

Применение онтологий для моделирования ИТ-инфраструктуры и описания систем энергетики*

Т. Н. Ворожцова, А. П. Костюченко, Н. Н. Макагонова, С. К. Скрипкин

Институт систем энергетики им. Л. А. Мелентьева СО РАН, Иркутск, Россия

e-mail: tnv@isem.sei.irk.ru, mak@isem.sei.irk.ru

The article addresses an application of the ontological approach for IT-infrastructure modeling and for description of the different aspects applied in studies of the energy and power systems.

В современном мире нормальное функционирование экономики страны, региона, города, предприятия или организации, а также комфортное существование человека невозможны без развитой инфраструктуры (экономической и социальной). Инфраструктуру образуют системы жизнеобеспечения, среди которых выделяют транспортные системы, системы электроснабжения и водоснабжения, сеть социальных учреждений: школы, больницы, аптеки, торговые учреждения, библиотеки, театры и т. д.

Сегодня на предприятиях особое внимание уделяют вопросам эффективной организации ИТ-инфраструктуры, которую образуют компьютеры, объединенные в локальные сети, используемые информационные технологии, а также информационные и автоматизированные системы, поддерживающие основные бизнес-процессы предприятия. От качества ИТ-инфраструктуры и от качества управления этой инфраструктурой во многом зависит эффективное выполнение этих процессов.

В России в последнее время ведутся работы по нескольким проектам, так или иначе связанным с ИТ-инфраструктурой:

— проект “Электронная Россия”, направленный на создание методологической и нормативной базы для полномасштабного применения информационно-коммуникационных технологий (ИКТ) в обществе, бизнесе, деятельности органов государственной власти и местного самоуправления;

— целевая программа “Информационно-телекоммуникационные ресурсы СО РАН”, в рамках которой выполнялась задача инвентаризации и описания информационных ресурсов Сибирского отделения Российской академии наук, а также предоставления доступа к этим ресурсам заинтересованным лицам;

— проект “Создание телекоммуникационной распределенной вычислительной инфраструктуры научных исследований: разработка методических основ и их применение для исследований в энергетике”.

Последний проект выполняется в Институте систем энергетики им. Л. А. Мелентьева СО РАН (ИСЭ СО РАН), и в нем ИТ-инфраструктура рассматривается как совокупность четырех взаимосвязанных компонентов: телекоммуникационной инфраструкту-

*Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (гранты № 07-07-00265а и № 08-07-00172а) и РГНФ (грант № 07-02-12112в).

© Институт вычислительных технологий Сибирского отделения Российской академии наук, 2008.

ры, вычислительной инфраструктуры, информационной инфраструктуры и интеллектуальной инфраструктуры [1], ее основное назначение — это обеспечение (техническое, программное/вычислительное и информационное) исследований, проводимых в институте (рис. 1). Выделение интеллектуальной составляющей инфраструктуры связано, во-первых, с потребностью в сохранении накопленных в институте знаний, представленных в виде методик проведения исследований, математических методов, алгоритмов, программ, баз данных, научных трудов, во-вторых, с необходимостью активного использования этих знаний при проведении исследований с целью формирования новых знаний и пополнения соответствующих баз знаний.

Для формализованного описания накопленных знаний, представленных в виде информационных ресурсов, был использован онтологический подход, позволяющий с помощью системы онтологий представить концепты предметной области и связи между ними на любом уровне иерархии с любой степенью детализации, а также отразить различные точки зрения на процессы предметной области.

В процессе проектирования информационной составляющей ИТ-инфраструктуры ИСЭМ СО РАН выявлены основные ресурсы, используемые исследователями при выполнении своих работ. Эти ресурсы объединены в следующие группы:

