

## Анкета участника

Информация	Описание
ФИО студента	Власенко Дарья Владимировна
Направление/специальность	080801.65 «Прикладная информатика (в экономике)»
Вуз	Государственное образовательное учреждение высшего профессионального образования Нижегородский государственный архитектурно-строительный университет
Вуз-партнер	Да
Город	Нижний Новгород
Кафедра	Информационные системы в экономике
ФИО зав. кафедрой	Папкина М.Д.
Тема ВКР	РАЗРАБОТКА ИНТЕГРИРОВАННОЙ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ РЕГИОНАЛЬНОГО УПРАВЛЕНИЯ
Версия Deductor	5.1
Дата защиты	24.06.2010
Оценка	Отлично
Руководитель ВКР	Канд. ф.-м. наук, доцент Прокопенко Наталья Юрьевна
Представлено	<ul style="list-style-type: none"> <li>▪ Анкета участника</li> <li>▪ Аннотация</li> <li>▪ Пояснительная записка</li> <li>▪ Сканированные титульные листы</li> <li>▪ Рецензия</li> <li>▪ Отзыв научного руководителя</li> <li>▪ Сценарии Deductor (с данными)</li> <li>▪ Презентация Power Point</li> </ul>

## Аннотация

### **РАЗРАБОТКА ИНТЕГРИРОВАННОЙ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ РЕГИОНАЛЬНОГО УПРАВЛЕНИЯ**

Регулирование жизнедеятельности региона невозможно без быстрого доступа к данным по важнейшим показателям развития муниципальных районов, без качественного прогнозирования этих показателей. Принятие обоснованных решений по вопросам развития региона базируется на обработке все увеличивающегося объема пространственно распределённой информации, на учете и анализе большого числа взаимосвязанных факторов. Необходимые для решения этих задач ведение баз данных, анализ данных и географически ориентированное представление данных может осуществляться с помощью интеграции различных информационных систем и технологий

Целью данной работы является разработка методологии использования различных информационных систем и технологий для интегрированной обработки экономической информации в системах поддержки принятия решений региональных органов управления.

Для достижения поставленной цели требовалось решить следующие задачи:

- определение общей структуры разрабатываемой информационно-аналитической системы;
- сбор и систематизация статистической информации;
- разработка структуры базы данных для социально-экономических показателей муниципальных районов Нижегородской области;
- создание интегрированного хранилища данных, а также организация обработки накопленной информации современными методами поддержки принятия решений;
- изучение и применение методов экономического прогнозирования;
- оценка текущего и будущего состояния социально-экономического развития муниципальных районов Нижегородской области;
- визуализация полученных результатов с использованием средств геоинформационных систем.

Объектом данного исследования является система управления экономикой Нижегородского региона.

Предмет исследования: информационные системы (офисные, интеллектуальные, геоинформационные).

Работа состоит из введения, трех глав, заключения, списка литературы, двух приложений.

В первой главе дипломного проекта излагаются теоретические аспекты автоматизации анализа экономической информации с использованием современных инструментальных средств, описываются цели, задачи, структура разрабатываемой ИАС.

Основной целью разрабатываемой системы является повышение эффективности принятия решений по развитию региона за счет внедрения комплексной ИС, которая способна оперативно проводить анализ информации и осуществлять прогнозирование основных макроэкономических показателей.

На ИАС возложены следующие основные функции: сбор, очистка, хранение, аналитическая обработка и группировка данных, построение краткосрочных прогнозов основных показателей

муниципальных районов и Нижегородского региона, получение аналитической и картографической отчетности.

Разработанная автоматизированная информационная система состоит из четырех подсистем, отвечающих за конкретные функции:

- подсистема сбора и хранения данных;
- подсистема очистки и подготовки данных к анализу и построению аналитических моделей;
- подсистема моделирования и прогнозирования;
- подсистема отображения данных.

В рамках данного дипломного проекта была определена система основных макроэкономических показателей региона, проведены сбор и систематизация статистических данных по районам Нижегородской области.

Информационно-эмпирическая база данного исследования формировалась на основе официальных данных федеральных и региональных органов государственной статистики, органов государственной власти Нижегородской области, а также данных, собранных на сайтах [www.icss.ac.ru](http://www.icss.ac.ru), [www.nizstat.sinn.ru](http://www.nizstat.sinn.ru).

