



УДК 004.9

© 2007 г. С.А. Голубев,
Д.И. Петров,

М.В. Петропавловский, д-р техн. наук
(Марийский государственный технический университет, Йошкар-Ола)

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И МАТЕМАТИЧЕСКИЕ МОДЕЛИ В СИСТЕМЕ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ ГОСУДАРСТВЕННОЙ АККРЕДИТАЦИИ ВУЗОВ

Рассматриваются информационно-аналитические технологии системы поддержки принятия решений при государственной аккредитации высших учебных заведений.

Введение

В системе профессионального образования России функционируют тысячи образовательных учреждений. Ежегодно несколько сотен экспертных комиссий проводят работу по оценке качества их деятельности. Решение о выдачи образовательному учреждению свидетельства о государственной аккредитации принимает Аккредитационная коллегия Федеральной службы по надзору в сфере образования и науки. Очевидно, что без информационных аналитических технологий невозможно реализовать систему поддержки принятия решений (СППР) по государственной аккредитации.

Основным рабочим документом аналитической поддержки принятия решений по государственной аккредитации является «Представление образовательного учреждения к государственной аккредитации» [1].

Информационно-аналитические технологии СППР при государственной аккредитации

Подготовка «Представления...» – достаточно длительный и трудоемкий процесс, предполагающий наличие сложной информационно-программной технологии. Такая технология была разработана на основе Центрального банка данных государственной аккредитации (ЦБД ГА) и используется в практике работы Национального аккредитационного агентства в сфере образования. Она включает в себя три основных уровня: сбор данных; хранение и администрирование данных; обработка данных (рис. 1).

