит» пользователю как компонент СУБД, независимо от того, нравится она ему или нет.

Одна из главных проблем выбора модели данных связана со степенью ее сложности. Чем проще модель данных, тем легче пользователю понять и применить ее.

Еще один аспект выбора модели данных – сравнение естественной структуры данных со структурными средствами модели данных. Выбор модели, подсказываемой текущей организацией данных, не всегда полностью соответствует операциям, которые необходимо выполнять над данными.

Анализ совокупности факторов помогает выбрать оптимальную модель, соответствующую каждой области моделирования данных. *Оптимальная* модель данных должна удовлетворять следующим критериям [7]:

1. *Структурная достоверность* – соответствие способу определения и организации информации в данной предметной области.

2. *Простота* – легкость понимания модели как профессионалами, так и обычными пользователями.

3. *Выразительность* – способность представлять отличия между разными типами данных, связей между данными и ограничений.

4. *Отсутствие избыточности* – исключение излишней информации.

5. *Способность к совместному использованию* – отсутствие принадлежности к какому-то особому приложению или технологии.

6. *Расширяемость* – способность эволюционировать с целью включения новых требований с минимальным влиянием на уже существующих пользователей.

7. *Целостность* – согласованность со способом использования и управления информацией внутри предметной области.

8. *Представление в виде диаграмм* – способность представления модели с помощью диаграммных обозначений.

Однако иногда эти критерии несовместимы друг с другом, и приходится идти на некоторый компромисс. Например, в погоне за наибольшей выразительностью модели данных можно утратить ее простоту.

Вопрос о том, что такое «оптимальная» или «наилучшая» модель, весьма спорен. Убедительно доказать, что некоторая модель данных во всем является наилучшей, трудно. Каждая модель данных имеет свои преимущества в конкретной сфере применения.

Проведем сравнительный анализ моделей данных по выявленным критериям с целью определения наиболее оптимальной модели для проектирования БД ЛТК (табл. 2).

Сравнительный анализ моделей данных по критериям для БД ЛТК

Критерий	Классические модели				Неклассические моде-		
	ранние модели	иерархическая	сетевая	реляционная	постреляционная	многомерная	объектно-ориентированная
Структурная достоверность	-	-	+	+	+	+	+
Простота	-	-	-	+	-	-	-
Выразительность	-	-	-	+	+	+	+
Отсутствие избыточности	-	-	-	+	-	+	+
Способность к совместному использованию	-	-	-	+	+	+	-
Расширяемость	-	-	-	+	+	+	+
Целостность	-	-	-	+	-	+	-
Представление в виде диаграмм	-	+	+	+	+	+	-
Результат	0	1	2	8	5	7	4

Анализ показал, что, несмотря на все ранее сказанные недостатки реляционной модели данных, она является наиболее оптимальной для БД ЛТК. Результаты проведенного анализа свидетельствуют также в пользу применения многомерной модели. Нет смысла использовать ранние модели ввиду их неконкурентоспособности по сравнению с реляционными, и нет смысла использовать объектно-ориентированную модель, ввиду того, что под ней нет теоретической основы.

Однозначно общепотребимой модели сейчас нет и разные модели сосуществуют. Более того, они существуют взаимосвязано либо же попытки такой взаимоувязки неустанно предпринимаются. Например, существуют мнения, что реляционная и объектно-ориентированная модели взаимоисключают друг друга и что они взаимодополняют друг друга. Атрибутами в реляционных таблицах могут быть объекты произвольно заданной сложности. С точки зрения реляционной модели они остаются атомарными, а все возможности работы с ними, проистекающие из наличия внутренней структуры, реализуются объектно-ориентированными методами. Существует выбор, какие свойства предметной области моделировать реляционными методами, а какие – объектными.

Другой аспект взаимной связи указанных двух моделей носит реализационный характер. Некоторые объектно-ориентированные системы сами реализованы на «реляционно-ориентированных» СУБД, как на системах, получивших доминирующее распространение на рынке СУБД, и вследствие этого наиболее продвинутых как промышленные изделия. В таких системах определения, заданные в рамках объектного подхода, переводятся в реляционные определения, или наоборот, объектно-ориентированные определения строятся как надстройки над реляционно-ориентированными системами.

