

МЕТОДЫ ИНТЕЛЛЕКТУАЛИЗАЦИИ ПРИКЛАДНЫХ ЗАДАЧ В ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ СИСТЕМЕ «ОБРАБОТКА СТАТИСТИЧЕСКОЙ ОТЧЕТНОСТИ»

Г.И. Храмов, канд. экон. наук,
Территориальный орган Росстата по Красноярскому краю,
С.М. Окладников,
Сибирский государственный технологический университет

В журнале «Вопросы статистики» № 3 за 2005 г. была опубликована наша обзорная статья об унифицированной информационно-аналитической системе (ИАС) «Обработка статистической отчетности» (ОСО). В связи с развитием корпоративной информационно-вычислительной системы Росстата, предстоящими сельскохозяйственной (2007 г.), экономической (2008 г.) переписями и переписью населения (2010 г.), нами подготовлен материал, который может быть использован при подготовке программного продукта автоматизированной обработки не только информации текущей статистики, но и всех единовременных выборочных и сплошных обследований, в особенности для районного, а далее и муниципального уровней. В этой статье содержится подробное описание автоматизированной справочной службы, реализованной в ИАС в виде функции, которая возвращает список оптимальных действий, совершаемых пользователем для достижения определенного результата в рамках ИАС «ОСО». Однако локальность данной функции не определяет ее специфичность, а минимизация или полное отсутствие специфики порождается ее универсальностью. Задача универсализации средств по обработке статистической отчетности была основополагающей при проектировании и реализации ИАС «ОСО». Ее решение определило унифицированный функциональный состав системы, в котором каждая отдельная функция является универсальной, то есть допускается ее реализация вне ИАС как самостоятельного приложения. Соответственно и функция оптимизации принимаемых решений является универсальной, а это значит, что ее применение не зависит от специфики предметной области. Гибкость формализации логики предметной области позволяет легко перенастраивать функциональную модель системы, используемую в качестве формальных параметров в рассматриваемой функции.

Значимость поиска оптимальных действий при работе с программными приложениями различных предметных областей, особенно с многофункциональными информационно-аналитическими системами, становится все более весомой. Насыщение ИАС многообразием различных функциональных возможностей влечет за собой создание сложных по иерархии справочных служб как неотъемлемой части любой системы. Справочные службы в гипер-

текстовом формате существенно упрощают решение пользовательских задач при работе с системой, но не всегда можно оперативно получить справку о решении частной задачи с набором определенных условий. В режиме регламентной работы¹ иногда просто не хватает рабочего времени на изучение определенного ряда функций системы. Эта проблема актуальна для работы в структуре Росстата. Частое обновление форм отчетности влечет за собой обновление программных продуктов. Проблема единообразной обработки статистической отчетности была решена с созданием унифицированной системы ИАС «ОСО». Однако динамика системы, ее функциональное «наращивание» создают определенные трудности у пользователей, что проявилось на этапе апробации.

Для оперативной ориентировки в большом наборе функций нами в систему была внедрена функция, условно называемая «интеллектуальный путеводитель». Данная функция позволяет оперативно получить ответ на вопросы, связанные с выбором тех или иных действий в рамках системы при наличии определенных входных параметров.

В основе алгоритма «интеллектуального путеводителя» использован аппарат Сетей Петри (СП). Формально Сеть Петри описывается множеством $PN = \{\Theta, P, T, F, M_0\}$,

где $\Theta = \{0, 1, 2, \dots\}$ - множество дискретных моментов времени;
 $P = \{p_1, p_2, \dots, p_n\}$ - непустое множество элементов Сети (позиций);
 $T = \{t_1, t_2, \dots, t_n\}$ - непустое множество элементов Сети (переходов);
 F - функция инцидентности;
 $F: (P \times T) \cup (T \times P) \rightarrow \{0, 1, 2, \dots, k, \dots\}$;
начальная разметка позиций $M_0: P \rightarrow \{0, 1, 2, \dots\}$.

Сеть Петри, как аппарат моделирования, изначально играла приоритетную роль, выражающуюся в имитационном моделировании параллельных систем. В настоящее время Сети Петри применяются не только для распараллеливания процессов, но и в лингвистике, верификации функционального моделирования и т. п. В рассматриваемой работе СП применяется для моделирования функционального состава ИАС. Построенная модель является аргументом в функции «интеллектуального путеводителя».

