

ГИС сегодня: состояние, перспективы, решения

Ю. И. Шокин¹, В. П. ПОТАПОВ^{2,*}

¹Институт вычислительных технологий СО РАН, Новосибирск, Россия

²Институт вычислительных технологий СО РАН, Кемеровский филиал, Россия

*Контактный e-mail: potapov@ict.sbras.ru

Дан краткий анализ современного состояния геоинформационных систем и рассмотрены перспективные направления их развития на основе новых информационных технологий (облачных и интеллектуальных систем, методов обработки больших данных), приведены примеры конкретных реализаций применительно к решению задач горнопромышленных регионов.

Ключевые слова: геоинформационная система, интеллектуальная геоинформационная система, дистанционное зондирование Земли, метаданные, BIG DATA, data mining, хранилище данных, облачный сервис, сервис-ориентированная архитектура, геопортал, извлечение знаний, мультиагентная система, онтология, база знаний.

Введение

Геоинформатика как наука находится на слиянии двух быстро развивающихся областей: информатики и науки о Земле. Сегодня исследования в этой области используют все новейшие методы вычислительной математики и компьютерной обработки больших объемов данных наблюдения Земли, реализуемые через географические информационные системы (ГИС), которые привлекают все большее внимание государства, промышленности и научно-исследовательских организаций. В настоящей работе авторы, не претендуя на полноту в этой стремительно развивающейся области знаний, делают попытку анализа состояния геоинформационных систем, рассматривают перспективные направления их развития исходя из результатов собственных исследований и многолетнего опыта их применения для решения конкретных задач.

В литературе встречается множество определений ГИС [1–6]:

ГИС — компьютерная система, которая используется для формирования, загрузки, хранения, визуализации и анализа пространственных данных [1];

ГИС — информационная система, предназначенная для работы с пространственными, или географическими, координатами [2];

ГИС — система для манипуляций с пространственными объектами, такими как точки, линии и полигоны, извлекающая данные с помощью специальных запросов с целью их последующего анализа [3];

ГИС — система, состоящая из пяти основных элементов: данных, аппаратных средств, программного обеспечения, процедур и людей, их использующих [4];

ГИС — программно-аппаратные комплексы, представляющие собой географический информационный сервис [5];

ГИС — информационная система, обеспечивающая сбор, хранение, обработку, доступ, отображение и распространение пространственно-кодированных данных [6].

Если проанализировать приведенные определения, то можно характеризовать ГИС как информационную систему, обеспечивающую сбор, хранение и анализ пространственной информации, т. е. как геоинформационный сервис.

История развития ГИС насчитывает уже более 50 лет [7]. Первые попытки их создания относятся к 1960-м гг., когда Р. Томлинсон (R.F. Tomlinson) в 1968 г. разработал первую Канадскую Географическую Информационную Систему для сельскохозяйственного агентства (CGIS). В Великобритании делались попытки создания автоматизированных картографических систем начиная с 1963 г. В США основным разработчиком методических вопросов ГИС стала Гарвардская лаборатория компьютерной графики, ее результаты послужили базой для создания промышленных ГИС. На этой основе в 1964 г. была создана первая система Sympar, авторы которой впоследствии возглавили известные теперь фирмы ESRI и Intergraph. В 1963 г. в Швеции была разработана ГИС для округа Upsala, направленная на решение задач землепользования.

В конце 1980-х гг. в США появились прикладные геоинформационные системы, ориентированные на решение природоохранных задач, и были выполнены работы по созданию электронных карт — так называемых природоохранных ГИС, а также работы по картированию лесов с использованием ГИС-технологий, аэро- и космической съемки. В начале 1990-х годов в США начаты комплексные работы по созданию на всю территорию страны системы по оценке биоразнообразия [7]. Несмотря на интенсивное развитие геоинформационных технологий, первые общедоступные полнофункциональные ГИС появились только после выхода на рынок персональных компьютеров, в первую очередь это связано с работой фирмы ESRI [8], которая выпустила свой основной продукт ArcView 2.0. Именно с этого момента и началось развитие ГИС как массовых программно-информационных комплексов для решения задач в различных прикладных областях.

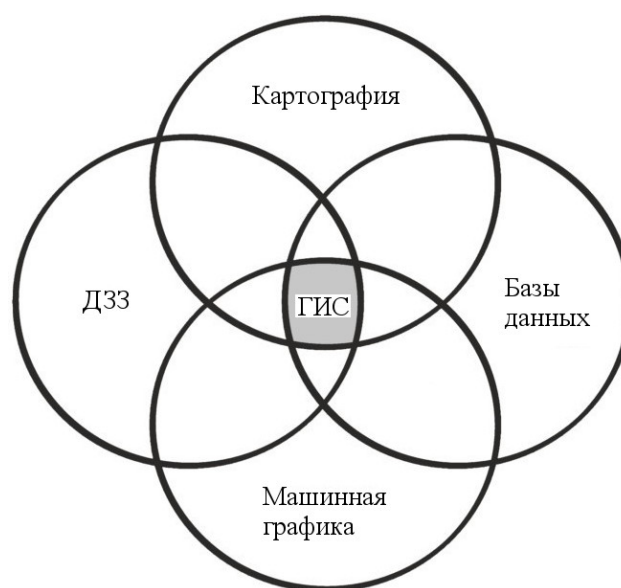


Рис. 1. Традиционные составляющие для разработки ГИС

Уже на первых этапах своего развития ГИС формировались как некие интегрированные программные комплексы, быстро реагирующие на изменения в области информационных технологий и ориентированные на решение широкого круга задач по обработке разнообразной пространственной информации (рис. 1).

1. Коммерческие и свободно распространяемые ГИС

В настоящее время на рынке ГИС существует не менее 100 коммерческих (проприетарных) систем и более 300 свободно распространяемых программных комплексов [9] для работы с пространственной информацией. В табл. 1 и 2 приведены наиболее популярные и широко распространенные системы этих двух больших классов. В данной работе мы не ставим перед собой задачу подробного анализа всего программного обеспечения ГИС, так как это уже сделано другими авторами, а опираемся на собственный многолетний опыт по использованию таких систем.

Остановимся на свободно распространяемых ГИС, которые находят все большее применение во всем мире, и это связано не только с их стоимостью и открытым кодом, но и с тем, что за последнее время их функциональность резко повысилась и некоторые из них, например QGIS, могут по ряду аспектов конкурировать даже с таким лидером рынка, как ESRI ArcGIS. Достаточно быстрая смена версий также привлекает пользователя, который может подключать свои собственные библиотеки или комплекс расчетных модулей для конкретных задач, формируя на базе таких систем персональную информационно-вычислительную среду обработки пространственных данных. Большую роль в развитии открытого программного обеспечения сыграл консорциум Open Source Initiative, который определил соответствие лицензии на программное обеспечение стандартам открытого программного обеспечения (ПО) (opensource.org) [5] и FreeSoftware Foundation (www.fsf.org). Именно благодаря их усилиям произошел быстрый рост свободно распространяемых ГИС, основные достоинства и недостатки которых приведены ниже.

Достоинства:

Большое количество разработчиков
Бесплатное приобретение
Частое обновление и исправление ошибок
Открытость кода позволяет пользователю самому вносить дополнения и разрабатывать собственную среду
Быстрое внедрение новых технологий и расширений системы
Широкий круг подключаемых библиотек и модулей, обеспеченных открытостью кода

Недостатки:

Недостаточно отлаженный программный код
Не всегда полная документация, сложное сопровождение
Иногда возникают проблемы качества ПО
Не все дополнения становятся общими для всех пользователей
Не всегда качественное ПО, особенно в первых версиях, недружественный интерфейс
Сопровождение их не всегда обеспечивается разработчиками

Бурное развитие ГИС и использование их для широкого круга задач в разных областях знаний, начиная от оборонных предприятий до медицины, заставили мировое сообщество задуматься о правилах и стандартах обработки пространственной информации, хотя эти процессы уже шли на начальных этапах развития ПО ГИС. Работы известных компаний и корпораций наиболее полно освещены в [5], поэтому здесь мы лишь кратко остановимся на более значимых из них, которые связаны как с инфраструктурой пространственных данных, так и со стандартами на пространственные данные OGC.