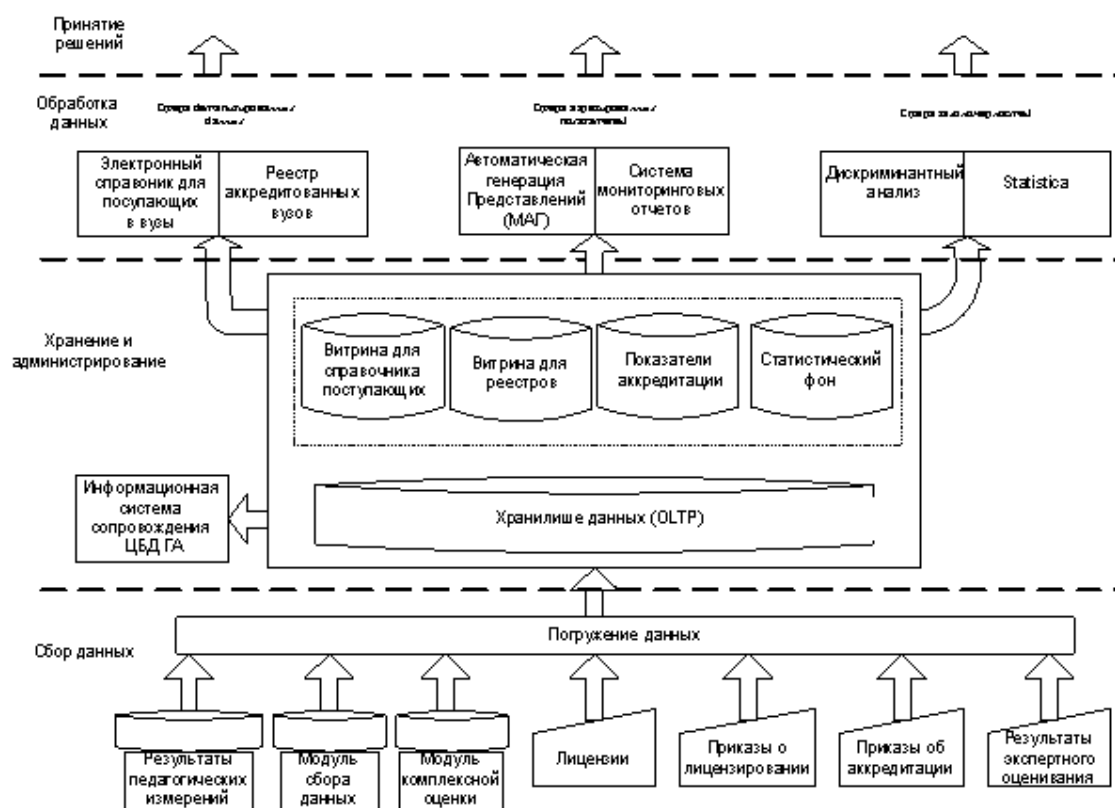


Рис. 1. Общая схема структуры ЦБД ГА как информационно-аналитической системы, построенной на основе многомерного хранилища данных.

Сбор данных происходит в полуавтоматическом режиме для информации, поставляемой в заданном формате с помощью электронных носителей («Модуль сбора данных», «Модуль комплексной оценки», «Результаты педагогических измерений» и т.д.) и в режиме ручного (операторского) ввода для информации, поставляемой на бумажном носителе (лицензии, приказы о лицензировании, приказы об аккредитации, результаты экспертного оценивания) [2].

Сбор данных завершается погружением данных в информационное хранилище (ЦБД ГА), построенное на основе реляционной СУБД MS SQL Server 2005.

На этапе сбора для доставки данных используется специализированное программное обеспечение «Модуль комплексной оценки» (МКО). Модуль предназначен для автоматизации работы экспертной комиссии, внесения информации в ЦБД ГА и подготовки представления образовательного учреждения (ОУ) к государственной аккредитации.

Хранение и администрирование заключаются в формировании витрин данных, специализированных для задач обработки данных, и осуществляются при помощи специально разработанного программного комплекса администрирования ЦБД ГА с использованием стандартных средств, пре-

доставляемых СУБД.

При обработке данных происходит формирование электронных документов, содержащих аналитические материалы, на основании которых в дальнейшем осуществляется принятие управленческих решений по государственной аккредитации образовательных учреждений.

Использование OLTP технологии позволяет осуществлять поддержку принятия решений с помощью просмотра и анализа данных в наиболее удобном виде, агрегирования и детализации информации, визуализации информации, генерации промежуточных и итоговых отчетов.

При проектировании системы аккредитации с количественными показателями возникает задача моделирования критериальных значений показателей [3]. Специальный программный модуль анализа данных – «Светофор», используя информацию ЦБД ГА, позволяет моделировать в диалоговом режиме процедуру установки критериев и визуализировать полученные результаты.

Для информационного сопровождения процедуры государственной аккредитации используется также «Модуль дискриминантного анализа», позволяющий определять вероятность принадлежности образовательного учреждения к каждой видовой группе вузов, а также положение образовательного учреждения на плоскости дискриминантных функций [3].

Одной из важнейших задач комплекса информационных технологий системы поддержки принятия решений по государственной аккредитации вузов является задача подготовки «Представления образовательного учреждения к государственной аккредитации». С этой целью Росаккредагентством разработана методика визуализации интегрированной информации в виде «Представления...». Для автоматизации подготовки «Представлений...» разработан программный модуль МАГ («Модуль автоматической генерации»), который позволяет проводить построение в автоматическом режиме проекта «Представления...», включая необходимые графические элементы – диаграммы ранжирования, диаграммы корреляции, эпюры и др. [4].

В настоящее время в системе поддержки принятия решений по государственной аккредитации используется более 10 разных видов «Представлений образовательных учреждений к государственной аккредитации».

По содержанию «Представления...» можно разделить на следующие:

установление, подтверждение или изменение государственного статуса образовательного учреждения;

аккредитация филиалов;

доаккредитация отдельных образовательных программ в уже аккредитованных образовательных учреждениях.

Также происходит разделение «Представлений...» по типу образовательного учреждения – высшее, среднее или дополнительное профессио-

нальное образование.