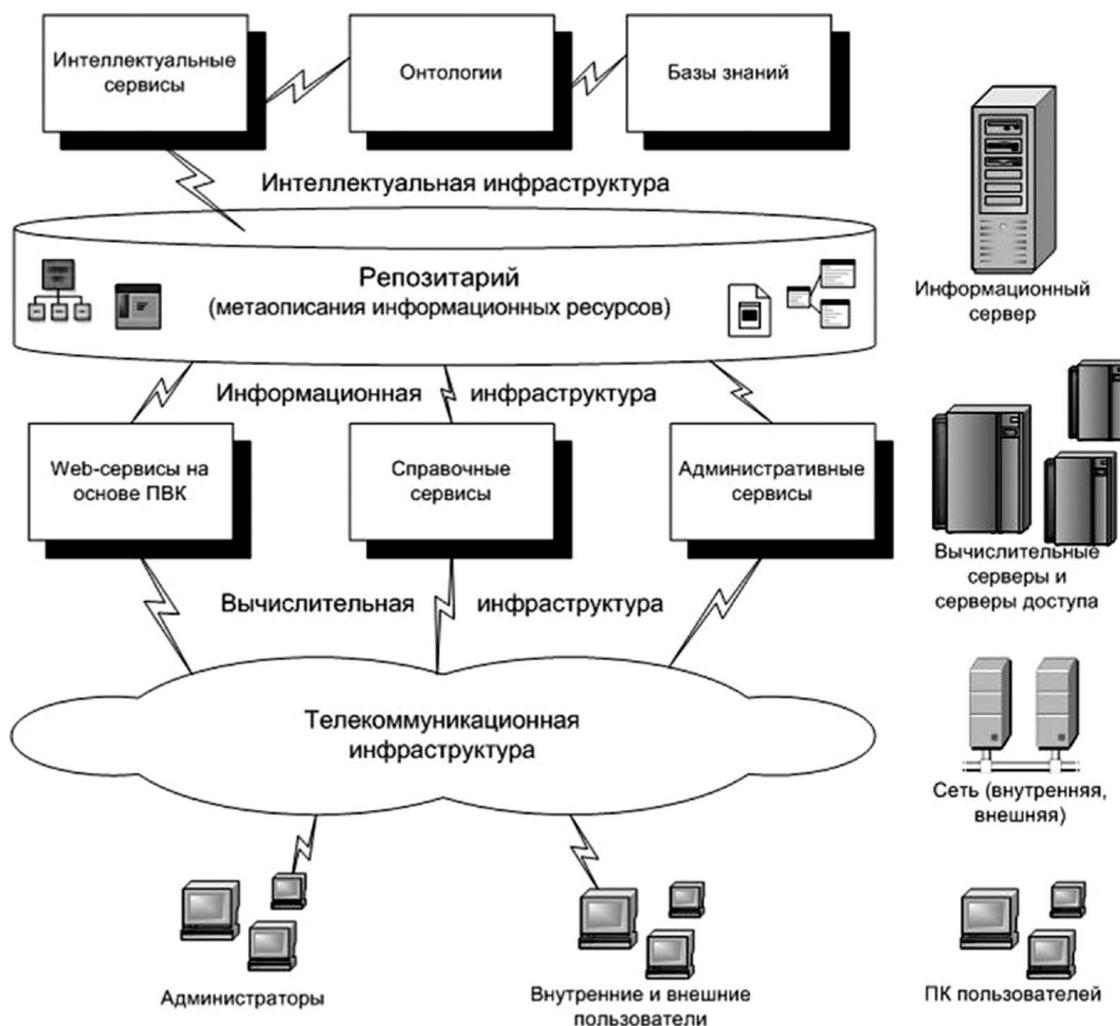


Рис. 1. ИТ-инфраструктура научных исследований

- *информационные* — базы данных, хранилища данных, наборы данных, онтологии;
 - *вычислительные* — программы и программные комплексы;
 - *трудовые* — научные сотрудники института, разработчики баз данных и программных комплексов;
 - *научно-исследовательские* — научно-исследовательские работы, проекты, гранты, конференции, семинары и т. п.;
 - *сервисные* — веб-сервисы, имеющиеся в институте;
 - *электронные* — публикации, отчеты;
 - *внутренние* — вспомогательная информация для работы сервисов.

На рис. 2 представлена онтология ИТ-инфраструктуры научной организации, отражающая ее основные компоненты: интеллектуальную, информационную, вычислительную и телекоммуникационную, показаны их связь и состав каждого из них. Для того чтобы представить различные аспекты функционирования инфраструктуры, необходимо разработать систему онтологий, описывающую как отдельный элемент инфраструктуры, так и совокупность взаимосвязанных элементов. В [2] предлагается выделять следующие типы онтологий: метаонтологии, онтологии предметной области или задач, онтологии приложений. Онтологию, на которой представлены основные концептуальные решения по структуре и организации ИТ-инфраструктуры, можно отнести к уровню метаонтологий. С помощью онтологий предметной области описываются информационные ресурсы, имеющиеся в институте, а также энергетические системы, исследования и терминология, используемая в области энергетики и энергетических исследований [3].

