

В. А. Беляков

Динамический web-сервер многомерных баз данных многоцелевого назначения

Предлагается новая информационная технология для создания корпоративных динамических web-систем многоцелевого назначения на примере информационно-аналитической системы оценки устойчивого развития г. Тюмени. Рассматриваются структура единой интегрированной многомерной базы данных, основные характеристики универсального программного комплекса ведения баз данных и специализированные информационные решения по оценке устойчивого развития г. Тюмени.

Введение

Конвенция об устойчивом развитии, принятая Генеральной Ассамблеей ООН 5 июня 1992 г. в Рио-де-Жанейро, провозгласила необходимость перехода мирового сообщества на путь устойчивого развития, обеспечивающего должный баланс между решением социально-экономических проблем и сохранением окружающей среды, удовлетворения основных жизненных потребностей нынешнего поколения с сохранением таких возможностей для будущих поколений. Договаривающиеся стороны констатировали необходимость системного сбора информации на всех уровнях с целью создания информационной базы обеспечения охраны окружающей среды и устойчивого развития, а также выработки достаточно универсальных показателей устойчивости развития.

Непременным условием устойчивого развития общества является сохранение окружающей среды, ее защищенность от воздействия вредных техногенных, природных, экологических и социальных факторов. При установлении целесообразного и приемлемого для общества уровня предельно допустимых нагрузок на экосистемы возникает необходимость в проведении многокритериального анализа условий и путей устойчивого развития общества.

Когда речь идет об устойчивом развитии, совершенно очевидно, что в его состав необходимо включить экологические, экономические, технологические, социальные, политические, психологические аспекты. Соответствующие системы посредством разнообразных и подчас критических способов соединяются в одну сложную общую систему. Для того чтобы общая система была жизнеспособной, жизнеспособной должна оставаться каждая из составляющих ее основных подсистем. Жизнеспособность общей системы зависит от надлежащего функционирования подсистем.

Практическая реализация задач устойчивого развития в первую очередь связана с созданием информационных систем на основе корпоративных многомерных баз данных, распределенных в глобальной сети Internet. Одна из главных целей построения междисциплинарной информационной системы состоит в том, чтобы выявить актуальные показатели для каждой подсистемы. Больше того, необходимо определить показатели, обеспечивающие информацию о вкладе каждой подсистемы в поддержание жизнеспособности общей системы, и на этой основе разработать интеграционные математические модели оценки устойчивого развития.

1. Основные цели и задачи

Разработка динамического web-сервера многомерных баз данных многоцелевого назначения позволит создать хранилище агрегированных показателей устойчивого развития. Эти показатели могут быть использованы как в управлении городом, так и в научных исследованиях.

Web-сервер разрабатывался в архитектуре вычислительного комплекса Client — Server в сети Internet/Intranet (рис. 1).

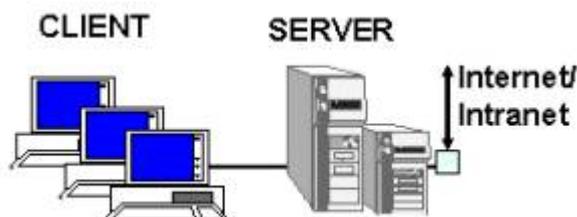


Рис. 1. Архитектура вычислительного комплекса web-сервера ИПОС СО РАН

Многомерная база данных — мощная база данных со специальной организацией хранения — кубами, позволяющая пользователям анализировать большие объемы данных. Система управления многомерными

базами данных (СУМБД) обеспечивает высокую скорость работы с данными, хранящимися как совокупность фактов, измерений и заранее вычисленных агрегатов.

Основные решения многомерных баз данных включали:

— создание универсальной технологии обработки и анализа результатов мониторинга — временных рядов и документальных материалов, что позволит выявить закономерности в поведении наблюдаемого процесса и прогнозировать его поведение в будущем;

— разработку технологий формирования электронной библиотеки по наиболее значимым публикациям, книгам, законам и т. п. в совокупности с аналитическими средствами информационной системы по рассматриваемым тематическим направлениям;

— разработку механизмов по защите информации, которые должны включать регистрацию и определение прав доступа пользователя к тематическим разделам и программам их обработки.

2. Предлагаемые методы и подходы

Функционирование web-сервера обеспечивается универсальным программным комплексом (рис. 2), включающим подсистемы: ADMINISTRATOR — ведение баз данных, CLIENT — интерфейс для специализированного программного комплекса SITE и работающие на единой интегрированной многомерной базе данных DATA BASE.