Т а б л и ц а 1. Коммерческие ГИС

Программный продукт	Производитель	Достоинства	Недостатки
1	2	3	4
ArcGIS 10.3	ESRI Inc., Калифорния, США http://www.esri.com	Широкая функциональность, встроенный язык программирования Python, многоплатформенность, мобильные приложения, поддержка разнообразных стандартов, в том числе OGC, сертифицикат ФЭСТЭК России	Высокая стоимость, сложность в эксплуатации, необходимость квалифицированного персонала
MapInfo 12.2	Pitney Bowes Inc., США http://www.pbinsight.com/welcome/marinfo/	Возможность решать основные задачи пространственного анализа, простота в использовании, низкая цена, наличие языка MapBasic, расширения Basic Microsoft, простой текстовый формат обмена данными, мобильные приложения для Microsoft Windows Mobile, обмен по протоколам WMS, WFS	Отсутствие полноценного 3D-модуля и поддержки других протоколов OGC, ограничения на создание условных обозначений, слабая многоплатформенность, сложность работы с веб-приложениями
Geomedia	Intergraph Corporation, США http://www.intergraph.com	Широкая функциональность, реализуемая различными дополнениями, неплохие возможности работы через Интернет, интеграция с пакетом данных ДЗЗ, реализация основных протоколов обмена пространственной информацией, возможность работать с геоинформационными БД большинства индустриальных форматов, эффективно интегрировать геоданные в единую информационную систему масштаба от рабочей группы до предприятия, полнофункциональность	Высокая стоимость, отсутствие мобильных приложений, слабая многоплатформенность, небольшой набор средств для анализа пространственной информации

Продолжение табл. 1

1	2	3	4
Autodesk AutoCad Map 3D/Civil 3D 2015	Autodesk Inc., США http://www.autodesk.com	Инструментальная ГИС. Обработка пространственных данных из различных источников, совмещение их с проектными данными из систем САПР в формате DWG семейства программ AutoCAD и возможность строить на основе этих данных сложные трехмерные модели, включая модели территории	Отсутствие собственного формата хранения пространственных данных, использование форматов других популярных ГИС или хранилищ пространственных данных. Для работы требуется дополнительно базовый пакет AutoCad. Отсутствие многоплатформенности и работы по стандартам обмена, достаточной сложности интерфейса, небольшой набор средств для анализа пространственной информации, сложная интеграция с другими пакетами. Относительно высокая стоимость
Bentley Map V8i	Bentley Inc., США www.bentley.com	Хорошо развитые функции 2D- и 3D-проектирования. Картографирование, планирование, проектирование. Поддержка создания, обработки, анализа и совместного использования 2D- и 3D-геопространственных данных. Поддержка стандарта Oracle Spatial обеспечивает централизованное хранение данных. Oracle Spatial Bentley Map позволяет редактировать 2D- и 3D-данные напрямую в любой стандартной среде Oracle Spatial. Корректное создание геопространственных объектов	Наличие базовой версии Microstation. Отсутствие мобильных приложений, слабая многоплатформенность и реализация стандартов обмена. Сложная организация связи с другими СУБД
ГИС Карта 2011	ЗАО "КБ Панорама", г. Москва, Россия http://www.gisinfo.ru/	Широкий набор программных продуктов универсального назначения, специализированные решения для различных задач обработки пространственной информации, частичная реализация протоколов обмена. Наилучший выбор для решения традиционных задач геодезии и картографии. Современный набор картографических сервисов, включая доступ Интернет и работу по стандартам обмена WMS, WFS, WMTS	Отсутствие полноценной возможности редактирования информации из внешних хранилищ. Поддержка стандартов OGC реализована только на уровне визуализации данных. Использование внутреннего стандарта SFX. Сложность интеграции с другими системами и хранилищами пространственных данных

Окончание табл. 1

1	2	3	4
GeoGraph/ GeoDraw/Geo Constructor	Центр геоинформационных исследований Института географии РАН, г. Москва, Россия	По своим возможностям близок к продукту ESRI ArcView, широко используется при обучении студентов основам ГИС	Отсутствие таких современных возможностей, как многоплатформенность, доступ и работа по стандартам обмена WMS, WFS, WMTS, мобильные платформы и др.
ГИС-платформа компаний CSoft	ЗАО "СиСофт-Терра", г. Москва, Россия http://www.csoft.ru/about/contacts.html	Современная инструментальная ГИС с широким набором прикладных модулей для различных инженерных приложений	Отсутствие возможности подключить внешние данные "на лету", как у большинства современных ГИС. Можно только импортировать пространственные данные внутрь хранилища. Сложности при экспорте растровых изображений в другие форматы вместе с координатной привязкой. Отсутствие многоплатформенности и работы по стандартным протоколам
ГИС-платформа ИнГео	ЗАО Центр системных исследований "Интегро", г. Уфа, Россия http://www.integro.ru	Инструментальная ГИС с разнообразным набором модулей для построения систем работы с пространственными данными по типу клиент-сервер	Отсутствие поддержки стандартов OGC (WMS, WFS и GML). Для разработки Интернет-приложений имеется собственный протокол, работающий вне стандартов OGC. Сложность работы с растрами, отсутствие мобильных приложений и многоплатформенности

Т а б л и ц а 2. Свободно распространяемые ГИС

ГИС	Разработчик	Особенности
1	2	3
GRASS 7.1 (Geographic Resources Analysis Support System)	GRASS (1982–1995 гг.) Исследовательская лаборатория CERL (Construction Engineering Research Laboratory), штат Иллинойс, США. GRASS 4.2 (1997 г.) Бэйлорский университет, штат Техас, США. GRASS 4.2.1 (1998 г.) Институт физической географии и экологии, Ганновер, Германия. Версия GRASS 5.0 (1999 г.) Официальный выпуск – 2002 г. (Centro per la Ricerca Scientifica e Tecnologica), Тренто, Италия http://gras.baylor.edu , http://gras.itc.it .	Среда GRASS – модульная система, предоставляющая доступ к более чем 300 программам для работы с двух- и трехмерными растровыми и векторными данными. По функциональным возможностям сравнима с продуктом ESRI ArcGIS уровня ArcInfo. По причине отсутствия удобного пользовательского графического интерфейса распространенность GRASS ограничена, используется преимущественно исследователями институтов и университетами. До недавнего времени рост числа пользователей сдерживала невозможность запуска GRASS на платформах MS-Windows без использования эмуляторов Linux или Unix платформ (например, Cygwin). Однако с выходом версии 6.3.0 эта проблема была решена
SAGA (System for Automated Geoscientific Analyses) v. 2.04	Геттингенский университет, Гамбургский университет, Германия http://www.saga-gis.org	Решение задач морфологического анализа, почвенное картирование и работа с изображениями, включая некоторые данные ДЗЗ. Наличие мощных средств работы с векторными данными, в том числе с расчетными сетками из других программ, большой набор вычислительных моделей
Kosmo	Результат объединения разработок компании SAIG (Испания) и ряда проектов с открытым исходным кодом (JUMP, JTS, GeoTools и др.)	Простота установки, понятный интерфейс, большой набор классических модулей для работы с пространственными данными и обширный набор поддерживаемых форматов и источников данных. Может быть использована для начального обучения

Продолжение табл. 2

1	2	3
<p>PLWIS (Integrated Land and Water Information System) v. 3.8 и 3.3 академическая</p>	<p>Компания ГТС (1980-е гг.), Голландия. В настоящее время основной координатор проекта — компания 52°North GmbH, Германия</p>	<p>Достаточно мощная ГИС с широким спектром обработки данных ДЗЗ. Решаются задачи сегментации и кластеризации карт. Реализованы все базовые операции пространственного анализа. Работает только с ОС семейства MS Windows</p>
<p>SADA (Spatial Analysis and Decision Assistance)</p>	<p>ORNL (Oak Ridge National Laboratory), штат Техас, США (www.ornl.gov)</p>	<p>Решение задач экологии, в том числе оценка риска и экономические модели. Включает все типовые операции для работы с пространственными данными, имеет возможность добавления собственных моделей</p>
<p>gvSIG</p>	<p>Развивается от правительственного гранта Испания (министерство транспорта Валенсии) с 2003 г. Первая рабочая версия программы появилась в 2004 г. http://www.gvsig.org/web</p>	<p>Система аналогична ESRI ArcView. Работает в Windows, Linux, MacOSX. Является клиентом для сервисов WMS, WFS и WCS. Достаточно просто писать дополнения к системе на языках Java, Python. Имеет практически все свойства базовых ГИС и собственный веб-браузер. Существует мобильная версия</p>
<p>Landserf v. 2.3 (2009)</p>	<p>Университет Сити Лондон, Великобритания, http://www.landserf.org</p>	<p>Наиболее приспособлена для анализа топоповерхностей, включая их морфологию. Манипулирует как векторными, так и растровыми моделями, включая 3D-представление. Имеет развитые функции для фрактального анализа. Интегрируется с GPS-приемниками. Среда разработки — Java. Работает в Windows, MacOSX, Unix, Linux и др.</p>
<p>MapWindow GIS v. 6.1.2</p>	<p>Geospatial Software Lab университета штата Айдахо, США, http://mapwindow6.codeplex.com/</p>	<p>Поддерживает VB.Net и C Sharp. Работает только в среде Windows. Имеет большой набор классических модулей для работы с пространственными данными. Единственный поддерживаемый формат векторных данных — shp-файлы. Проста в обращении. Легко встраивается под Microsoft Office</p>