При разработке онтологии информационной составляющей ИТ-инфраструктуры были выделены следующие типы ресурсов и определены их характеристики:

- базы данных (БД) (название, назначение, дата создания, дата последней модификации, разработчик, программный комплекс, использующий БД, используемая СУБД, модель данных (ER-модель), расположение БД, доступ к БД);
 - программный комплекс (название, назначение, описание архитектуры, дата раз-



Рис. 2. Онтология информационных ресурсов ИТ-инфраструктуры

работки, дата последней модификации, разработчик, используемая БД, язык программирования, среда программирования, расположение ПК, доступ к ПК);

- онтологии (название, описание, составитель, расположение, инструментарий, формат представления, местонахождение онтологии);

- сотрудники (ФИО, должность, степень, звание, электронная почта, веб-страница);

- публикации (название, тип публикации, автор, выходные данные, аннотация, местонахождение ресурса, дата создания ресурса, тип ресурса, язык ресурса);

- исследования (название, описание, дата проведения, исполнитель, онтология исследования, используемый ПК, используемая БД, результаты исследований — отчет, публикации, местонахождение);

- научные конференции, семинары (название, дата проведения, место проведения, тематика, программа конференции, отчет по конференции, труды конференции);

- веб-сервисы (название, характеристика, разработчик, технология реализации, дата размещения сервиса, адрес сервиса, доступ к сервису);

- пользователи (имя, пароль, электронный адрес, организация, права, дата регистрации);

- сессии (пользователь, дата входа, дата окончания, вид завершения).

На рис. 3 представлен фрагмент онтологии, описывающий информационные ресурсы института и взаимосвязи, существующие между ними.

Для хранения основных элементов онтологии (концептов, атрибутов и связей) используется репозитарий [4, 5], являющийся базовым компонентом информационной инфраструктуры. С помощью ряда сервисов организован просмотр метаданных информационных ресурсов либо в виде онтологий, представленных в графическом формате системы СmapTools [6] (рис. 4, а), либо в формате HTML, т. е. в виде веб-страниц (рис. 4, б).

Онтологии предметной области были отнесены к информационным ресурсам, описание которых хранится в БД метаданных ресурсов, но одновременно сами онтологии в виде файлов хранятся в файловой системе репозитария. Инstrumentальная система СmapTools позволяет организовать систему онтологий через реализацию механизма гиперссылок.