Многомерную модель данных также нецелесообразно использовать в чистом виде из-за ее сложности и требуемого большого объема памяти для реализации простейших задач. Так же как иерархическую и сетевую, многомерную модель целесообразно дополнить механизмом ссылок в виде внешних ключей, связывающих данные, расположенные в разных гиперкубах. Введение в навигационную модель данных механизма ссылок в виде внешних ключей позволяет гибко формировать ключ поиска при сложных запросах к навигационной базе данных. Такая модель получила определение интегрированной и вобрала лучшие качества иерархической (механизм поиска), сетевой (произвольный доступ) и реляционной (логическое разделение первичных структур хранения информации) моделей данных. Использование интегрированной модели позволяет отказаться от двойного представления информации, реализованного в гибридных многомерно-реляционных базах данных, предлагаемых ведущими производителями данного класса программного обеспечения. Это дополнительно увеличивает скорость доступа к данным.

Интегрированная модель данных имеет широкие перспективы развития в направлении многоверсионной архитектуры, замены процедуры, опережающей записи в журнал транзакций, на специальные алгоритмы контроля полноты и достоверности контроля без обращения к энергонезависимому носителю информации, управления отдельными экземплярами атрибутов данных и т.д.

Выводы.

1. Проведена сравнительная качественная оценка основных моделей БД и их характеристик.
2. Выявлены критерии оценки модели данных для выбора оптимальной модели.
3. Проведен сравнительный анализ моделей данных по критериям для БД ЛТК. Результаты проведенного анализа свидетельствуют в пользу выбора реляционной или многомерной модели данных.

Предлагается дополнять реляционную и многомерную модель разработками объектно-ориентированного подхода. Полученную модель данных можно определить как интегрированную модель. Ее использование для БД ЛТК позволяет отказаться от двойного представления информации, реализованного в гибридных многомерно-реляционных базах данных. Это дополнительно увеличивает скорость доступа к данным.

Список литературы

1. Когаловский М. Р. Перспективные технологии информационных систем. – М. : ДМК Пресс; Компания АйТи, 2003. – 288 с.
2. Дейт К. Дж. Введение в системы баз данных. – 8-е изд. – М. : «Вильямс», 2006.
3. Экономическая информатика / Под ред. П. В. Конюховского и Д. Н. Колесова. – СПб : Питер, 2001
4. Коннолли Т. Базы данных: проектирование, реализация и сопровождение. Теория и практика / Т. Коннолли, К. Бегг, А. Страчан; пер. с англ. – 2-е изд. – М. : Издательский дом «Вильямс», 2001.
5. Модели данных [Электронный ресурс]. – Режим доступа : <http://inftech.webservis.ru/it/diploms/dip1/c41.html/> Модели данных
6. Есин В. И. Универсальная модель данных и ее математические основы // Системи обробки інформації. – Х. : Харківський університет Повітряних Сил, 2011. – № 2 (92).
7. Карасев В. Жизненные циклы базы данных [Электронный ресурс]. – Режим доступа : <http://wiki.auditory.ru/> Жизненные циклы базы данных

References

1. Kogalovskij M. R. Perspective technologies of information systems. – М. : DMK Press; IT Company, 2003. – 288 с.
2. Date C. J. An Introduction to Database Systems. – 8th edition. – М. : “Williams”, 2006.
3. Economic Informatics. Ed. P.V. Konjuhovskij i D.N. Kolesov. – St. Petersburg : Piter, 2001.
4. Connolly T. Database Systems : Practical Approach to Design, Implementation, and Management. Theory and practice / T. Connolly, C. Begg and A. Strachan, trans. from English. – 2nd edition. – Moscow: Publishing House “Williams”, 2001.
5. Models of data [Electronic resource]. – Access mode : <http://inftech.webservis.ru/it/diploms/dip1/c41.html/> Models of data
6. Esin V. I. (2011). Universal model of data and its mathematical bases. *Information processing systems*, (2 (92)).
7. Karasev V. Life cycles of a database [Electronic resource]. – Access mode : <http://wiki.auditory.ru/> Life cycles of a database

Стаття надійшла до редакції 04.03.2013.

Відомості про автора:

Нечипоренко О.В., кандидат технічних наук, доцент, доцент кафедри спеціалізованих комп'ютерних систем, Черкаський державний технологічний університет.

Миценко С.А., аспірант кафедри спеціалізованих комп'ютерних систем, Черкаський державний технологічний університет.