Окончание табл. 2

1	2	3
OSGeoLive	Консорциум OSGF или OSGeo. Все проекты Фонда распространяются под открытыми лицензиями, сертифицированными OSI	Набор инструментальных средств для создания ГИС. Свыше 50 различных приложений, тестовые наборы данных. Работает в среде Linux под виртуальной машиной. Используя систему, можно построить собственную среду, опираясь на базовые ГИС, входящие в набор, так как в ней содержатся все современные средства
QGIS (Quantum GIS) 2.10 (Pisa)	http://www.qgis.org Развивается благодаря усилиям волонтеров из различных стран мира	Одна из самых многофункциональных ГИС, дружественная к пользователю, с открытым исходным кодом, позволяющая управлять гео данными, отображать, редактировать и анализировать их, а также создавать макеты карт. Поддерживает различные СУБД. Благодаря интеграции с GRASS обладает мощным аналитическим функционалом. QGIS работает в Linux, Unix, Mac OSX и Windows. Намечается реализация под Android. Работает с ДЗЗ. Постоянно обновляется, включает множество встраиваемых моделей для морфологического анализа, решения экологических задач. Имеются интерфейсы к языкам Python, Java, R

Стандартизация необходима для систематизации геоинформационных ресурсов, заключающейся в необходимости обеспечения доступности пространственных данных, информирования пользователей об их существовании, возможности поиска данных нужной тематики на определенную территорию и предоставления данных и их интеграции. Решается эта проблема путем создания инфраструктуры пространственных данных (ИПД, SDI) — систем базовых пространственных данных, метаданных, географических информационных узлов и стандартов для обеспечения доступа и обмена географическими информационными ресурсами [10]. В нашей стране ситуация стала меняться, когда в 2004 г. ГИС-Ассоциация выиграла конкурс Министерства экономического развития РФ на разработку Концепции создания российской ИПД. В 2006 г. Концепция была одобрена Распоряжением Правительства Российской Федерации, что вывело проблему ИПД на государственный уровень. Реализацию Концепции планировалось осуществить в 2006–2015 гг., однако этот процесс продвигается крайне медленно. Так, в период с 2006 по 2009 г. были разработаны только ГОСТ Р 52573–2006 “Географическая информация. Метаданные” и ГОСТ Р 53339–2009 “Данные пространственные базовые” и ряд других [11]. Сам термин “инфраструктура пространственных данных” впервые предложен Национальным исследовательским Обществом США еще в 1993 г. Именно тогда за рубежом в этом направлении были начаты работы, и сегодня наиболее развитые страны уже имеют такую структуру. Пространственная инфраструктура данных включает: взаимосвязанные, взаимодействующие системы и пространственные базы данных, содержащие пространственные данные и метаданные подходящего содержания и качества, информацию и геоинформационные технологии в соответствии с общепринятыми стандартами, законодательством, организационными структурами, экономическими решениями и человеческие ресурсы, а также пользователей, создающих геоинформационное общество [12]. Пространственная инфраструктура данных имеется в более чем 30 странах и используется повсеместно:

- NSDI (National Spatial Data Infrastructure) в США;
- NCGI (National Clearinghouse Geoinformatie) в Нидерландах;
- ASDI (Australian Spatial Data Infrastructure) в Австралии;
- GDI-DE (Geodateninfrastruktur Deutschland) в Германии;
- INSPIRE (Infrastructure for Spatial Information in Europe) в странах Европы.

Эти работы и послужили толчком к развитию геоинформационных систем, особенно с открытым кодом. Данные системы используются в различных областях знаний, что обусловлено их преимуществами [9–13].

На сегодняшний день существует несколько крупных ГИС-проектов, координирующих разработки в данной области. Первый из них — Open Geospatial Consortium (OGC). Его цель — разработка открытых стандартов для служб, основанных на геолокации. Создание открытых стандартов упрощает взаимодействие между проектами и дает возможность кому угодно разрабатывать ГИС без уплаты отчислений.

Второй крупный проект — Open-Source Geospatial Foundation (OSGeo) — ставит своей целью создание и поддержку свободного геоинформационного программного обеспечения высокого качества. Программное обеспечение, разработанное и поддерживаемое OSGeo, включает ГИС GRASS, а также GDAL и OGR (библиотеки для работы с растровыми и векторными геоданными), MapServer, PostGIS (геоинформационную базу данных на основе СУБД PostgreSQL), Quantum ГИС и множество других. Еще один проект — OSM или OpenStreetMap — свободная карта мира, в которую может вносить исправления любой желающий. OSM следует основным принципам Википедии.

Т а б л и ц а 3. Распространенность ГИС

Свободные ГИС	Количество ссылок	Коммерческие ГИС	Количество ссылок
PostGIS	9144	ArcGIS	81533
GRASS	7155	ArcView	72426
GeoDa	6928	ArcInfo	43257
ILWIS	6544	MapInfo	23666
QGIS	3726	ERDAS	23336
gvSIG	1980	IDRISI	8624
DivaGIS	1674	ENVI	8294
GDAL	1666	eCognition	6591
uDIG	1224	Oracle Spatial	4337
SAGA	1046	PCI Geomatica	1354
OpenModeller	524	GlobalMapper	945
SpatialLite	380	—	—
SEXTANTE	362	—	—
ORFEOtoolbox	204	—	—

Из табл. 3 для различных ГИС, полученной с использованием Google Scholar [14], отчетливо видна распространенность наиболее популярных коммерческих и свободных ГИС.

2. Технологии создания распределенных ГИС (интерактивные онлайн-овые картографические сервисы, геопорталы, облачные сервисы, кластерные системы)

Уже на первых этапах своего развития ГИС подразумевали работу в режиме, когда клиент пользовался не своим локальным компьютером, а работал с удаленным по сети сервером. Иными словами, использовалась привычная сегодня модель клиент-сервер. При этом пользователь фактически получал замену своего настольного приложения через сервер, что позволяло организовывать коллективную работу. Появился и соответствующий термин “распределенные пространственные вычисления” (Digital Geospatial Computing — DGC), который относится именно к архитектуре клиент-сервер и обозначает возможность работы в удаленном режиме, например через протокол HTTP (передача гипертекстовых сообщений). Исторически первая пространственная информация была передана по сети DARPA, и первый картографический сервер был разработан в это же время фирмой Hexox [7]. В 1994 г. был основан Федеральный комитет США по географическим данным, который занялся вопросами распределенных геовычислений и передачи пространственной информации. Впоследствии были созданы консорциум OGC (Open Geospatial Consortium) и технический комитет 211 ISO [5], призванные решать вопросы стандартизации интерфейсов для работы и обмена пространственной информацией в распределенной вычислительной среде. В 1995 г. созданы первые приложения типа Mapquest. В 1996 г. ESRI, Intergraph и другие компании начали реализацию первых программных продуктов, соответствующих разрабатываемым стандартам OGC, а в 1998 г. появилась первая разработка по виртуальной цифровой Земле, сегодня известной как Google Earth. При реализации алгоритмов для этих систем применялись

математическая статистика, вычислительная геометрия, комбинаторная топология, которые затем стали типовыми методами, используемыми при разработке распределенных геоинформационных систем.

Одной из ведущих фирм, развивающихся в этом направлении, остается ESRI, которая имеет большой комплекс продуктов для создания распределенных ГИС. Это серверные продукты, одни из них существуют давно (ArcSDE, ArcIMS), другие только начинают осваиваться рынком (ArcGIS Server, ГИС Portal Toolkit). Значительно позже появились разработки фирм Intergraph, Autodesk и др. [15]. В связи с развитием телекоммуникационных и информационных систем, прежде всего обусловленных Интернет-технологиями, таких как веб-сервисы [16], GRID-системы [17], облачные сервисы [18], распределенные ГИС перешли от локальных сетей на использование больших объемов геоданных, разбросанных по всему миру.

Многие авторы по-разному классифицируют этапы развития ГИС [5], но все признают решающую роль стандартизации, которую проводят консорциум OSG и ISO. Именно благодаря их усилиям сегодня можно выделить несколько направлений, по которым развиваются распределенные ГИС, — это онлайн-картографические сервисы, порталы, системы облачных сервисов и использование вычислительных кластеров.