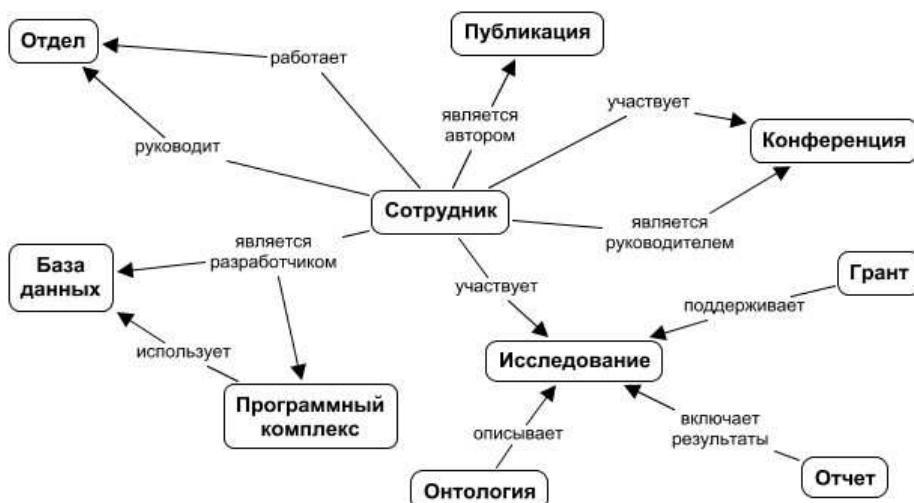


Рис. 3. Онтология информационных ресурсов ИТ-инфраструктуры

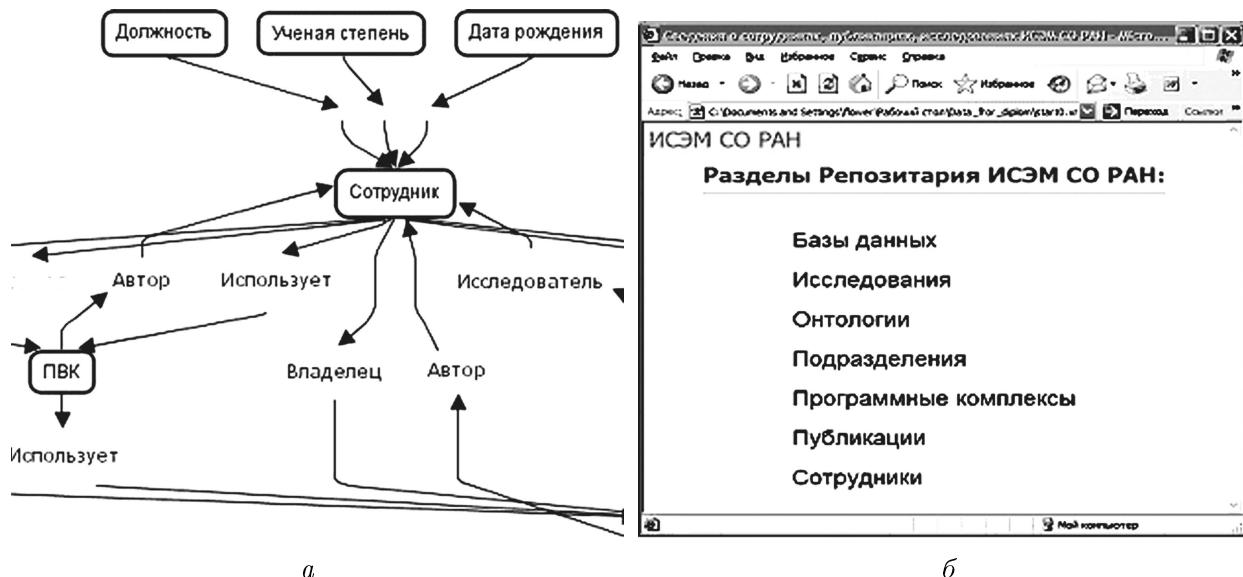


Рис. 4. Фрагмент модели метаданных в виде онтологии (*а*) и в виде HTML-страницы (*б*)

На рис. 5 представлена онтология топливно-энергетического комплекса, описывающая иерархию систем и объектов, входящих в его состав. Топливно-энергетический комплекс включает отдельные энергетические системы: электроэнергетические, углеснабжения, газоснабжения, нефте- и нефтепродуктоснабжения, ядерно-энергетические системы и системы теплоснабжения. Каждая энергетическая система включает совокупность объектов энергетики, подразделяющихся на объекты добычи, переработки, транспорта, хранения и потребления. Каждый объект онтологии (концепт) может быть связан с внешним файлом (текстовым или графическим), хранящим дополнительную информацию об объекте. Также это может быть ссылка и на другую онтологию, например, на рис. 5 концепт “углеснабжение” содержит ссылку на онтологию, в которой более