В первой главе были рассмотрены две идеологии хранения информации: база данных и хранилище данных. В разрабатываемой ИАС были реализованы и использованы достоинства каждого из подходов: компактное хранение детализированных данных и поддержка больших баз данных, обеспечиваемые реляционными СУБД, а также простота настройки и хорошие времена отклика при работе с агрегированными данными, обеспечиваемые ХД.

В процессе создания ИАС и ее подсистемы – информационного хранилища как важная проблема выделяется обеспечение необходимого качества данных. Подсистема очистки и подготовки данных к анализу занимается выявлением и удалением ошибок и несоответствий в данных с целью улучшения качества данных.

Центральной подсистемой разрабатываемой ИАС является подсистема моделирования и прогнозирования. Построение моделей – универсальный способ анализа показателей, а прогноз является основополагающей составляющей планов всех субъектов экономической деятельности. В дипломном проекте для исследования динамики макроэкономических показателей и ВВП была выбрана модель регрессионного анализа, для решения задачи классификации районов по уровню развития использовались модели карта Кохонена и деревья решений, при построении прогнозной модели в данной подсистеме используются нейронные сети. Для реализации этих моделей были изучены методики расчета и анализа социально-экономических показателей, а также статистические и интеллектуальные методы для их прогнозирования.

Одной из важнейших составляющих аналитических технологий является визуализация – представление данных в виде, который обеспечивает наиболее эффективную работу лиц, принимающих решения. Выбранный способ визуализации должен максимально полно отражать поведение данных, содержащуюся в них информацию, тенденции, закономерности и т.д. Целью создания подсистемы аналитической отчетности и отображения данных в разрабатываемой ИАС является визуализация результатов анализа и прогнозирования: получение аналитических отчетов OLAP-средствами и карт с помощью ГИС.

Информационно-аналитическая система – это не готовый продукт или семейство продуктов. Успех реализации информационно-аналитической системы зависит от правильного выбора и максимально тесной интеграции используемых для его реализации средств. Для реализации указанных подсистем требуются специализированные программные инструментальные средства. В первой главе дается обзор и обоснование выбора инструментария.

Разрабатываемая интегрированная информационно-аналитическая система поддержки принятия решений региональных органов государственного управления строится на базе трех

систем: СУБД MS Access, аналитической системы Deductor и геоинформационной системы MapInfo. Ввод информации выполняется в приложении СУБД MS Access, очистка, предобработка, построение моделей и отчетов в Deductor Academic, а географическая визуализация результатов кластеризации в MapInfo.

Во второй главе дипломного проекта рассмотрена практическая реализация ИАС принятия решений регионального управления, состоящая из 4 подсистем, каждая из которых решает свои задачи. В соответствии с разработанной в главе 1 структурной моделью и с учетом специфики выбранного программного обеспечения создан пилотный вариант ИАС.

На этапе проектирования ИАС были рассмотрены три технологии моделирования: метод функционального моделирования IDEF0, метод построения диаграмм потоков данных (DFD) и метод описания бизнес-процессов IDEF3.

Подсистема сбора и хранения данных реализована средствами MS Access и АП Deductor. Структура базы данных в MS Access для социально-экономических показателей муниципальных районов Нижегородской области разработана в соответствии со стандартами проектирования. Были выделены основные сущности и атрибуты, определены первичные ключи, была создана инфологическая модель, датологическая модель, представлена схема данных.

Создание интегрированного хранилища данных «регион», а также организация обработки накопленной информации было реализовано на базе Deductor Warehouse. Так как в проекте предусмотрено использование учебной версии Academic Deductor Studio, где возможность прямого экспорта данных из MS Access отсутствует, то данные в ХД загружались из текстовых файлов районы.txt и показатели.txt. На этапе загрузки данных в хранилище было произведено редактирование данных с помощью инструмента «Парциальная предобработка», в результате чего были восстановлены пропущенные данные и удалены аномалии.

Подсистема моделирования и прогнозирования реализована на базе Deductor Studio. Она включает несколько прогностических и классификационных моделей.

Так как входной набор данных представляет собой годовые макроэкономические показатели 48 муниципальных районов Нижегородской области, то решено было построить комплекс моделей анализа и прогноза на примере Ардатовского района, который впоследствии можно будет применить ко всем остальным районам и таким образом автоматизировать процесс анализа и прогнозирования показателей районов. Для автоматизации в АП Deductor есть 2 возможности: с помощью «Динамического фильтра» при выгрузке информации из ХД и при помощи обработчика «Групповая обработка». Используя «Динамический фильтр», при каждом выполнении узла импорта данных из ХД будет выводиться окно запроса, где можно будет указать название района, информация по которому требуется для анализа и прогноза. Обработчик «Групповая обработка» позволяет создавать очень гибкие сценарии благодаря тому, что входной набор делится на части по указанным группам, и затем каждая группа отдельно «прогоняется» через копию цепочки узлов обработки прогностической модели.