Задача функции - определить оптимальные действия пользователя для достижения заданных выходных параметров. На рис. 1 изображена модель ИАС «ОСО» в нота-

¹ Регламентной работой будем называть режим, определенный особыми временными ограничениями на выполнение той или иной функции системы.

Таблица 1

Таблица 1

*Время, необходимое для выполнения соответствующего события, задано в условных единицах.

*Время, необходимое для выполнения соответствующего события, задано в условных единицах.

Таблица 2

Рассмотрим работу «интеллектуального путеводителя», реализованного в ИАС «ОСО».

Решение задачи. Задаем начальную разметку Сети Петри [101000010000100], в которой определены входные и выходные параметры модели. Процесс моделирования производится до определенной глубины. Затем анализи-



руются векторы свободного языка СП. Анализ включает в себя нахождение минимальной суммы вектора. Вектор представляет собой одномерный массив, хранящий временные характеристики последовательности срабатывания переходов. В данной задаче будут сопоставляться векторы, в которых заполнена позиция p_{11} , что сигнализирует о выводе протокола отправки телеграммы. Сеть Петри, моделирующая функциональность ИАС «ОСО», имеет особенности построения. Двухнаправленная стрелка обозначает для одного перехода одновременно две дуги: исходящую и входящую.

На рис. 2 изображена форма ИАС «ОСО» для работы с «интеллектуальным путеводителем». В списке, расположенном слева, отмечены позиции, имеющиеся в наличии. По умолчанию, список формируется автоматически, исходя из фактического наличия данных позиций в системе. В центральном списке формы отмечаются те позиции, которые нужно получить в ходе работы с ИАС. И в окне справа отображаются результаты работы «интеллектуального путеводителя» в виде списка действий, которые пользователю необходимо выполнить для достижения цели.



Рис. 2. Форма «интеллектуального путеводителя»

В результате работы модели до глубины, равной 5, был сформирован свободный язык Сети Петри, состоящий из 237 векторов. Среди них был найден единственный вектор, после срабатывания переходов которого в позиции p_{11} появилась фишка: $\text{Vector}=[t_1, t_2, t_3, t_4, t_5]$. Заполним данный вектор совокупностью соответствующих значений функций времени из таблицы 1: $\text{Vector}=[5, 6, 8, 11, 1]$. Полученная длина вектора $L=\sum_{i=1}^n(\text{vector}[i])=31$.

Таким образом, имея в наличии данные статистического регистра, инструктивные письма для заполнения статистической отчетности, штатное расписание и первичную отчетность, для отправки телеграммы пользователь должен совершить следующие действия:

- 1) определить регламент;
- 2) создать электронную форму;
- 3) создать макет телеграммы;
- 4) ввести данные;
- 5) отправить телеграмму.

Зная оптимальный набор действий, можно исключить вероятность выполнения лишних функций в системе, что приводит к экономии рабочего времени и уменьшению трудоемкости.

В данной статье использование методов интеллектуализации принятия оптимальных решений рассмотрено нами на примере ИАС. Однако практическое применение разработанных алгоритмов может быть более широким.

Реализация Федеральной программы статистических работ Росстата в целом, и особенно проведение важных государственных мероприятий по сплошному статистическому наблюдению, таких, как прошедшая перепись на-

селения, предстоящие Всероссийская сельскохозяйственная и экономическая переписи, предполагают объемную подготовительную работу.

Для реализации поставленных задач привлекаются внештатные сотрудники, для большинства которых статистика является новой областью в их трудовой деятельности. Анализ централизованного обучения переписных кадров на территории Красноярского края показал, что около 50% районных органов статистики с трудом воспринимали информацию, касающуюся автоматизированной обработки списков субъектов сельскохозяйственной переписи. В ходе обучения возникало множество вопросов, связанных с проблемой составления списков в АРМ «Агрорегистр-1», которые решались уже на межрайонном уровне. Приходилось повторно организовывать обучение и консультации в межрайонных отделах специалистов из близлежащих районов.