Из-за того, что все виды «Представлений...» имеют разную структуру и содержат разную информацию, возникает задача формализации их содержимого, которая решается при помощи XML-технологии.

XML – расширяемый язык разметки, фактически представляющий собой свод общих синтаксических правил. Технология XML предназначена для хранения структурированных данных или для обмена информацией между программами. Целью создания XML было обеспечение совместимости при передаче структурированных данных между разными системами обработки информации.

Формат XML – это человеко-ориентированный формат, одновременно понятный и человеку и компьютеру, он имеет строго определенный синтаксис и требования к анализу, что позволяет ему оставаться простым, эффективным и непротиворечивым, а иерархическая структура XML подходит для описания практически любых типов документов.

Созданная XML-структура описания содержимого «Представлений образовательных учреждений к государственной аккредитации» имеет иерархическую структуру, представленную ниже:

```
<представления>
  <тип ОУ>
    <вид ОУ>
      <элемент в представлении/>
      ....
      <элемент в представлении/>
    <вид ОУ/>
    ....
  <тип ОУ/>
  ....
</представления/>
```

В созданной XML-структуре были описаны все виды «Представлений...» и элементы, в них входящие. Выделены основные типы элементов: текст, графический элемент, таблица, список и т.д.

Технология генерации html-презентации как элемент СППР при государственной аккредитации

Для информационно-аналитического сопровождения работы коллегии подготавливаются не только «Представления...» на бумажных носителях, созданные при помощи программного модуля МАГ, но и html-презентация «Представлений...», подготовленная при помощи программного модуля «Программа автоматического формирования html-презентации «Представления образовательных учреждений к государственной аккредитации».

Html-презентация – электронная копия бумажных документов, кото-

рая позволяет повысить информативность и динамичность проведения коллегий. Используя такое техническое средство как проектор, можно демонстрировать на большом экране интегрированные аналитические данные того или иного образовательного учреждения, представленного к государственной аккредитации.

Технологически презентация выполняется в виде совокупности статических HTML-документов (веб-сайта). Такая технология позволяет запускать презентацию на любой операционной системе семейства Windows, не требовательна к вычислительным ресурсам компьютера и наиболее удобна для автоматической генерации презентации на основе заданных шаблонов. Кроме того, данная технология позволяет размещать информацию об образовательном учреждении на Интернет-ресурсе, что делает разноплановым применение подготовленной презентации.

Программный модуль формирования html-презентации позволяет на основе информации «Представлений образовательных учреждений к государственной аккредитации» (в электронном варианте формата MS Word) автоматически сформировать html-документы, в которых дублируется информация, указанная «Представлениях...». Принцип работы программного модуля приведен на рис. 2.