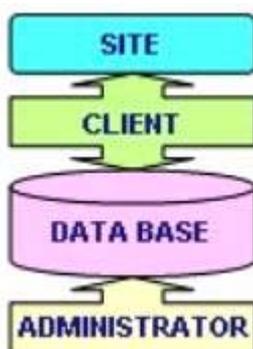


Рис. 2. Концептуальная схема управления многомерными базами данных

Предлагается за основу создания динамического web-сервера принять структуру многомерных баз данных, где массивам каждой базы данных присваиваются условные наименования и функции (рис. 3):

(P) — «Тема/Раздел» (тематические направления);

(X) — «Объекты» (географические объекты наблюдений/главы книг/персонал организации);

(Y) — «Параметры» (показатели/свойства «объектов» наблюдений);

(Z) — «Задачи» (методы обработки информационных материалов);

(M) — «Материалы» (информационные материалы: временные ряды («**таблицы**») или текстово-графические HTML-документы («**документы**»));

(S) — «Права доступа» пользователей к тематическим разделам и программам их обработки;

(B, b) — ссылки на источники информации (библиографии).

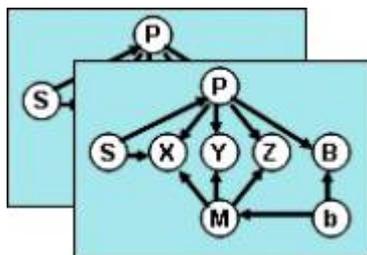


Рис. 3. Концептуальная схема многомерных баз данных

Исходя из определений многомерной базы данных, в качестве осей измерений используются таблицы с иерархическими метками (дескрипторами) X, Y, Z (1–3 уровня) по каждому из тематических разделов P (1–2 уровня). Создание дополнительной оси Z обеспечит объединение таблиц временных рядов и документов в

одной структуре баз данных (М), включая и нормативно-справочные данные, а также настройку ранга таблицы временных рядов (определение количества столбцов). Соглашение по единой структуре ключевых полей таблиц (Р, Х, Y, Z) и таблицы библиографий (В) определит разработку единой функции формирования дерева многоуровневых иерархических строк описаний (меток/дескрипторов).

Наращивание тематических направлений может привести к необходимости реализации распределенного характера многомерных баз данных на основе единых структур БД и режимов их обработки. Создание технологии формирования распределенных многомерных баз данных является важным фактором, определяющим тиражируемость типовых структур БД многоцелевого назначения.

Программный комплекс **ADMINISTRATOR** (см. рис. 1) предназначен для ведения корпоративной многомерной базы данных, главное двухуровневое меню которого включает следующие режимы:

1) **«Классификатор»**, обеспечивает формирование массивов — **«Р-Тема»**, **«Х-Объект»**, **«Y-Параметр»**, **«Z-Задача»**, **«В, в-Библиография»**.

2) **«Материалы»**, обеспечивает формирование массива информационных материалов (М) (подрежимы **Учет/Поиск**) с представлением (обработкой) их в виде документов или таблиц (временных рядов).

3) **«Права доступа»**, обеспечивает: а) **регистрацию** пользователей (Х); б) определение **прав доступа** (S) пользователей (Х) к информационным ресурсам/тематическим разделам (Р).

4) **«Статистика»**, обеспечивает формирование отчетов статистических данных по заполнению базы данных: а) **по темам/тематическим разделам**; б) **по операторам**.

Режим **«Классификатор»**. Структура окна классификатора состоит из трех рабочих зон, где 1-я зона предназначена для формирования строк дескрипторов массива (Р), 2-я зона — для формирования строк дескрипторов массивов Х/Y/Z, 3-я зона — для управляющих кнопок (модификация, ввод, удаление строк дескриптора, выход). Подрежим «Тема» обеспечивает формирование и кодировку двухуровневого дескриптора массива (Р). Кодирование осуществляется двухразрядным литерным кодом, например — JA. Текст наименования дескрипторов (название) содержит не более 100 знаков. Массив публикуется как двухуровневое дерево дескрипторов. Формирование массивов (Х, Y, Z, В) по тематическому разделу (Р) осуществляется соответственно подрежимами «Объект», «Параметр», «Задача», «Библиография» по единому алгоритму.

Кодировка дескрипторов цифровая шестизначная, два знака на уровень. Текст наименования дескрипторов содержит до 250 знаков. «Объект» и «Параметр» определяют классификацию основных объектов наблюдений и их свойств, в трехуровневой иерархии дескриптора подразделяются как группа/подгруппа/объект или параметр.

При формировании массивов-классификаторов приняты соглашения.

1) По массиву «Z-Задача» принято соглашение на структуру кода дескриптора — NNZZPP, где NN — код наименования задачи (NN = 00–99); ZZ определяет вид информационного материала (ZZ = 01 — документ, ZZ = 02–09 — таблицы временных рядов); PP = 01–99 — колонки таблиц.