Наиболее известным решением в области создания распределенных ГИС являются онлайн-картографические сервисы. Самые распространенные из них — Google Maps, Google Earth и NASA WorldWind — содержат информацию по всей земной поверхности. В нашей стране аналогичные задачи решает система “Космоснимки” (<http://www.Kosmosnimki.ru>), созданная Федеральным космическим агентством “Роскосмос”. Отметим, что вместо обычных интерактивных ГИС-карт такие системы не только отображают картографическую информацию, но и актуализируются за счет космоснимков, оперативно получаемых из соответствующих баз данных. Кроме того, в некоторых системах обеспечивается трехмерная визуализация. В этом направлении развиваются и другие системы, особенно связанные с получением данных ДЗЗ. Именно интерактивные карты применяются для построения различных, в том числе и локальных, систем, связанных с обработкой геоинформации. Наиболее часто используется так называемая Mashup-технология [19], представляющая собой набор интерфейсов для создания интерактивных карт. Общая схема работы Mashup-интерфейсов приведена на рис. 2. Ее дальнейшее развитие предусматривает:

- переход от индивидуальных решений к платформам для интеграции разнообразных динамических источников;
- развитие семантики в направлении таксономии, создание новых типов метаданных, контролируемых словарей, переход к естественным категориям;
- расширение инфраструктуры, интеграцию содержания с минимальными технологическими элементами (XML);
- тесное взаимодействие с системами искусственного интеллекта;
- комбинирование интерфейсов API и семантической платформы для включения неструктурируемого содержимого в Mashup.

Данная технология не единственная для создания интерактивных карт-приложений [5], однако она имеет широкое распространение и развивается в рассмотренных выше направлениях.

Онлайн-картографические сервисы получили развитие и в технологии геопорталов, описанию и систематизации которых посвящено множество статей, наиболее полной из них, по нашему мнению, является работа [20]. Геопортал можно определить как



Рис. 2. Схема работы Mashup-технологии

средство доступа к распределенным сетевым ресурсам пространственных данных и гео-сервисов, которые могут быть найдены на геопортале как исходной точке входа в сеть серверов [21]. В общем виде, согласно этому определению, геопортал является распределенной информационной системой, предоставляющей информацию по определенному типу запроса, своего рода системным интегратором различных типов геопро пространственных данных с последующим представлением сервисов определенного вида.

Рассмотрим некоторые аспекты создания геопорталов. Так, для полного описания геоинформационной системы (например, системы мониторинга конкретной направленности) требуются самые разнообразные и неоднородные геоданные, которые можно условно разделить на локальные и глобальные. Под локальными данными понимают в основном точечные экспериментальные замеры тех или иных параметров, зависящие в том числе и от времени, т. е. данные, представляющие собой временные ряды любой длительности либо произвольное непронумерованное множество данных. Глобальные данные чаще всего могут представлять собой карты, схемы, космоснимки, которые имеют значительный пространственный охват и различную временную привязку.

Отметим, что каждый тип данных обрабатывается по собственному алгоритму и может иметь различные формы отчетности (графики, таблицы, карты, зависимости). Учитывая все вышесказанное, информационную сервисную геоинформационную систему можно представить в виде

$$G = \{H(L, B), M, P\}\{S\}, \quad (1)$$

где L — множество локальных геоданных; B — множество глобальных геоданных; M — множество моделей процессов; H — хранилище данных (множество структурированных баз); P — набор решающих правил; S — множество результатов и выводов (сервисов). Такое описание системы позволяет четко разграничить ее блоки, а также определить их функции.

Из формулы (1) становится понятной архитектура системы с точки зрения ее функциональных особенностей. Первые два множества $H(L, B)$ фактически представляют собой структурированную систему (хранилище) баз данных, формируемых динамически из информационных потоков по любой пространственной обстановке. При этом для нее совершенно безразлично, где находятся эти данные. Однако должно безоговорочно выполняться такое условие, как наличие широкополосного канала (Интернет) для передачи данных. В хранилище данных происходят концентрирование и обработка геоданных на уровне соответствующих баз. Важным является сам объект описания, он может динамически изменяться, для него формируется набор правил, представляющих собой соответствующий сервис, а хранилище данных — своеобразный концентратор и конвертор информационных потоков.

Таким образом, создается специализированная информационная среда, которую можно представить как геопортал. Он является, с одной стороны, интегратором, актуализирующим данные по конкретному объекту, а с другой — определяет правила и методы, по которым ведется обработка геоданных, или сервисы.

Множество M представляет собой набор алгоритмов и программ, которые предназначаются для обработки геоданных с учетом особенностей конкретного объекта. Эти элементы должны располагаться в пределах виртуальной сети, формируемой геопорталом. Иными словами, геопортал имеет набор описаний моделей (метаописаний) для вычислений, которые располагаются в пределах доступной части сети. Фактически в геопортале создается некоторая виртуальная лаборатория, представляющая собой специализированный вычислительный кластер, позволяющий решать множество задач и реализуемый в виде сервисов. Для самого пользователя эта часть фактически закрыта, так как он должен только определить методы, которые считает необходимыми при решении его задач. В геопортале по умолчанию может действовать типовой конвейер вычислений для конкретного типа пространственных данных, определяемый набором правил, “защитых” в портале.

Множество решающих правил P представляет собой достаточно новый инструмент, до сегодняшнего времени практически не применявшийся при создании геопорталов, используемый в интеллектуальных ГИС. Это обусловлено, с одной стороны, появлением совершенно новых типов геоданных, используемых для решения разнообразных задач, а с другой — трудоемкостью и нестандартностью их обработки. К таким данным в первую очередь относятся данные зондирования Земли, а также радарные данные и “облака” пространственных точек. За счет них создаются значительные объемы данных, которые растут при обработке путем генерирования специфических промежуточных данных. Большие наборы таких геоданных требуют в основном “ручной” обработки с заданием соответствующих параметров. Для них может работать принцип “по аналогии”, но это не всегда удается, поскольку для решения таких задач требуется достаточно опытный специалист. Именно поэтому специфика новых типов данных требует создания некоторого набора решающих правил, которые устанавливают как последовательность действий, так и задание конкретного набора вычислительных параметров, отработанных при расчетах для подобных объектов и множества данных.

Решающие правила физически представляют собой набор “подсказок”, необходимых для получения конкретного результата. Введение множества решающих правил обусловлено и тем простым фактом, что для обработки пространственных данных, включая ДЗЗ, существует огромное количество алгоритмов и пакетов прикладных программ, в которых пользователь может просто утонуть. Поэтому в геопортале может быть ре-

лизован принцип краудсорсинга (crowdsourcing), работающий по схеме “данные → услуга”. Иными словами, пользователь геопортала поставляет свои геоданные и на выходе получает готовый результат, формируемый в соответствии со спецификой исследуемого объекта.

Множество S представляет собой специфические генераторы отчетов, которые помогают пользователю в простой и наглядной форме получать результаты обработки множества геоданных в соответствии с их спецификой. Это могут быть карты, графики, таблицы, логические выводы и другие формы представления итоговых знаний. Особенности построения геопортала позволяют вести обработку как “сырых” данных, так и результатов работы любого из блоков, а также создавать как специализированные, так и универсальные, легко расширяемые и настраиваемые подсистемы, решающие разнообразные задачи. Такая система может быть использована также при создании различных учебных курсов для обучения студентов в интерактивном режиме с обработкой конкретных моделей и расчетов.

В работе [22] рассмотрены различные схемы создания геопорталов и приведены примеры их работы. В нашем подходе, согласно формуле (1), геопортал имеет расширенные функции за счет введения множеств M , P , S . Именно они обеспечивают основное отличие предлагаемого нами подхода к созданию геопорталов, которые становятся активной информационно-вычислительной средой для обработки и анализа пространственных данных по сравнению с [21], где среда обеспечивает в основном поиск и простейший анализ геоданных на уровне типовых средств пространственного анализа. При этом в соответствии со спецификой задачи могут создаваться виртуальные лаборатории, ориентированные на ее решение. Особенностью предлагаемого подхода является активное использование различных видов ДЗЗ, которые пока еще не в полной степени используются для решения прикладных задач.

Рассмотрим такую технологию, как *облачные вычисления*. Описанный выше подход к созданию геопорталов опирается именно на облачные вычисления, точнее, на определенные виды сервисов, предоставляемых ими [23–25].

Облачный сервис SAAS (software as service) предполагает предоставление определенного программного обеспечения, в том числе конкретных базовых ГИС и расчетных систем, доступ к которым может быть реализован средствами Mashup. Для реализации данного вида сервиса необходимо иметь соответствующие каталоги доступного (в том числе и по Интернет) программного обеспечения, метаописания геоданных и некоторый внутренний стандарт для работы с таким видом сервиса.