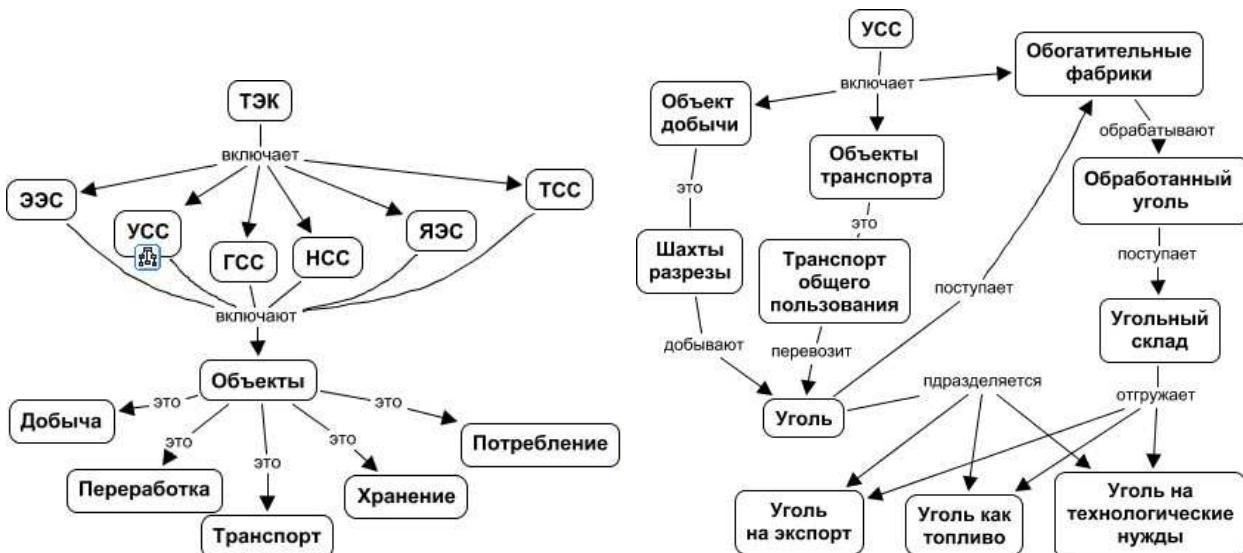


Рис. 5. Онтология топливно-энергетического комплекса

Рис. 6. Онтология углеснабжающей системы

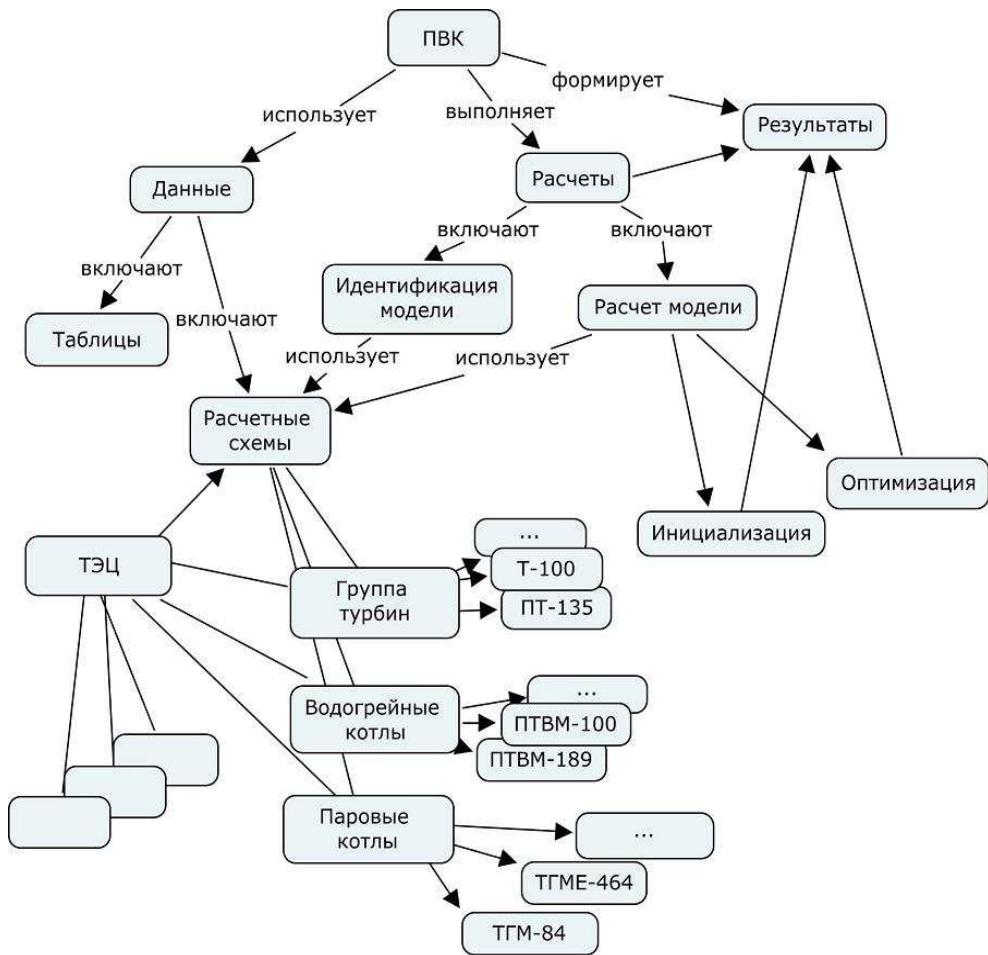


Рис. 7. Онтология приложения для расчета технологических схем

подробно рассмотрена углеснабжающая система (рис. 6). Она включает объекты добычи, транспорта и обогатительные фабрики. Объекты добычи — это шахты и разрезы, добывающие уголь, который с помощью транспорта общего пользования доставляется на обогатительные фабрики, обогащенный уголь поступает на угольный склад, откуда он отгружается либо на экспорт, либо на технологические нужды, либо как топливо потребителям.