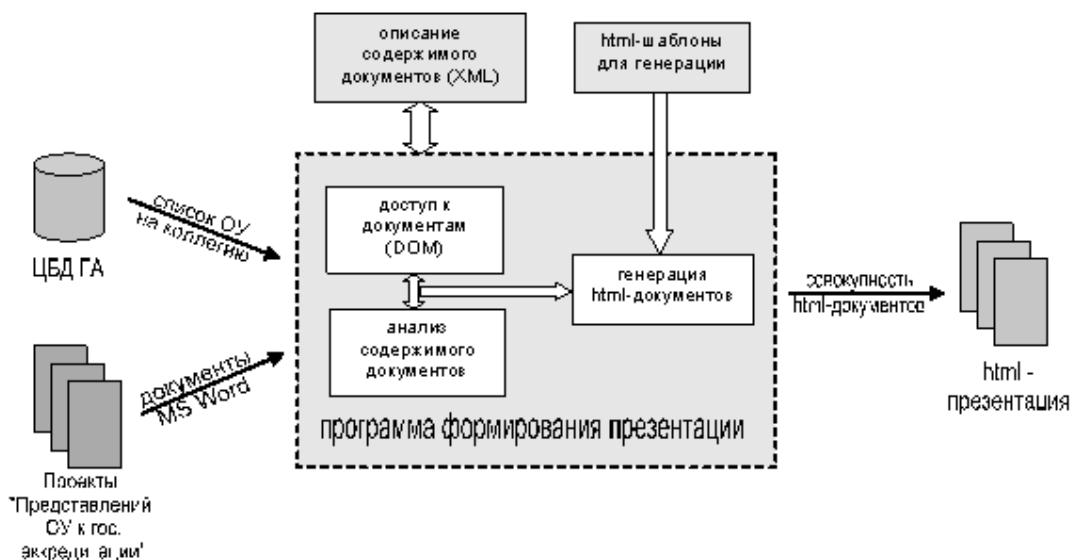


Рис. 2. Принцип работы программы автоматического формирования html-презентации «Представления образовательных учреждений к государственной аккредитации».

Входные данные программного модуля – информация Центрального банка данных государственной аккредитации о списке образовательных учреждений, выходящих на коллегия, и электронные варианты «Представлений образовательных учреждений к государственной аккредитации», созданных по технологии автоматической генерации документов.

Для доступа к информации Центрального банка данных государственной аккредитации используется технология ADO (ActiveX Data Objects)

– интерфейс программирования приложений для доступа к данным, разработанная компанией Microsoft и основанная на технологии компонентов ActiveX.

Для доступа и работы с электронными вариантами «Представлений образовательных учреждений к государственной аккредитации» (документами в формате MS Word) используется технология DOM (Document Object Model – «объектная модель документов»), реализованная зарегистрированными в операционной системе классами объектов и интерфейсов MS Office. Технология DOM представляет собой платформу-независимый программный интерфейс, позволяющий программам управлять содержимым документов, а также изменять их структуру и оформление. Так, любой документ известной структуры (в нашем случае документы MS Word) с помощью DOM может быть представлен в виде дерева узлов, каждый узел которого содержит элемент, атрибут, текстовый, графический или любой другой объект, а узлы связаны между собой отношениями родитель-потомок.

В процессе работы программы производится анализ документов «Представлений...» (в электронном варианте в формате MS Word). Программа пытается распознать указанные пользователем документы как «Представления» конкретных образовательных учреждений и производит в них поиск элементов (диаграмм, таблиц и т.д.), необходимых для формирования html-презентации. В программе реализован универсальный, настраиваемый поиск элементов документов, определяемый типом элемента, ключевыми словами поиска и направлением поиска.

Анализ содержимого документов и поиск в них нужных элементов, а также генерация html-презентации в программе формирования презентации осуществляются согласно XML-структуре описания содержимого «Представлений...», изложенного выше.

Также в программе реализована процедура корректировки информации: если программа в процессе анализа документов не смогла автоматически распознать документ или найти в нем необходимые элементы, то пользователь имеет возможность прямо в программе открыть нужный документ и вручную указать недостающую информацию.

В процессе генерации html-презентации программой создается совокупность графических файлов формата jpg (соответствующих графическим элементам «Представлений...») и файлов формата html, сохраняемых в указанной пользователем директории.

Структура html файлов определяется специально разработанными html-шаблонами для генерации – для каждого типа элемента документа разработан отдельный шаблон.

Html-презентация имеет строгую иерархическую структуру: список образовательных учреждений (по типу) → образовательное учреждение → элементы «Представления...».

Математические модели СППР при государственной аккредитации вузов

Задача формализации данных в законодательной базе определений видов вузов решается на основе моделирования критериальных значений системы показателей на массиве данных ЦБД ГА.