Сервис DAAS (data as service) представляет собой способ получения различных геоданных, в том числе генерируемых на основе расчетов. К этому виду сервиса можно отнести получение данных ДЗЗ, включая уже обработанные по определенным алгоритмам снимки. Особенно это важно для гипер- и мультиспектральных, а также радарных данных, так как их тематическая обработка может выполняться только высококвалифицированными специалистами. При этом пользователь освобождается от сложной и затратной по вычислительным ресурсам обработки и получает гарантированный результат. Для такого вида сервиса можно применить подход на основе краудсорсинга, когда пользователь предоставляет свои данные, получая взамен соответствующий вид сервиса (обработку данных, карты, графики, таблицы и т. п.). Используя сервисы, которые реализуются с помощью облачных систем, мы превращаем информационную среду в сервисную, в которой работают слабые связи, с возможностью многократного использования и выявления отдельных компонентов. Иными словами, мы переходим

в среду SOA (Service-Oriented Architecture) [26], которая строится на сервисах, а не на приложениях. SOA основана на принципе слабой связи, где сервис — это дискретные программные компоненты, имеющие определенную функциональность и используемые в составе многих приложений. В этом плане облачные вычисления предоставляют больше возможностей и сервисов. Принципы SOA оказываются достаточно полезными при реализации распределенных ГИС, например, так реализована прикладная геоинформационная система CUSHI HIS [27]. Отметим, что облачные вычисления позволили существенно увеличить возможности распределенных геоинформационных систем и значительно упростить их разработку.

В настоящее время практически все исследователи отмечают взрывной рост информационных потоков, при этом уже существующие технологии, например реляционные СУБД, не всегда справляются с ними и не только из-за их бурного роста. Так, например, в некоторых областях наук о Земле сложно найти и проанализировать результаты проведенных на протяжении многих десятков лет экспериментальных исследований и расчетов, так как они хранятся где угодно и как угодно, поэтому следует признать, что сформировать современные системы на основе хранилищ данных (ETL-системы — Extract, Transform, Load) [28] даже с учетом облачных сервисов вряд ли удастся. Многие организации хранят свою расчетную и экспериментальную информацию в различных, часто собственных форматах, в результате чего либо она просто теряется, либо для ее использования требуется предпринять определенные усилия, связанные с решением проблемы “хаоса данных”. Кроме того, сегодня появились совершенно новые типы данных, такие как облака пространственных точек, получаемые при лазерном сканировании, или данные радарной интерферометрии, которые можно отнести к классу плохо структурированных данных. Поэтому перспективной для создания распределенных систем обработки пространственных данных становится активно развиваемая *технология больших данных* [29]. Ее применение для реализации распределенных ГИС не только позволит решать многие вопросы, связанные с катастрофическим ростом объемов геоданных, но и на некоторой единой методологической основе создавать системы для сбора, хранения и обработки разнородной и неструктурированной геоинформации (например, по отдельным горнопромышленным районам).

Так как объем и скорость поступления геоданных ежегодно увеличиваются на порядки, для хранения произвольных научных данных, структурированных и неструктурированных, требуются совершенно другие модели баз данных, а также новые методы, основанные на Data Science [30], в которых преобладают синтезирующие теории, а статистические методы применяются при работе с огромными объемами и потоками данных.

Рассмотрим для примера современную геодинамику, для которой можно выделить следующие группы больших данных:

- экспериментальные данные, получаемые в режиме реального времени и представляющие собой некоторые временные ряды, связанные с изменением состояния массива горных пород вследствие техногенных нагрузок;
- облака точек, получаемые с помощью приборов лазерного сканирования;
- растровые изображения, отражающие динамику изменения состояния массива, получаемые с помощью средств дистанционного зондирования;
- результаты расчетов различной пространственной размерности, получаемые на основе геомеханических математических моделей и в большинстве случаев представляющие собой набор сеточных данных (grid-модели);

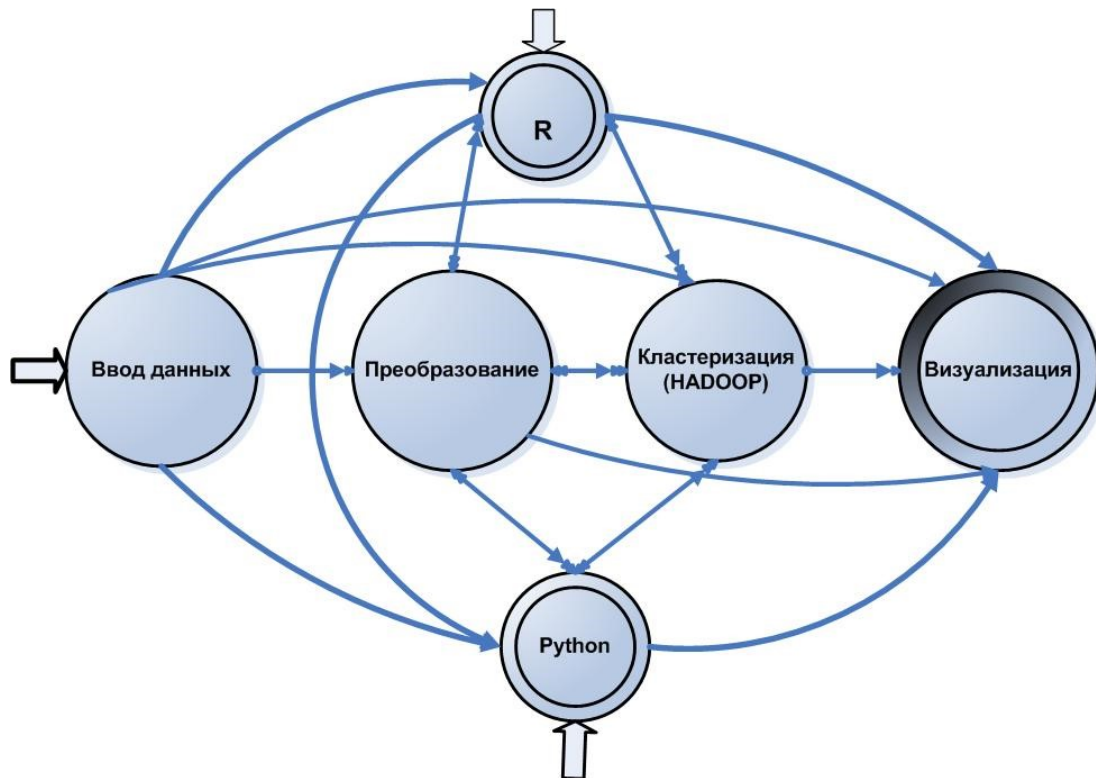


Рис. 3. Схема обработки геоданных на основе Hadoop-технологии

- аудио и видео мультимедийные данные, полученные либо с помощью новейших приборов, либо на основе результатов расчетов;
- архивные данные различной природы и форматов.

Такое многообразие данных ставит не только ряд новых проблем по их обработке и хранению, но и дает право говорить о создании совершенно новых систем, ориентированных именно на большие данные. Существует классическое определение больших данных — “четыре V”: объем (volume), скорость обработки (velocity), достоверность (veracity) и разнообразие (variety) [31].

Развитие вычислительных кластеров позволяет развивать такие системы. Однако встает ряд вопросов, связанных как с доступностью этих систем, так и со специализированной операционной средой, для которой необходимо разрабатывать соответствующие методы распараллеливания уже имеющихся задач и моделей. Так как для большинства организаций это является сложной проблемой, сегодня трудно найти удачные примеры реализации геоинформационных вычислительных комплексов, доступных большинству пользователей.

Для решения этих и других проблем, вызванных появлением распределенных геоинформационно-вычислительных систем, нами предложен подход, основанный на создании вычислительных кластеров, использующих достаточно простые программно-технические комплексы на базе серверов стандартной архитектуры, объединяемых с помощью так называемой Hadoop-технологии [32] в рамках модели программирования MapReduce [33]. Данный подход позволяет повысить не только производительность обработки различной геоинформации, в том числе и неструктурированной, за счет применения параллельных вычислительных систем, но и эффективность разработки новых приложений, связанных с большими данными.

К основным преимуществам Hadoop-технологии [32] относятся:

- автоматическое распараллеливание задач на кластере из серверов стандартной архитектуры;
- распределение нагрузки между узлами кластера;
- защита оборудования от сбоев за счет перезапуска задачи на другом кластере;
- распределенная файловая система для хранения данных на внутренних серверах кластера.

Общая схема обработки потоков геоданных представлена на рис. 3. Здесь блоки Python и R приведены в качестве примера как наиболее популярные при реализации методов обработки геоданных, они могут быть заменены на другие инструментальные средства. В последнее время появилось много реализаций геоинформационных систем с применением Hadoop, например как в работе [34]. Таким образом, использование такого подхода для реализации распределенных ГИС можно считать перспективным.

3. Интеграция ГИС и систем дистанционного зондирования Земли

Интенсивное развитие систем дистанционного зондирования Земли привело к резкому увеличению информационных потоков пространственных данных. Так, разными странами запущено около 300 космических аппаратов, которые поставляют различную информацию о состоянии земной поверхности [35]. Потоки геоданных, формируемых различными системами, условно можно разделить на мульти-, гиперспектральные и радарные, каждый из которых имеет собственные форматы, и, соответственно, с их помощью можно определять те или иные характеристики земной поверхности и динамику их изменения. Развитие получили и программные средства для их обработки. Как показано на рис. 1, средства обработки данных ДЗЗ — одна из составляющих ГИС, но, к сожалению, нет уверенности, что любое последующее развитие этих систем обязательно приведет к их прочному взаимодействию как в теоретическом, так и практическом плане и что ГИС и дистанционное зондирование будут интегрированы в большей степени, чем это имеет место в настоящее время.