2) Дескриптор «В-Библиография» формируется следующим образом: код первого уровня определяет вид библиографии (например, литературные источники, монографии, сборники статей, периодические издания и т. п.), код второго уровня — алфавитную группу по первым двум литерам наименования источника информации, код третьего уровня — наименование «библиографии/источника информации».

Режим **«Материалы»**.

Структура рабочего окна состоит из пяти зон:

1 — зона для формирования карточки запроса (дескрипторы массивов Раздел/Объект/Задача /Параметр);

2 — зона для отображения текущего дескриптора;

3 — зона отображения информационных материалов;

4 — зона для отображения ссылок на источник информации включаемого в базу данных информационного материала;

5 — зона для управляющих кнопок, включающая привязку HTML-документа, удаление ссылки на HTML-документ, привязка библиографии, удаление ссылки на библиографию, выход на главное меню.

Режим **«Права доступа»** — функция web-администратора, включающая подрежимы «Регистрация» и «Доступ». Подрежим «Регистрация» обеспечивает формирование базы пользователей и регистрацию пользователя в качестве оператора базы данных. Подрежим «Доступ» обеспечивает определение прав доступа зарегистрированному пользователю к тематическим разделам и режимам главного меню.

Режим **«Статистика»** — функция web-администратора, обеспечивающая выдачу справок заполнения базы данных. Запрос включает выбор типа информационных материалов (документ/показатель), временной

период и вид отчета («по темам» или «по операторам»).

Программный комплекс CLIENT — это набор подпрограмм/фреймов, обеспечивающих web-разработку интерфейса доступа к универсальной базе данных при создании специализированного web-сайта. Подпрограммы обеспечивают предоставление из базы данных «материалов» и источников информации по интерактивному запросу, формирующемуся исходя из выбора соответствующей строки («клик») древовидного дескриптора в ключевых таблицах «объект» — obj / «параметр» — prop / «задача» — app. Примеры вызова подпрограмм (URL адреса) приведены ниже.

M1. <http://<URL>/RICS/m1.pl?db=1&app=ZQ020100>

M2. <http://<URL>/RICS/m2.pl?db=2&app=YC020100&prop=YC100000>

M3. <http://<URL>/RICS/m3.pl?db=3&app=SH102000&obj=SH102501>

M4. <http://<URL>/RICS/m4.pl?db=4&app=ZQ020100&obj=ZQ020402&prop=ZQ050000>

M5. <http://<URL>/RICS/m5.pl?db=5&app=XA010100&prop=XA100000>

M6. <http://<URL>/RICS/m6.pl?db=6&app=SH102000&obj=SH102501>

Специализированный программный комплекс SITE структурного оформления главной странички сайта рассматривается здесь как средство для предоставления информации из универсальной базы с использованием интерфейсного механизма Client. В общем случае этот механизм гиперссылок встраивается в любую web-форму.

3. Использование механизма многомерных баз данных в создании информационно-аналитической системы устойчивого развития г. Тюмени

Разработка и использование четкой концептуальной структуры для ориентации процесса оценки чрезвычайно важны. В случае уместной концептуальной структуры показатели проявляются более естественным образом и могут быть адаптированы к нуждам данной местности или группы лиц, ответственных за принятие политических решений. В данном проекте предлагается структура оценки устойчивого развития, базирующаяся на следующих ключевых массивах[1].

1. Классификатор тематических разделов (ТЕМЫ/ОРИЕНТИРЫ):

- [J] Устойчивое развитие
 - [A] Индивидуальное развитие
 - [B] Социальная система
 - [C] Система политического управления
 - [D] Экономическая система
 - [E] Система инфраструктуры
 - [F] Система окружающей среды и ресурсов

2. Классификатор «объектов» (г. Тюмень):

- [10] Показатели устойчивого развития: индивидуальное развитие
 - [25] Тюменский район
 - § [01] г. Тюмень

3. Пример формирования классификатора показателей устойчивого развития («параметров»):

Индивидуальное развитие:

- [01] Существование
 - [02] Функционирование системы
 - § [02] Индекс развития человеческого потенциала
 - § [04] Доля лиц, имеющих отдельную комнату в благоустроенном жилье, отнесенных к общей численности

Методы оценки позволяют графически отобразить зависимость между состоянием показателей и удовлетворением соответствующих ориентиров на основе обработки временных рядов. Оценка ведется исходя из динамики данных мониторингов. В соответствии с разрабатываемой схемой, требуется осуществление двух видов оценки: первая оценка имеет дело с жизнеспособностью конкретной составляющей системы, вторая — с тем вкладом, который вносит данная составляющая в поддержание жизнеспособности общей системы. Первые шаги в познании математической оценки устойчивого развития

осуществляются следующими математическими моделями.