Для того чтобы наглядно проследить динамику значений социально-экономических показателей Ардатовского района были построены OLAP-куб и на основе его различные кросс-диаграммы.

Для оценки зависимости макроэкономических показателей был проведен корреляционно-регрессионный анализ при помощи обработчика «Корреляционный анализ». После проведения корреляционного анализа становится доступным визуализатор «[Матрица корреляции](#)». В работе представлена матрица корреляции (а именно, результат объединения нескольких таблиц с помощью обработчика Слияние) для исходных данных. Анализ матрицы корреляции позволил сделать предположение, что на факторы «Доход бюджета» и «Расход бюджета», оказывают сильное влияние факторы «инвестиции», «сельское хозяйство» и «промышленность».

Для исследования зависимости «Расхода» от указанных факторов была построена множественная линейная регрессия по данным Ардатовского района. Для наглядной оценки качества построенной модели служит визуализатор Диаграмма рассеяния. Все расчетные теоретические значения находятся внутри доверительного интервала, качество модели множественной регрессии можно считать удовлетворительным, поэтому был построен прогноз

показателя» расход» на 2009 год (в качестве исходной информации были собраны данные за период 1997-2008 г.г.).

Была построена линейная регрессия для исследования зависимости показателя «доход бюджета» от факторов «инвестиции», «сельское хозяйство» и «промышленность».

Рассматривая имеющие годовые данные по макроэкономическим показателям «доход» и «расход» как временные ряды, было рассмотрено три подхода к решению задачи моделирования и краткосрочного прогнозирования их будущих значений. Были построены модель «Нейросетевой прогноз временного ряда», модель «Прогнозирование временного ряда с использованием парциальной обработки и скользящего окна», модель «Прогнозирование временного ряда на основе линейной регрессии». Все три модели были созданы в АП Deductor Academic. Для того чтобы выбрать наиболее точную модель прогноза сначала было проведено прогнозирование, используя данные за 1997-2007, для каждой модели были сопоставлены полученные прогнозные значения на 2008г. с реальными данными. В результате был сделан вывод, что наиболее точная модель для прогноза дохода и для расхода бюджета получилась при помощи методики «прогнозирование временного ряда с использованием парциальной обработки и скользящего окна».

Задача разбиения муниципальных районов Нижегородской области на классы была решена в АП Deductor Academic, используя обработчик «Карта Кохонена».

В результате кластеризации получилось три кластера:

- 0 класс «развитые районы» – средний уровень с/х, средний объем промышленности, умеренная з/п, низкий уровень безработицы, высокие доходы и расходы бюджета, средние инвестиции в предприятия;
- 1 класс «развивающиеся районы» – средний уровень с/х, высокий объем промышленности, средняя з/п, низкий уровень безработицы, средние доходы и расходы бюджета, высокие инвестиции в предприятия;
- 2 класс «наименее развитые районы» – низкий уровень с/х, низкий объем промышленности, умеренная з/п, средний уровень безработицы, низкие доходы и расходы бюджета, низкие инвестиции в предприятия.

Задача деления всех районов на различные классы по уровню дохода бюджета была решена также при помощи обработчиков «Квантование» и «Дерево решений». Для этого были найдены средние значения показателей по каждому району за весь исследуемый период. Значения поля «доход бюджета» при помощи обработчика «Квантование» было разбито на три диапазона «низкий доход», «средний доход», «высокий доход». В результате применения обработчика «Дерево решений» были получены правила, применяя которые можно определить к какому из трех возможных уровней дохода будет относиться произвольный район. Таблица сопряженности позволяет наиболее наглядно оценить результаты классификации, полученные с помощью модели «Дерево решений».

В рамках системы моделирования и прогнозирования была построена модель динамики ВРП МРО и ВВП России и выявлена общая динамика темпов данных показателей.

Информационная база для анализа динамики была получена с сайтов [www.icss.ac.ru](http://www.icss.ac.ru), [www.land-in.ru](http://www.land-in.ru), [www.nizstat.sinn.ru](http://www.nizstat.sinn.ru).