С позиции интеграции систем дистанционное зондирование обеспечивает сбор пространственных данных для баз ГИС, это позволяет осуществлять обновление и проверку тематических покрытий, используя различные типы геоданных: мульти- и гиперспектральные, радарные, данные беспилотных летательных аппаратов и лазерного зондирования. При этом данные ГИС являются вспомогательной информацией для обработки изображений. Существует много методов и алгоритмов совместной обработки векторных и растровых данных, которые позволяют получать комбинацию этих слоев и решать различные задачи, такие как определение тематических кластеров, сегментация изображений с привязкой к конкретным областям карты, вычисление некоторых физических характеристик, связанных с локальными и глобальными изменениями топографических поверхностей, а также другие, в том числе комбинированные аналитические функции. Эти методы и алгоритмы включают основные пространственные запросы, оверлей статистических и тематических атрибутов ГИС и дистанционного зондирования, используя булеву и нечеткую логику, создание экспертных систем для многократной обработки. В работе [36] выделены три этапа развития интеграции ГИС и ДЗЗ.

Этап 1. Раздельная, но синхронная разработка баз пространственных данных для каждой технологии. Данные передаются преимущественно в векторном (для ГИС) и растровом (для дистанционного зондирования) формате и могут быть одновременно визуализированы оверлейными программами. Их анализ может быть ограничен обновлением слоев ГИС, сравнением тематических атрибутов, сгенерированных системами дистанционного зондирования, или использованием данных ГИС для того, чтобы обеспечить геопривязку изображений.

Этап 2. Дальнейшее развитие баз пространственных данных, но при этом каждая технология должна позволять использовать некоторую общую схему данных, для которой любые данные могут быть преобразованы в другие через векторизацию, растрезацию и соответствующее изменение пространственных и временных атрибутов.

Этап 3. Заключительный этап полной интеграции. По существу, ГИС и дистанционное зондирование становятся одной неразличимой системой, в которой растровые и векторные данные преобразуются в некоторую метамодель, позволяющую осуществлять весь набор операций, свойственных как ГИС, так и системам обработки ДЗЗ. Полная интеграция сегодня, хотя и теоретически возможна, но на практике ее не существует. Однако проведено большое количество исследований, на основе которых созданы приложения, интегрирующие ГИС и системы дистанционного зондирования.

Иными словами, в современных ГИС появились разделы, связанные с обработкой данных дистанционного зондирования, и в то же время в программные комплексы обработки данных ДЗЗ стали включать модули обработки пространственной информации, традиционно принадлежащие ГИС. Это отчетливо видно при анализе последних версий таких многофункциональных пакетов, как ARCGIS фирмы ESRI, а также ENVI (Exelis) и ERDAS IMAGINE фирмы Intergraph, свободно распространяемых ГИС SAGA, QGIS, GRASS. Сам характер интеграции в этих пакетах различный. Иногда в ГИС дополнительно включаются модули обработки изображений или появляются некоторые специализированные интерфейсы, вызывающие внешние процедуры обработки данных соответствующего формата и типа, как это сделано в ENVI, когда при генерации пакета пользователь просто указывает наличие системы ARCGIS на его машине, и система проводит соответствующие изменения в настройках toolbox, включая в него необходимые модули. Так как в обеих системах имеются интерфейсы с языком Python, интерфейс выполняется его средствами.

Это направление интеграции находится еще в стадии разработки, и не всегда средствами ARCGIS возможно решить задачи, которые с легкостью решаются в ENVI. Намного проще связь систем осуществляется на уровне конверторов форматов различных пространственных данных, которые просто импортируют и экспортируют их из ГИС в пакеты данных ДЗЗ, и наоборот. По этому пути пошли разработчики свободно распространяемых ГИС, в которые включены средства импорта и экспорта данных ДЗЗ, полученных с различных космических аппаратов, и ряд вычислительных модулей для их обработки, например, кластеризации и сегментации растровых данных. Недостатком такого подхода является то, что каждый пакет может обрабатывать не все данные космических аппаратов, а только некоторые (например, систем Landsat или SPOT) в зависимости от уровня их первоначальной обработки [37], что существенно ограничивает его функциональность. Выходом может служить формирование данных в формате GeoTIFF, что позволяет работать с большим множеством систем. Такой подход применяется при работе с отечественными системами дистанционного зондирования “Ресурс-П” и “Канопус-В”.

Наиболее простой путь интеграции между ГИС состоит в том, что данные различных систем могут использоваться как некоторые слои для создания комплексных электронных карт, применяемых впоследствии для как стационарных, так и распределенных ГИС. Именно за счет ДЗЗ сегодня происходит актуализация множества различных электронных атласов во всем мире, применяемых для решения таких практических задач, как:

- мониторинг фактического использования земель, актуализация информации о состоянии территории;
- оценка состояния объектов строительства;
- создание и обновление цифровой картографической основы земельного кадастра;
- мониторинг транспортной сети;
- мониторинг состояния полигонов бытовых отходов;
- выявление несанкционированных свалок;
- картографирование лесного фонда;
- определение породного состава лесов;
- изучение природных условий, анализ биоразнообразия территорий;
- мониторинг инфраструктуры объектов добычи полезных ископаемых и транспортировки;
- мониторинг зоны контроля подземных магистральных трубопроводов;
- мониторинг экологического состояния территорий в районах добычи, переработки, транспортировки нефти и газа;
- выявление территорий, загрязненных нефтепродуктами, мониторинг аварийных разливов нефти, контроль темпов и оценка эффективности рекультивационных мероприятий;
- оперативное определение районов аварий и изучение их транспортной доступности;
- определение смещений земной поверхности на разрабатываемых месторождениях полезных ископаемых.

Среди основных направлений использования ДЗЗ для геологических исследований можно выделить:

- структурный анализ поверхности Земли;
- анализ локальных и глобальных аномалий;
- геологическое изучение площадей и составление геологических карт;
- геоэкологический мониторинг экзогенных геологических процессов;
- выявление потенциальных месторождений полезных ископаемых;
- определение линейных структур (разломов, контрастных геологических границ и т. д.);
- уточнение границ геологических тел, выходящих на дневную поверхность;
- тектоническое районирование;
- кинематическая классификация разрывных нарушений и связанных с ними локальных структур;
- оценка угроз тектонического и сейсмологического характера на основе оценки подвижек разломов и их динамики (с высокой точностью на основе данных радиолокационной интерферометрии);
- оценка горизонтальных и вертикальных смещений вдоль разломов;
- геологическое и геоморфологическое картографирование.

Т а б л и ц а 4. Интегрированные ГИС и системы обработки данных ДЗЗ

Название продукта	Производитель	Основной тип	Примечание
ESRI ARCGIS v. 10.3	ESRI Inc. (США) http://www.esri.com	Многофункциональная ГИС	На уровне toolbox достаточно развитые средства обработки изображений, в том числе обеспечивающие работу с некоторыми типами данных ДЗЗ
Erdas Imagine v. 15.0	Hexagon Geospatial (США) www.geospatial.com	Обработка данных ДЗЗ. Один из самых мощных пакетов	Встроенный пакет для создания виртуальных ГИС и работы с 3D-моделями
ENVI v. 5.3	Exelis VIS (США) www.exelisvis.com	На сегодняшний день самый многофункциональный пакет для обработки данных ДЗЗ	Отдельные элементы для пространственного анализа базовыми средствами ГИС. Интеграция с ESRI ARCGIS
Geomatica 2015	PCI Geomatics (Канада) www.pcigeomatics.com	Обработка данных ДЗЗ	Набор модулей для пространственного анализа, в основном 2D
Idrisi Selva 17.02, TerrSet 18.07/ April 2015	Университет Кларка (США) www.clarkslab.org	Система для пространственного мониторинга и моделирования	Полный ГИС-пакет, для базового и расширенного пространственного анализа, наличие программ для решения экологических задач
SAGA v. 2.1.4	Гамбургский университет (Германия) http://www.saga-gis.org	Расширяемая, свободно распространяемая ГИС	Набор средств для работы с изображениями, в основном промежуточных форматов, включая морфологический анализ, классификацию, сегментацию, фотограмметрию
Quantum ГИС QGIS 2.10 (Pisa)	http://www.qgis.org Развивается благодаря усилиям волонтеров из различных стран мира	Свободно распространяемая ГИС	Постоянно расширяемый набор модулей для работы с изображениями, интегрирована с некоторыми системами обработки изображений
GRASS 7.1	http://gras.baylor.edu , http://gras.itc.it	Одна из первых свободно распространяемых ГИС	Интегрирована с другими ГИС, имеет модули обработки изображений, в том числе спутниковых

В табл. 4 приведены наиболее распространенные программные комплексы ГИС и обработки ДЗЗ, в которых реализована интеграция того или иного вида.