Большую помощь специалистам различной квалификации могли бы оказать интерактивные справочные службы. Обычно службы автоматизированной помощи, интегрированные в программные комплексы, организованы традиционно иерархическим способом в гипертекстовом виде. Методологические указания, прилагаемые, в частности, к АРМ «Агрорегистр-1», представлены в текстовом виде в формате MS Word. В большинстве случаев в различных структурах организация представления информации, имеющая общую семантическую интерпретацию, может легко дезориентировать пользователя в дальнейших действиях. Более приемлема для стандартного мышления человека интерактивная форма информационного

обмена. Именно такую форму обмена информацией и предлагают методы интеллектуализации принятия оптимальных решений, реализованные в информационно-аналитической системе «Обработка статистической отчетности», которые могут применяться в различных сферах в силу их универсальности. В этом случае можно оперативно получать конкретную информацию на заданный системе вопрос.

Большинство интерактивных справочных служб применяются в специфичной области. Преимущество рассмотренных в данной статье методов в том, что они по-

зволяют довольно гибко адаптироваться к предметной области с помощью изменения функционального состава системы. При этом функциональная модель должна задаваться централизованно. Таким образом, рассмотренная методика, как уже упоминалось, может применяться вне комплекса ИАС, в виде отдельного приложения для обеспечения оперативного получения необходимой справочной информации в различных сферах деятельности, в частности для получения итоговой информации предстоящих переписей на любом уровне [Росстат, территориальные и районные (городские) статистические органы].

ПРИМЕНЕНИЕ ТЕХНОЛОГИИ ХРАНИЛИЩ ДАННЫХ НА ПРИМЕРЕ СОЗДАНИЯ «ЦЕНТРАЛИЗОВАННОГО ХРАНИЛИЩА ДАННЫХ ПОКАЗАТЕЛЕЙ СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО РАЗВИТИЯ»

Е.В. Миронова,

Территориальный орган Росстата по Республике Саха (Якутия)

Целью внедрения технологии хранилищ данных в Саха(Якутия)стате является создание единого информационного пространства, объединяющего все статистические и ведомственные потоки информации в один мощный корпоративный управляемый ресурс - Централизованное хранилище данных (ЦХД) показателей социально-экономического развития Республики Саха (Якутия).

Основой для реализации этого проекта послужили новые требования, предъявляемые к организации информационных ресурсов, связанные с принятием региональных программ информатизации органов власти в рамках Федеральной программы «Электронная Россия» на 2002-2010 гг.

На сегодняшний день региональная информация разбросана по ведомственным и муниципальным системам, отсутствует централизованный сбор, накопление и контроль данных для проведения комплексного анализа и мониторинга состояния региона. Вследствие этого управленческие структуры не успевают перерабатывать и обобщать большие потоки информации.

Эффективную поддержку принятия управленческих решений призвана обеспечить информационно-аналитическая система, созданная в территориальных органах Росстата, так как именно в статистической системе существует сложившийся, отлаженный процесс сбора, хранения, обработки и контроля территориальной информации с использованием стандартизованных общероссийских классификаторов.

Возможности системы ощутимо облегчают постоянно возникаемую проблему отсутствия гибкого и универсального программного модуля в части представления выходной информации, так как сегодня пользователь хочет иметь возможность самостоятельно строить сложные многоуровневые запросы по своему макету.

Новая информационная система ориентирована на

современную технологию интегрированных хранилищ данных и углубленную аналитическую обработку больших объемов накопленной информации современными методами поддержки принятия решений.

Успешному внедрению проекта способствовал профессиональный уровень разработчиков системы и многолетний опыт специалистов территориальных органов Росстата по применению информационно-аналитических систем и различных СУБД.

Работы по проекту начаты в конце 2003 г. и реализованы в инструментальной среде Аналитического комплекса «Прогноз» (ЗАО «Прогноз», г. Пермь), основанной на технологии WareHouse с использованием СУБД ORACLE Enterprise Edition Server 9i.

Разработка нормативной базы. На начальном этапе работ была проведена параметризация и согласование с заказчиком этапов проекта, определены финансовые, материальные и временные ресурсы.