Онтологии как средство формализованного описания понятий и отношений в некоторой предметной области используются не только для представления метаинформации и формализованного объединения отдельных разнородных компонентов ИТ-инфраструктуры, но и могут являться базой знаний о данной предметной области, описывая более детально отдельные ее компоненты на разных уровнях. На рис. 7 представлена онтология, используемая программным комплексом при разработке адаптируемых моделей паровых турбин, котлов и других компонентов теплоэнергетических установок. Этую онтологию можно отнести к онтологиям уровня приложений, которые описывают технологию проведения расчетов с указанием необходимых программ и данных [7].

Представленная на рис. 7 онтология используется для хранения необходимых данных о технологических схемах моделей тепловых электростанций и группах элементов схем, являющихся компонентами иерархии, а также о видах проводимых расчетов и вычисляемых параметрах по каждому элементу схемы.

Заключение

В разных предметных областях используются похожие подходы и решения для формализованного описания объектов и их взаимосвязей, причем как в текстовых форматах представления (XML, RDF, OWL), так и в графических. Онтологии являются универсальным средством представления разнородной информации, позволяющим использовать ее как человеку, так и компьютеру в различных целях, например: для формализованного описания информации (метаописания), для представления структур данных, для описания взаимосвязей между разнородными информационными объектами, а также для хранения и поиска данных. Дальнейшее развитие онтологического подхода для моделирования ИТ-инфраструктуры и описания предметной области исследований энергетики должно идти в двух направлениях: а) расширения объемов представляющей информации и наполнения репозитария; б) детализации формализованного описания взаимосвязей между онтологиями верхнего уровня и онтологиями приложений. При достаточно полном онтологическом описании предметной области исследований возможно применение онтологий для принятия управленческих решений и получения новых знаний.

Список литературы

- [1] МАССЕЛЬ Л.В. Методы и технологии создания ИТ-инфраструктуры научных исследований // Информационные и математические технологии в науке, технике и образовании / Тр. Х Байкальской Всерос. конф. “Информационные и математические технологии в науке, технике и образовании”. Ч. 1. Иркутск: ИСЭМ СО РАН, 2005. С. 57–65.
- [2] Тузовский А.Ф., Чириков С.В., Ямпольский В.З. Системы управления знаниями (методы и технологии). Томск: Изд-во НТЛ, 2005. 260 с.
- [3] Ворожцова Т.Н., Макагонова Н.Н., Скрипкин С.К. Применение онтологий для описания исследований энергетики // Информационные и математические технологии в науке и управлении / Тр. XII Байкальской Всерос. конф. “Информационные и математические технологии в науке и управлении”. Ч. 3. Иркутск: ИСЭМ СО РАН, 2007. С. 127–134.
- [4] Копайгородский А.Н., МАССЕЛЬ Л.В. Организация доступа к информационным объектам информационной инфраструктуры научных исследований // Тр. Междунар. конф. “Вычислительные и информационные технологии в науке, технике и образовании”. Т. 2. Павлодар: ТОО НПФ “ЭКО”, 2006. С. 473–480.
- [5] Копайгородский А.Н., МАССЕЛЬ Л.В. Разработка и интеграция основных компонентов информационной инфраструктуры научных исследований // Вест. ИрГТУ. 2006. № 2(26). С. 20–24.
- [6] IHMC CmapTools. <http://cmap.ihmc.us/>
- [7] Ворожцова Т.Н., Скрипкин С.К. Моделирование программно-вычислительного комплекса на основе онтологий // Информационные и математические технологии в науке, технике и образовании / Тр. Х Байкальской Всерос. конф. “Информационные и математические технологии в науке, технике и образовании”. Ч. 1. Иркутск: ИСЭМ СО РАН, 2005. С. 87–94.