Алгоритм оценки первого вида. **Оценка динамики изменений интегрального показателя или группы показателей временного ряда**

Оценка состояния процесса реализуется функцией интегрального показателя $Pijt = Kijt \cdot Xijt / Yijt$, вычисляемого по усредненному показателю $Xijt$ (временной ряд) в сравнении с соответствующим усредненным нормативом/ограничением/базовым показателем $Yijt$ (при отсутствии массива $Yijt = 1$), где i — территория; j — показатель; t — период. $Kijt$ — поправочный коэффициент (при отсутствии массива $Kijt = 1$). По результатам вычислений строится график.

Алгоритм оценки второго вида. **Метод оценки ориентиров, тематических направлений для устойчивого развития**

Данная проблема решается здесь за счет выбора показателей из базы данных, содержащей временные ряды нескольких сотен показателей. Из этой базы данных можно получить перечень показателей, адекватным образом охватывающих различные аспекты оценки удовлетворения базовых ориентиров. Указанный метод использует особенности иерархического формирования показателей (3 уровня) по двухуровневой иерархии ориентиров/тематических направлений. По каждому разделу/подсистеме рассматриваемого ориентира рассчитывается вектор (вершина), равный усредненному интегральному показателю $Pijt = Kijt \cdot Xijt / Yijt$ (см. описание выше). График оценки представляет замкнутый контур («звезда») вершин ориентиров равноделящих окружность усредненной нормы ориентиров (усредненная устойчивость), как правило 6 векторов. По смещению «звезды» от центра и вершин от нормы можно оценить устойчивость ориентира на основе практических наблюдений и анализа фактического материала.

Заключение

По предлагаемой технологии также создан web-сервер ИПОС СО РАН (<http://www.ipdn.tmn.ru>). Сервер представляет базы данных комплексного мониторинга Тюменской области по тематическим направлениям СОЦИАЛЬНАЯ ЭКОНОМИКА, АРХЕОЛОГИЯ, ЭТНОГРАФИЯ, БИОРАЗНООБРАЗИЕ, ЗАГРЯЗНЕНИЕ [1–6] и включает:

- динамику показателей социально-экономического развития юга Тюменской области и г. Тюмени за 1996–2000 гг.;
- сведения о биоразнообразии п-ова Ямал, включающие описания видов птиц, млекопитающих, грибов, мхов, лишайников, цветковых;
- сведения об археологических памятниках по районам Ямало-Ненецкого автономного округа и юга Тюменской области;
- сведения о коренных народах п-ова Ямал;
- данные мониторинга загрязнения химическими элементами Среднего Приобья и г. Тюмени.

Наблюдение за эксплуатацией системы более года показало, что использование этой технологии обеспечивает:

- 1) возможность формировать корпоративную библиотеку больших объемов специалистами, имеющими минимальные знания в области информационных систем;
- 2) упрощение функций web-администрирования;
- 3) возможность интеграции с другими web-системами.

Наибольшую эффективность предлагаемая технология многомерных баз данных даст в интеграции в распределенной среде Internet с web-системами, включающими метаданные на ресурсы многомерных баз данных и использующими свои аналитические и поисковые технологии.

Литература

1. Цибульский В. Р., Беляков В. А. Информационно-аналитическая система природных и социальных условий полуострова Ямал // Информ. бюл. ГИС ассоциации № 2(9). 1997. С. 64.
2. Беляков В. А. Региональная геоинформационная система устойчивого развития на примере Тюменской области // Материалы междунар. конф. «ENVIROMIS 2000». Томск, Институт оптики атмосферы, 23–28 октября 2000. Томск, 2000. С. 61.
3. Беляков В. А. Региональная web-геоинформационная система устойчивого развития на примере Тюменской области // Материалы междунар. конф. «Новые Internet технологии». Петрозаводск, 25–29 июня 2000.
4. Цибульский В. Р., Беляков В. А. Динамический геоинформационный WWW сайт природных и социальных условий Тюменской области // Электронные библиотеки. 1999. Т. 2. Вып. 4.
5. Беляков В. А. Динамический геоинформационный web-сервер многоцелевого назначения // IV Всерос. учеб.-практ. конф. Геоинформатика в нефтегазовой и горной отраслях. Гипротюменьнефтегаз, Тюмень, 2–4 апреля 2001 г.
6. Беляков В. А. Динамический геоинформационный web-сервер многоцелевого назначения // VI Рабочее совещание по электронным публикациям «Ei-Pub-2001», 25–27 апреля 2001 г., Новосибирск, Академгородок.

[1] Полный перечень показателей смотри в отчете по теме: «Разработка системы показателей и методики оценки социально-экономического развития г. Тюмени как начального этапа разработки стратегического плана развития г. Тюмени». ИПОС СО РАН, 2001. Прил. 2.