Для этого разработаны и исследованы различные модели интегрального индекса соответствия критериям аккредитации $J = J(x)$ [3]:

линейная модель:

$$J(x) = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{p_i}, \quad (1)$$

где $x = (x_1, \dots, x_n)^T$ – вектор значений характеристик вуза; $p = (p_1, \dots, p_n)^T$ – вектор соответствующих критериальных значений;

линейная модель с ограничением компенсации значений показателей:

$$J(x) = \frac{1}{n} \sum_{i=1}^n \begin{cases} \frac{x_i}{p_i}, & \text{если } \frac{x_i}{p_i} \leq 2, \\ 2 & \text{иначе.} \end{cases} \quad (2)$$

Данная модель индекса позволяет ограничить компенсацию значений показателей: существенное превышение значения одного показателя может компенсировать только одно значение «провального» показателя. Кроме того, индекс ограничен на значение сверху ($J(x) \leq 2$);

логистическая модель:

$$J(x) = c + b \sum_{i=1}^n \frac{1}{1 + e^{-1.7(x_i - p_i)}}. \quad (3)$$

Использование логистических функций основано на анализе закона распределения показателей. Выбор параметров модели осуществляется из условий

$$\begin{cases} J(0) = 0, \\ J(p) = 1. \end{cases} \Rightarrow \begin{cases} c = 1 - \frac{n}{2} b, \\ b = \frac{1}{\frac{n}{2} - \sum_{i=1}^n \frac{1}{1 + e^{1.7 p_i}}}; \end{cases} \quad (4)$$

логистическая модель с ограничением компенсаций значений показателей:

$$J(x) = b \sum_{i=1}^n \frac{1}{1 + e^{-1.7a(x_i - p_i)}}, \quad (5)$$

$$\begin{cases} J(p) = 1, \\ J(2p) = 2 - d. \end{cases} \Rightarrow \begin{cases} b = \frac{2}{n}, \\ a = \frac{\ln(\frac{2}{d} - 1)}{1.7 \max p_i}, \end{cases} \quad (6)$$

где $d > 0$ – достаточно малый параметр. Модификация предыдущей модели позволяет понизить компенсацию значений показателей за счет введения дополнительного ограничения на значение индекса в точке $2p$ (при двукратном превышении пороговых значений по всем показателям индекс не превышает двух);

многомерная модель:

$$J(x) = c + b \frac{1}{1 + e^{-1.7 \sum_{i=1}^n (x_i - p_i)}}. \quad (7)$$

Модель основана на многомерном нормальном законе распределения показателей. Параметры модели вычисляются из следующих условий:

$$\begin{cases} J(0) = 0, \\ J(p) = 1. \end{cases} \Rightarrow \begin{cases} c = 1 - 0,5b, \\ b = \frac{1}{0,5 - \frac{1}{1 + e^{\frac{1.7 \sum_{i=1}^n p_i}}}}. \end{cases} \quad (8)$$

Результаты проведенного сравнения и анализа расчетов на данных ЦБД ГА позволили выделить в качестве основной, используемой для поддержки принятия решения по аккредитации, линейную модель, построенную на факторах. В этой модели задача определения весовых коэффициентов для показателей решается методом главных компонент на основе статистического анализа массива данных ЦБД ГА.

Определены факторы для всех трех видов вузов и дана их интерпретация. Первый фактор интерпретируется как «Научно-исследовательская работа», второй – «Качественный состав ППС», третий – «Научно-методическая работа». Четвертый фактор характеризует широту спектра научной и учебной деятельности университетов.

Исследование процедуры определения критериальных значений аккредитационных показателей показало: для того чтобы критериальное значение играло стимулирующую роль, оно должно находиться в интервале от нижней квартили до медианы, причем ближе к нижней квартили (рис. 3). Использование порядковых статистик (квартили, медиана) позволяет более объективно оценивать распределение образовательных организаций по показателям, так как данные статистики являются устойчивыми к наличию статистических «выбросов» в выборке.