4. Интеллектуальные ГИС

В последние годы все большую популярность начинают приобретать интеллектуальные ГИС (ИГИС), для которых пока нет устоявшейся терминологии. Некоторые исследователи под интеллектуальными ГИС понимают системы обработки пространственной

информации, в которых используются различные методы искусственного интеллекта. Другие под интеллектуальной ГИС понимают программно реализованную систему искусственного интеллекта, которая может наряду с другими использовать в том числе и пространственные данные, динамически генерируя соответствующие человеко-машинные интерфейсы. Здесь основное внимание уделяется созданию, моделированию и оценке эффективности качественных аспектов географической информации, а также воздействию обработки пространственной информации на вычислительные модели тех или иных процессов. Интеллектуальные ГИС такого типа обычно осуществляют динамическую настройку на предметную область, используя системы логического вывода и интерпретацию сценариев обработки, с применением баз онтологий, настраиваемых объектных моделей метаданных, продукционных моделей, методов извлечения знаний, интеллектуального анализа данных, построения разнообразных хранилищ и др. Одна из первых работ в этом направлении — [38]. Сегодня ИГИС начинают находить все более широкое применение в областях, связанных с анализом транспортных сетей, логистики и принятия решений в динамическом режиме генерации возможных пространственных вариантов с соответствующей маршрутизацией [39, 40].

Для большего понимания структуры и принципов построения ИГИС необходимо рассмотреть их отдельные элементы и технологии их составляющих, учитывая оба подхода и основные схемы построения классических ГИС. Иногда для построения ИГИС применяется концепция сервисно-ориентированной архитектуры (SOA), в рамках которой используется декомпозиция системы на отдельные функциональные блоки, где каждый из них связан с определенным видом сервиса, при этом один из блоков построен на принципах систем искусственного интеллекта, он фактически выполняет функции экспертной системы, которая осуществляет настройку приложений из других блоков на конкретную задачу. Применение такого подхода для определенных приложений можно считать целесообразным. Однако современные ИГИС имеют сложную структуру и чаще всего строятся как распределенные, включая в себя различные технологии [41]. Рассмотрим основные технологии, применяемые для создания блоков ИГИС.

Data Mining — технология интеллектуального анализа данных, для которой существует множество определений [42] с разной степенью подробности, но применительно к ИГИС целесообразно использовать следующее. *Data Mining* — это процесс извлечения полезной информации из больших наборов данных, при этом информация может быть дескриптивной (акцент делается на понимание структуры данных или их отдельных образцов) и предиктивной (выполняется предсказание одной или более переменных на основе других) [43]. Такое определение удобно для понимания методов интеллектуального анализа и четко разделяет его на две большие области, которые используют различные модели для обработки пространственной информации.

Методы *Data Mining* сегодня используются в самых различных приложениях: от медицины до геологии. В зависимости от вида информации в системе выделяют подсистемы, которые выполняют следующие функции:

- анализ текстовой информации;
- анализ и извлечение информации из веб-сайтов;
- анализ изображений;
- манипуляции пространственными объектами (векторными или растровыми);
- классификацию пространственных данных;
- кластеризацию пространственных наборов данных;
- построение решающих правил.

Последние три подсистемы наиболее часто используются при обработке данных ДЗЗ, поэтому они включены во многие современные программные комплексы.

Хранилища данных (data warehouse) — субъектно-ориентированная многовариантная по времени интегрированная коллекция доступных данных для принятия решений [44] (рис. 4). Хранилища формируются путем интеграции информации из оперативных баз данных с помощью преобразований, очистки, извлечения зависящих от времени наборов данных. Поступившие в хранилище данные уже невозможно редактировать, поэтому иногда для некоторого их подмножества применяется термин “информационные гранулы”. Проблемно-ориентированные множества хранилищ иногда организуются в “киоски данных” (data marts), с которыми пользователь может взаимодействовать. Хранилища данных особенно удобны для накопления больших наборов плохо структурированной информации. Сегодня они находят широкое применение, особенно в областях, связанных с мультимедийной информацией и обработкой изображений. Для ИГИС хранилища данных являются обязательным элементом.

Экспертные системы (машины логического вывода). Из всего множества различных элементов экспертных систем в ИГИС чаще всего используются машины логического вывода, программно реализующие механизмы рассуждений и оперирующие данными с целью получения новых знаний об интересующих объектах. Как правило, используются механизмы дедуктивного логического вывода, поиска решений в сети фреймов или семантические сети, продукционные модели. Такие системы имеют узкую направленность и обычно разрабатываются под конкретную задачу [45] (например, транспортную логистику, проблемы энергетики), поэтому пока не находят широкого применения.

Мультиагентные системы — это новый класс систем, сегодня только начинающий применяться в конкретных реализациях. Рассмотрим некоторые базовые определения.

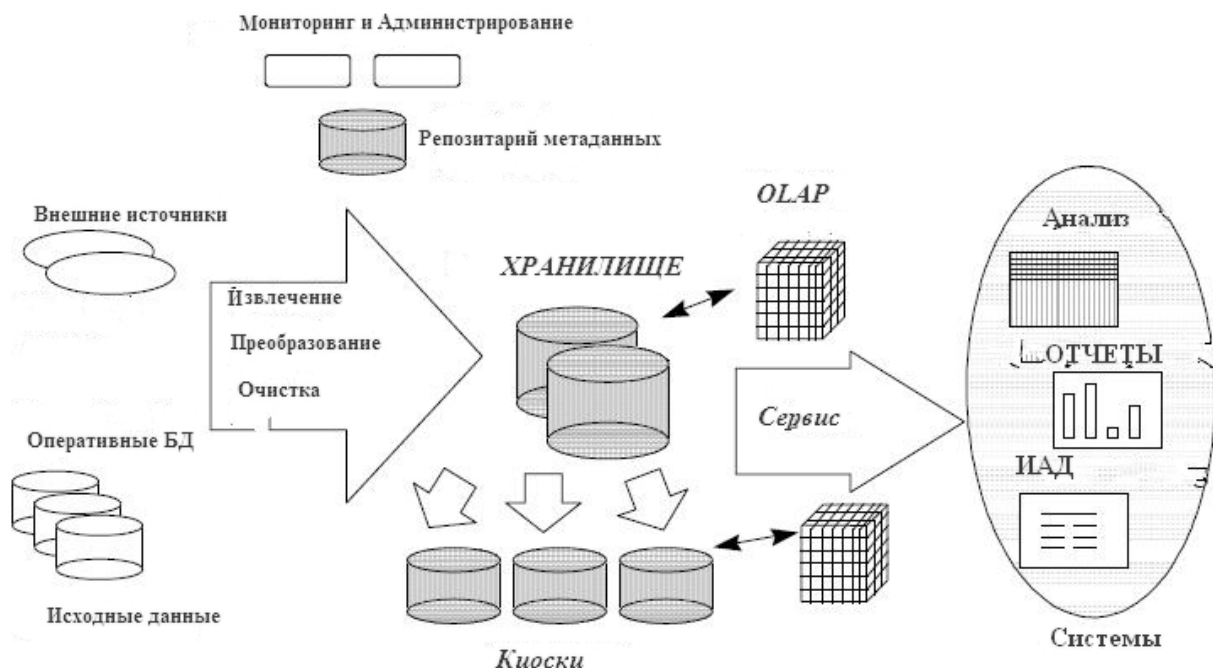


Рис. 4. Схема обработки геоданных на основе Hadoop

Интеллектуальный агент — программа, самостоятельно работающая по заданию пользователя в течение некоторого длительного промежутка времени. Иногда под интеллектуальными агентами понимают сущности, получающие разнообразную информацию о состоянии управляемых ими процессов и осуществляющие необходимые действия для достижения конкретных параметров. Отличительной особенностью мультиагентных систем является переход от локального типа к распределенному. Агенты как программные системы обладают такими свойствами, как: автономность, мобильность, реактивность, взаимодействие, индивидуальность, корпоративность, интеллектуальность поведения.

С точки зрения программной реализации агенты можно рассматривать как комплекс функций и соответствующих интерфейсов, которые могут получать вопросы и посылать ответы на них. Существует несколько типов агентов: агенты с простым поведением, с поведением, основанным на некоторой модели, целенаправленные, практические и обучающие [46].

