

КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ ИНТЕЛЛЕКТУАЛЬНОЙ БАЗЫ ДАННЫХ ЭЛЕКТРОМАГНИТНОГО БИОКОМПЛЕКСА

к.т.н. А.Р. Корсунов, д.т.н., проф. П.Ф. Поляков

В статье рассматривается концептуальная модель интеллектуальной базы данных, что обеспечивается введением в систему модели пользователя и формированием файлов в глубину.

В качестве модели данных в теории баз данных используются различные алгебраические структуры, обобщающие классические модели данных: реляционную, иерархическую и сетевую. Дальнейшее развитие подобных моделей связано с созданием файлов порядка более чем первый в “башне” файлов. Указанный термин определяет систему сложно структурированных данных по иерархическим и реляционным моделям.

Однако в данной статье разрабатываются концепции построения информационной системы (ИС) на основе базы данных (БД) с приданием ей свойств интеллектуальности. В функционировании ИС предусматривается не просто хранение информации. Благодаря встроенным моделям, а также структурированию и нормализации массивов БД, в операционной оболочке ИС инициализируется процесс управления решением диагностических и инструментальных задач. Таким образом, в рамках интеллектуализации ИС электромагнитного биокомплекса предусматривается продуцирование знаний, что является ступенью совместной эволюции компьютерных сетей, сетей связи, биоэлектроники и ведёт к созданию инфэлбиосферы (сращиванию вычислительной техники, средств связи и биоэлектронной инженерии).

Для работы с базой данных в режиме клиент-сервер используется язык структурированных запросов. При организации БД необходимо соблюдение следующих требований:

- связность базы, т.е. контроль и реализация отношений в базе реализуется на уровне структуры базы;
- хранению подлежит только то, что не может вычисляться;
- не должно быть в одной системе двух объектов, содержащих одинаковые данные.

Указанным требованиям соответствуют ряд программ - серверов, в частности, MS SQL Server.

Проектирование БД, как единого целого, определяется рядом правил, в частности, следующими:

- разделением объектов на ряд уровней по сложности;
- введением правил идентификации и связей между объектами;

– разработкой механизма поддержки операций над объектами.

При формировании структуры БД отдельным вопросом ставится проблема определения структуры данных архива и правила переноса оперативных данных в архив. При этом принято архив составлять из двух частей: реляционная база и данные для анализа.

Первая содержит те данные, которые считаются неизменными, и объекты, существующие в оперативной базе, но необходимые для обеспечения связности данных; вторая – данные, подготовленные в форме гиперкуба (OLAP базы). Она используется для формирования встроенной модели пользователя в БД, в частности систематизации параметров диагностических наборов электромагнитного биокомплекса и представляется в виде

$$\forall Z \in [1..n]: S^Z \subset S,$$

где $\forall Z$ – набор диагностики биообъекта; S - множество параметров, состоящее из отдельных разделов S^Z ; Z - общее число разделов (показатели температуры, показатели кровеносной системы и т.д.). Раздел представляет совокупность подразделов (систолическое давление, диагностическое давление и т.п.), для каждого из которых задаётся мера значимости

$$S^Z = \left\{ [S_1^Z]^{C_1^Z}, [S_1^Z]^{C_2^Z}, \dots, [S_1^Z]^{C_v^Z} \right\}; C_v^Z \in [0,1], v \in [1..n], \quad (1)$$

где v - число подразделов в разделе S^Z ; C_v^Z - величина, характеризующая степень важности при диагностике специалистом по исследованию из подраздела S_v^Z .

Если $S_1^Z, S_2^Z, \dots, S_v^Z$ подмножества (необязательно не пересекающиеся) диагностических задач одного раздела, то

$$\forall v \in [1..n]: S_v^Z \subset S^Z; S_v^Z = \left[S_{v_1}^Z, S_{v_2}^Z, \dots, S_{v_{m_0}}^Z \right], \quad (2)$$

где $S_{v_1}^Z, S_{v_2}^Z, \dots, S_{v_{m_0}}^Z$ - диагностические задания в подразделе S_v^Z .

Учитывая выражения (1) и (2), записываем полный набор параметров, необходимых при диагностике по заданному разделу:

$$S^Z \left\{ [S_{11}^Z, S_{12}^Z, \dots, S_{1m_1}^Z]^{C_1^Z}, [S_{21}^Z, S_{22}^Z, \dots, S_{2m_2}^Z]^{C_2^Z}, \dots, [S_{v_1}^Z, S_{v_2}^Z, \dots, S_{v_{m_v}}^Z]^{C_v^Z} \right\}.$$

Полные наборы диагностических заданий для каждого из разделов формируются группой экспертов и заносятся в БД.

Для получения адекватной оценки диагностики в систему вводится модель пользователя, формально представляемая в виде объекта

$$MS = \left\{ \mathbf{ID}, (a_1, < p_{-a_1} >), (a_2, < p_{-a_2} >), \dots, \left[S_{v_1}^Z, S_{v_2}^Z, \dots, S_{v_{m_v}}^Z \right]^{C_v^Z} \right\},$$

где \mathbf{ID} - идентификатор; a_1, a_2, \dots, a_i - имена полей; $p_{-a_1}, p_{-a_2}, p_{-a_i}$ - значения полей.