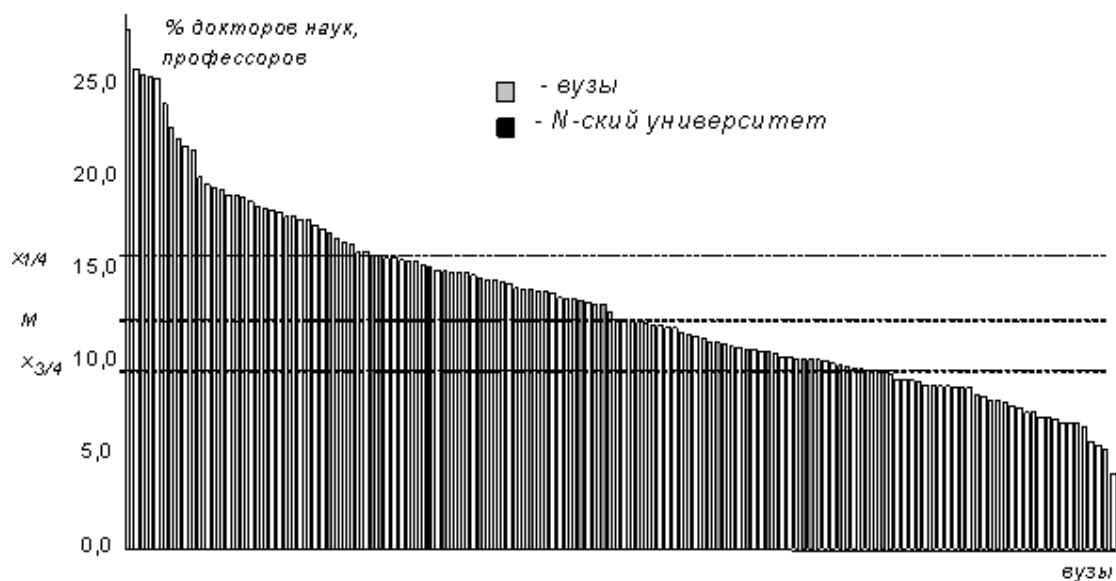


Рис. 3. Диаграмма ранжирования университетов по проценту лиц с ученой степенью доктора наук / ученым званием профессора в числе ППС.

В результате моделирования определены количественные критериальные значения показателей аккредитации для каждого вида, что позволило сделать процедуру принятия решений по государственной аккредитации в части установления вида объективной.

Таким образом, на основе расчетов индексов соответствия критериям аккредитации и критериальных значений показателей аккредитации формализованы прописанные в законе «О высшем и послевузовском образовании» виды высших учебных заведений: «университет», «академия», «институт».

На основе аппарата многомерного статистического анализа разработаны математические модели видовой классификации учреждений высшего профессионального образования [3], позволяющие определять интегральные характеристики вида (вероятность, координаты на дискриминантной плоскости) (рис. 4).

Использование методов дискриминантного анализа обусловлено тем, что решение об аккредитационном статусе учебного заведения принимается на основе обработки большого объема информации, при этом для части показателей критериальные значения не установлены, хотя они используются для анализа эффективности деятельности учебного заведения. Например, для университета необходимым условием является наличие эффективно работающих аспирантуры, докторантуры, диссертационных советов, реализация программ переподготовки и (или) повышения квалификации работников высшей квалификации, научных и научно-педагогических работников.

ЛИТЕРАТУРА

1. *Геворкян Е.Н.* Комплексная оценка высших учебных заведений / Е.Н. Геворкян, В.Г. Наводнов, Г.Н. Мотова, М.В. Петропавловский. – Изд.2-е. – М.: Центр государственной аккредитации, 2003.
2. *Наводнов В.Г., Петров Д.И.* Аналитические информационные технологии в системе государственной аккредитации образовательных учреждений: препринт. – Йошкар-Ола: Национальное аккредитационное агентство в сфере образования, 2006.
3. *Петропавловский М.В.* Математические модели государственной аккредитации учреждений профессионального образования. Монография. – Йошкар-Ола: Центр государственной аккредитации, 2004.
4. *Петропавловский М., Полевицков Д.А.* Язык генерации документов и деловой графики на основе xml-форматов Microsoft Word // IX Вавиловские чтения: Материалы постоянно действующей Всероссийской междисциплинарной научной конференции с международным участием: В 2-х ч. – Йошкар-Ола: МарГТУ, 2005.