Для сравнительного анализа динамики ВВП РФ и ВРП Нижегородской области по абсолютным показателям были рассчитаны и приведены к базовому периоду относительные показатели темпа прироста ВВП и ВРП. Сравнительный анализ ВРП Нижегородской области и ВВП России показал, что показатели имеют одинаковую тенденцию изменения.

Для того чтобы отследить будущую динамику ВРП Нижегородской области был выделен набор объясняющих переменных данного показателя:

- X1 – численность занятых, тыс. чел.;
- X2 – численность безработных, тыс. чел.;

- X3 – среднегодовые основные фонды в ценах 2000 года, млрд. руб.;
- X4 – капитальные вложения в ценах 2000 года, млн. руб.;
- X5 – экспорт товаров, млн. долл.;
- X6 – импорт товаров, млн. долл.

С помощью обработчиков в Deductor был проведен корреляционный анализ отобранных показателей. В работе представлена матрица корреляции (а именно, результат объединения нескольких таблиц с помощью обработчика Слияние). При анализе матрицы корреляции выявлена тесная взаимосвязь объясняющих переменных, которая делает модель неустойчивой и неадекватной. В этой связи из модели были исключены сильнокоррелируемые признаки, а именно численность безработных, капитальные вложения, экспорт товаров, импорт товаров.

При построении множественной линейной регрессии учитывалось влияние на ВРП двух факторов: основные фонды и число занятого населения. Для расчета в АП Deductor коэффициентов линейной множественной регрессии был использован обработчик Линейная регрессия. Для оценки качества построенной модели была построена Диаграмма рассеяния. На основе модели множественной линейной регрессии был построен прогноз ВРП Нижегородского региона на 2010г.

Подсистема построения отчетов предполагает графическое представление данных в виде графиков, диаграмм, карт Кохонена, а также предполагает нанесение результатов анализа и прогнозирования экономических показателей на карту Нижегородской области для сравнительного анализа состояний муниципальных районов

Подсистема аналитической отчетности и географического отображения данных реализована средствами АП Deductor и ГИС MapInfo.

Отчеты играют большую роль в принятии решений, так как они реализуют одну из поставленных целей системы, а именно повышение эффективности регулирования экономики региона.

Отчеты в Deductor Studio представлены в виде древовидного иерархического списка, каждым узлом которого является отдельный аналитический отчет или папка, содержащая несколько отчетов со своим способом отображения информации (таблица, кросс-таблица, гистограмма, кросс-диаграмма).

Подсистема отчетности должна предоставить возможность наглядно увидеть распределение муниципальных районов по степени развития на карте Нижегородской области.

Визуализация полученных результатов кластеризации была выполнена средствами ГИС MapInfo. В результате на карте Нижегородской области районы представлены в цветовой градации по уровню развития.

В данном дипломном проекте разрабатывается идеология интеграции информационных систем, имеющих разные специализации. Информационно-аналитическая система поддержки принятия решений регионального управления реализуется на базе трех систем СУБД MS Access, аналитической системы Deductor и геоинформационной системы MapInfo. Интеграцию всех подсистем можно осуществить средствами MS Access. Для этого создана главная кнопочная форма, которая является панелью управления ИАС и позволяет работать в трех режимах в зависимости от решаемых задач: режим ввода в Access, режим моделирования и прогнозирования в Deductor и режим географического отображения данных в MapInfo.

На данный момент каждая из подсистем работает независимо от других, так как в работе были использованы учебные версии MapInfo и Deductor.

В третьей главе дипломного проекта рассмотрена эффективность внедрения ИАС управления региона, которая обуславливается действием ряда факторов: организационного, информационного и экономического характера.

Организационный эффект проявляется в освобождении работников от рутинных операций по систематизации и группировке статистических данных и многочисленных расчетов, увеличив тем самым время для проведения анализа, прогноза и оценки принимаемых управленческих решений.

Информационный фактор эффективности выражается в повышении уровня информированности лиц, принимающих решения.

Экономический фактор проявляется в том, что статистическая информация, имеющая целью полное и своевременное отражение и состояние объекта и причин, влияющих на его развитие, в конечном счете, направлена на точность принятия решений.

Для принятия правильных управленческих решений необходимо наличие интегрированного в систему мощного блока прогноза и моделирования, являющегося центральной компонентой информационно-аналитической системы регионального управления. Разрабатываемая в данной работе комплексная система анализа и прогноза экономических показателей Нижегородского региона имеет требуемую подсистему, реализованную на базе аналитической платформы Deductor.