Исходя из информационной потребности круга пользователей, был определен состав показателей и проведена идентификация со статистическим справочником, определен перечень источников данных для загрузки. Таким образом, на этом этапе была проделана следующая работа:

- создана служба эксплуатации ЦХД, деятельность которой закреплена регламентом службы эксплуатации ЦХД, утвержденным руководителем Саха(Якутия)стата и согласованным с заказчиком - Министерством экономического развития Республики Саха (Якутия);
- составлен «Регламент работ по созданию, взаимодействию и использованию ЦХД», который был утвержден постановлением Правительства Республики Саха (Якутия);
- согласован с заказчиком перечень показателей, сопоставлен и идентифицирован со стандартизированным статистическим классификатором показателей - КСП (один из самых трудоемких процессов создания ЦХД);

- определены источники данных для загрузки:

а) базы данных показателей: оперативные базы данных федеральной статистической отчетности, региональные базы данных Республики Саха (Якутия), базы данных Дальневосточного федерального округа;

б) базы готовых документов (БГД) Республики Саха (Якутия) и России.

Реализация проекта. Система ЦХД реализована в среде Аналитического комплекса «Прогноз», который включает в себя универсальные мастера и модули, обеспечивающие свободную модернизацию компонентов системы, и состоит из следующих функциональных блоков:

- мастера конструирования баз данных;
- модули экспорта-импорта данных различных источников;
- OLAP-средства работы с данными;
- мастера конструирования диалоговых форм;
- Web-компонента отображения данных в Интернете;
- конструктор баз данных и отчетов;
- конструктор пользовательских АРМов (Дизайнер форм);
- конструктор математических моделей;
- средство администрирования;
- OLAP-навигатор.

На этапе внедрения системы в эксплуатацию были определены требования к конфигурации и количеству клиентских мест, произведена установка и настройка оборудования. Тщательно проанализированы информационные потребности внутренних и внешних пользователей, на основании этого обследования были сформированы требования к разработчику по функциональным возможностям:

- по способам загрузки и хранения данных;
- по представлению данных;
- по видам протоколов (контроля, загрузки);
- по способам разграничения прав.

Далее службой эксплуатации ЦХД были произведены следующие работы:

- определение структуры и содержания нормативно-справочной информации (НСИ), каталогов;
- определение состава, содержания и глубины динамики загружаемых данных;
- определение схем загрузки абсолютных и производных показателей;

- определение логической структуры витрин данных;
- выполнение работ по наполнению репозитория;
- загрузка в хранилище данных годовой и оперативной периодичности;

- построение витрин данных для внутренних и внешних пользователей.

На сегодняшний день в ЦХД загружены агрегированные итоги в различной разрезности начиная с 2000 г. (годовые), оперативные - с начала 2004 г., по перечню показателей (около 2,5 тыс. показателей), предусмотренных согласно объема регламентированного рамками Госзаказа, утвержденного Правительством республики. Все работы по диспетчеризации и организации ведения ЦХД обеспечивает служба эксплуатации ЦХД, осуществляющая:

- администрирование системы;
- разграничение прав пользователей;
- загрузку и актуализацию данных;
- загрузку и актуализацию НСИ;
- построение витрин представления данных.

Доступ к ЦХД осуществляется следующим образом:

- для специалистов Саха (Якутия) - через ЛВС [сервер ЦХД (основной) подключен к локальной сети] доступ к показателям предоставляется в полном объеме;

- для специалистов министерств - по опто-волоконному каналу к Web-серверу ЦХД. В министерствах установлена клиентская часть ЦХД. Доступ к показателям регламентирован;

- для удаленных пользователей реализована Web-версия ЦХД, доступ в таком случае предоставляется через глобальную сеть Интернет и включает несколько ограниченный набор функциональностей ЦХД.