В качестве имен полей используются статические характеристики пользователя. Значения полей – шкалированные оценки этих характеристик. В БД системы вводится набор функций преобразования, обеспечивающий формирование множества требований, предъявляемых при диагностике

$$S \xrightarrow{f_Z((p_{-a_i}))} \rightarrow S^Z \left\{ \left[S_{11}^Z, S_{12}^Z, \dots, S_{1m_1}^Z \right]^{C_1^Z}, \left[S_{21}^Z, S_{22}^Z, \dots, S_{2m_2}^Z \right]^{C_2^Z}, \dots, \left[S_{v_1}^Z, S_{v_2}^Z, \dots, S_{v_{m_v}}^Z \right]^{C_v^Z} \right\}.$$

Оценка уровня достигнутых параметров при электромагнитном воздействии на биообъект в пределах одного раздела выполняется в соответствии с критерием

$$\theta^Z = \frac{1}{V} \sum_{v=1}^V \frac{C_v}{m_v} \sum_{m=1}^{m_v} T_{mv} * S_{mv},$$

где V - число подразделов; C_v - показатель значимости для подраздела; m_v - число заданий в подразделе; T_{mv} - мера трудности задания; S_{mv} - булева переменная, равная 1 в случае достижения положительного эффекта при электромагнитном воздействии и 0 - в противоположном случае.

Условием успешного проведения процесса электромагнитного воздействия является выполнение неравенства

$$\theta_{\min}^Z \leq \theta^Z,$$

где θ_{\min}^Z - заданное в режиме формирования требований минимально допустимое значение критерия для данного раздела.

В качестве моделей данных принято использовать различные алгебраические структуры, обобщающие классические модели данных: реляционную, иерархическую и сетевую. Реляционная модель представления данных имеет наиболее развитый математический аппарат теории множеств со следующими недостатками:

- 1) жёсткость структуры данных (строки таблицы, например, не могут быть переменной длины);
- 2) размер БД влияет на скорость работы системы (из-за необходимости просматривать всю базу).

Иерархические и сетевые модели ускоряют доступ к необходимым данным. Однако каждый элемент данных должен содержать в своей структуре описание связей с другими элементами, что требует значи-

тельных затрат основной памяти ЭВМ. К снижению скорости обработки ведёт необходимость поддержки БД через основную память.

В связи с указанным, в существующую модель реляционных систем вводятся специальные средства ускоренного доступа к информации, действующие на произвольной глубине l файла порядка больше чем первого в “башне” файлов. Анализ указанных средств начнём с рассмотрения n множеств $D_1, D_2, \dots, D_i, \dots, D_n$. Отношение данных множеств R представляет собой множество элементов вида

$$\langle d_1, d_2, \dots, d_i, \dots, d_n \rangle,$$

где $d_i \in D_i, i=1, \dots, n$.

В теории множеств R - это подмножество декартова произведения исходных множеств D_1, \dots, D_n , называемых доменами отношения R . В свою очередь, элементы отношения d_i называются кортежами. Элементами декартова произведения $D \times B$ являются двухэлементные кортежи вида (d, b) , т.е.

$$D \times B = \{(d, b) / d \in D, b \in B\}.$$

Отношение представляется как таблица, где строки – это кортежи, а каждый столбец имеет имя, определяемое как атрибут. Рассмотренные табличные отображения в “башне” файлов ведут к l - глубинным табличным отображениям через фиксацию файла $R_n \in \text{Rel}(U)$ и строку $L \in C\Omega(R_n)$, где $C\Omega$ - оператор общего насыщения. Далее для указанной пары создаётся отображение $\hat{g}(R_n, l)e : \text{Rel}(U) \rightarrow \text{Rel}(U)$, исходя из следующего определения:

$$\hat{g}(R_n, l)e(R_m) = \{L \times e\beta \mid R_n \times e\beta \subseteq R_m\}, \forall R_m \in \text{Rel} \quad (3)$$

Данное отображение принято называть l - глубинным табличным отображением пары (R_n, l) .

Наибольший интерес среди всех l - глубинных отображений, $l \in N$, представляет класс BV -отображений ($l = 1$). Подобное отображение найдёт применение в адаптивных алгоритмах динамической репликации данных в распределённых информационно - управляющих системах.

ЛИТЕРАТУРА

1. Корсунов А.Р. Информационное и структурное воздействие на биоструктуры СВЧ-сигнала, регулируемого цифровым аттенуатором // Вестник НТУ “ХПИ”. Автоматика и приборостроение. – X. : НТУ “ХПИ”. – 2001. – № 4. – С. 117 - 119.
2. Корсунов А.Р., Хуторненко С.В. Анализ математических и электронных моделей биоструктур // Системы обробки інформації. – X. : НАНУ, ПАНМ, ХВУ. – 2001. – Вип.6(16). – С. 70 - 73.

Поступила 25.03.2002

КОРСУНОВ Анатолий Рувимович, канд. техн. наук, доцент кафедры автоматике и радиоэлектронике Украинской инженерно - педагогической академии. В 1968 году окончил

ХГУ. Область научных интересов – взаимодействие электромагнитных сигналов с биоструктурами.

ПОЛЯКОВ Петр Федорович – доктор техн. наук, профессор, зав. кафедрой “Транспортная связь” Харьковской государственной железнодорожной академии. В 1970 году окончил ХИРЭ. Область научных интересов – системы телекоммуникации и связь.