Создание ЦХД в системе Саха (Якутия) - это позволило перейти на более качественный уровень организации работ по хранению, актуализации и представлению статистической информации. Появился мощный инструмент, позволяющий:

- объединить показатели большого множества статистических баз в один глобальный ресурс со стандартизованным подходом ко всем показателям;

- оперативно отображать данные в режиме Экспресс-просмотра по любому показателю, загруженному в базу данных ЦХД. Окно экспресс-просмотра включает в себя иерархический список социально-экономических показателей в виде дерева (см. рис. 1);



Рис. 1

- самим пользователям, не дожидаясь выхода статбюллетеня, оперативно формировать собственные выходные таблицы (графики, диаграммы) по желаемым макетам в любой динамике, так как режим «Витрин данных» позволяет самостоятельно строить и сохранять сложные многоуровневые аналитические запросы в различных разрезах с дальнейшей выгрузкой в другие программы;

- экспортировать информацию хранилища данных в установленные форматы и базы данных (ежемесячно осуществляется выгрузка данных в Ситуационный центр Минэкономразвития республики посредством использования служебных таблиц сопоставления показателей);

- в режиме представления данных выбрать различные методы анализа для конкретных категорий пользователей. Разработчики постоянно расширяют возможности статистического и интеллектуального анализа (структурный, сравнительный, динамический, кластерный, регрессионный).

Следует отметить, что особенной популярностью пользуется встроенное в разделы Экспресс-просмотра и Витрин данных картографическое и графическое отображение данных. ГИС-карта наглядно формирует основу для оперативного мониторинга и анализа социально-экономического состояния региона и районов (см. рис. 2).

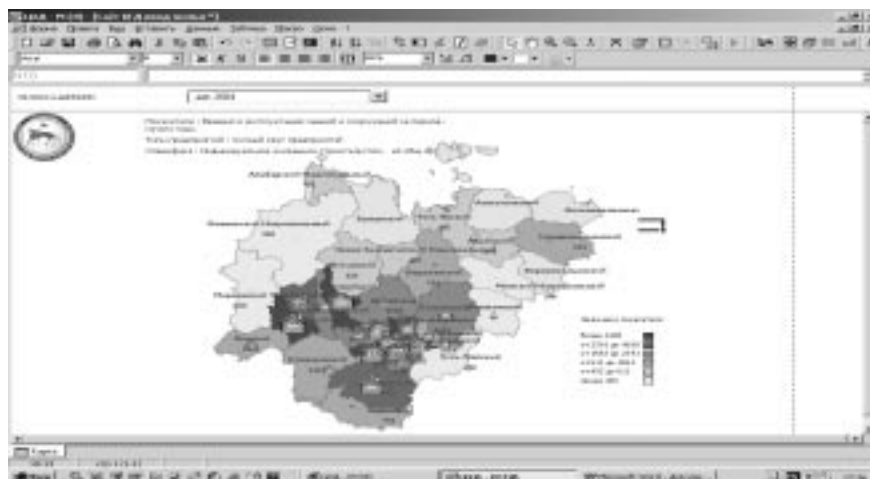


Рис. 2

Создание ЦХД позволило сформировать единое информационное пространство для подготовки и принятия управленческих решений, наладить обмен между различными ведомствами, тем самым повысить эффективность стратегического менеджмента за счет применения инструментальных средств анализа и моделирования процессов, происходящих на территории республики.

В настоящее время к серверам ЦХД подключено около 150 рабочих станций в Саха(Якутия)стате, более 50 в Минэкономразвития Республики Саха (Якутия), а также рабочие места в Информационном центре при Президенте РС(Я) и других министерствах республики. Таким образом, посредством доступа к ЦХД достигнут синтез управленческих структур, организован постоянный, интенсивный обмен выверенной информацией.

В перспективе информацию ЦХД предполагается до-

полнить блоком показателей муниципальной статистики. Решение этого вопроса позволит предоставить муниципальным органам управления оперативный доступ к данным ЦХД, в том числе использовать их для ведения паспортов муниципальных образований. Организация доступа администраций улусов к ЦХД, несомненно, будет способствовать развитию муниципальной статистики на высоком технологическом уровне.

Демонстрация работы ЦХД получила одобрительные отзывы участников семинара-совещания «Типовые решения внедрения ИКТ в регионах», проходившего в феврале 2005 г. в г. Якутске в рамках мероприятий федеральной целевой программы «Электронная Россия». Семинар был организован при участии специалистов Министерства экономического развития и торговли РФ, Росстата и ряда областей России.

Продолжается подписка на 1-е полугодие 2006 года!

Подписные индексы по каталогу агентства «Роспечать»:

71807 - для индивидуальных подписчиков; 70127 - для предприятий и организаций.

Подписной индекс по Объединенному каталогу

«Почта России» (том 1) - 41254.