Мультиагентные системы можно применять для решения самых разнообразных задач по обработке и анализу пространственной информации, например:

- интерактивная работа с хранилищами данных;
- создание гибридных моделей для обработки пространственных данных;
- анализ геоэкологической информации;
- разработка кадастров для планирования и управления территориями;
- прогноз развития городов;
- анализ демографических процессов;
- логистика транспортных систем;
- моделирование социальных систем;
- создание экономических моделей регионов;
- моделирование распространения эпидемий;
- планирование развития промышленных агломераций;
- анализ последствий катастроф;
- исследование динамики геосистем и др.

Уже простое перечисление задач, решаемых с помощью мультиагентных систем, указывает на их универсальность и широкое использование для различных приложений, связанных с пространственной информацией. Это объясняется тем, что они способны осуществлять дезинтеграцию сложных систем на индивидуальные компоненты, каждый из которых обладает собственным поведением с эволюцией в пространстве и во времени. Наиболее известны из моделей таких систем клеточные автоматы, которые уже сравнительно давно используются в различных приложениях [47]. Достоинством этого подхода является и тот факт, что имеется значительное количество инструментальных программных средств (Geomason, Repast), позволяющих отвлечься от математического аспекта теории мультиагентных систем и использовать их для решения собственных задач [48]. Несомненно, такие системы имеют широкие перспективы, особенно в области развития ИГИС.

Онтологии и базы знаний. Не касаясь чисто философского понятия “онтология”, которая занимается анализом категорий, определяющих сущность реального мира, и их особенностей, в computer science онтологию можно рассматривать как технологию создания моделей некоторой предметной области на основе артефактов, организуя и классифицируя их для использования в программах. В более упрощенном понимании онтология занимается созданием концептуальных моделей, основываясь на семантике предметной области. Онтологии часто используются для формирования баз знаний в каче-

стве “каркаса” для представления их концептуальной модели в какой-либо предметной области. При этом другая часть базы знаний является хранилищем экземпляров абстракций реальных объектов. Для общей иллюстрации этого элемента ИГИС приведем схему из [49] (рис. 5).

В области ГИС техника онтологий чаще всего применяется для создания настраиваемых объектных моделей с использованием метаданных, а также определения основных понятий и взаимоотношений между ними, для описания геопространственных знаний, сценариев обработки и алгоритмов моделирования. При этом метаданные рассматриваются как знания о данных, которые интерпретируют информацию для конкретной области решения задачи по обработке пространственных данных. Иногда онтологии используются как первый этап гармонизации пространственных данных, определяемой как приведение семантических характеристик метрической системы в соответствие с другими пространственными данными для обеспечения их взаимозаменяемости [50]. Применение онтологий позволяет использовать накопленный в предметной области кон-

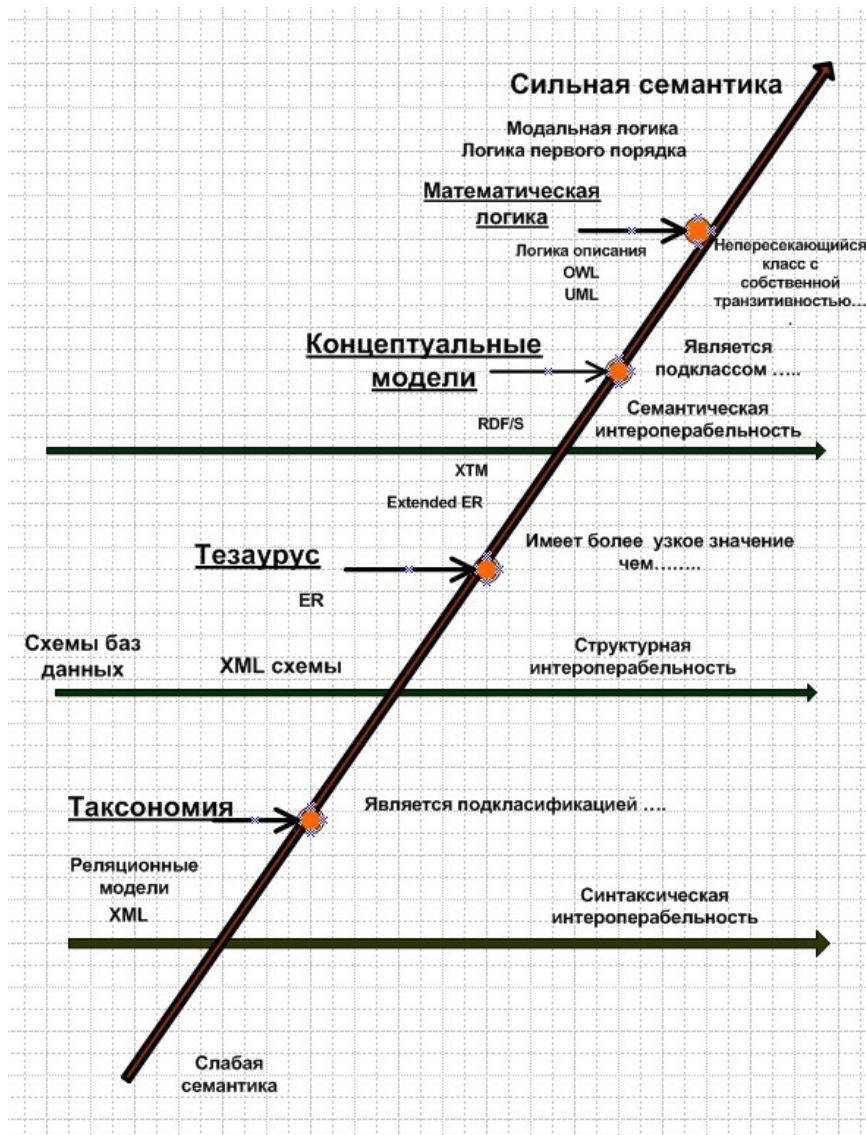


Рис. 5. Спектр различных онтологий

кредный опыт и изменять библиотеку соответствующих моделей, минимизируя тем самым время разработки приложений. Для создания онтологий можно использовать различные инструментальные средства [48], в том числе библиотеки на Python, Java, C# и ряд других.

Методы извлечения знаний (*Knowledge Discovery in Database — KDD*) также используются для создания ИГИС. В отличие от data mining эти методы являются более общими и занимаются извлечением полезных знаний из данных, в то время как методы интеллектуального анализа для извлечения образов из данных используют специфические алгоритмы. Эти технологии частично пересекаются с распознаванием образов и машинным обучением. Они фокусируют свое внимание на нахождении понятных образов, которые могут быть интерпретированы как потенциально полезные и интересные знания, содержащиеся в данных, имеют много общего со статистикой, иногда используют ее элементы на отдельных этапах извлечения знаний.

Традиционные методы извлечения знаний обычно опираются на опыт экспертов, работающих в конкретной прикладной области, но современное взрывное увеличение объемов анализируемых данных давно не позволяет их использовать, особенно с учетом возросшего количества различных типов данных, включая, например, облака точек или мультимедийную информацию. Извлечение знаний является более обширной технологией, включающей в себя методы интеллектуального анализа, которые работают на отдельных стадиях всего процесса. Именно эта технология вобрала в себя все лучшее из методов искусственного интеллекта, статистики, методов визуализации, мультиагентных систем и др. Общая схема извлечения знаний показана на рис. 6.

Из данной схемы видна разница между методами интеллектуального анализа и извлечения знаний на каждом этапе обработки информации. В настоящее время методы извлечения знаний входят в более общую технологию больших данных (BIG DATA), которая является своего рода “метатехнологией”, использующей вышеописанные подходы.



Рис. 6. Схема извлечения знаний

Эвристические методы находят закономерности, свойственные выборкам данных, которые могут быть компактно выражены в понятной человеку форме. Обычно используют некоторые последовательности или комбинации известных методов, не ограничиваясь рамками априорных предположений о структуре выборки и виде распределения значений анализируемых показателей или другими формализованными моделями. При этом отображение структуры анализируемых данных и получаемых решений существенно влияет на адекватность решений для различных задач, особенно в мониторинге и контроле состояния динамически изменяющихся объектов. Последние достижения в этой области позволяют отображать пространственные данные различной размерности и больших объемов, что существенно ускоряет процессы их анализа.

5. Примеры практической реализации

На протяжении многих лет в Институте вычислительных технологий СО РАН проводились исследования по разработке ГИС различной направленности. За это время отработаны различные технологии их создания, начиная от desktop-вариантов и заканчивая распределенными системами с интеллектуальным наполнением. Рассмотрим примеры реализации решения двух больших классов задач:

- разработка геоинформационных систем для мониторинга экологической безопасности в крупных горнопромышленных регионах;
- геоинформационная система оценки сейсмических явлений и их последствий.

Система мониторинга экологической безопасности создавалась как распределенная информационно-вычислительная среда, реализуемая как геопортал, использующий элементы облачного сервиса. Общая схема системы представлена на рис. 7.

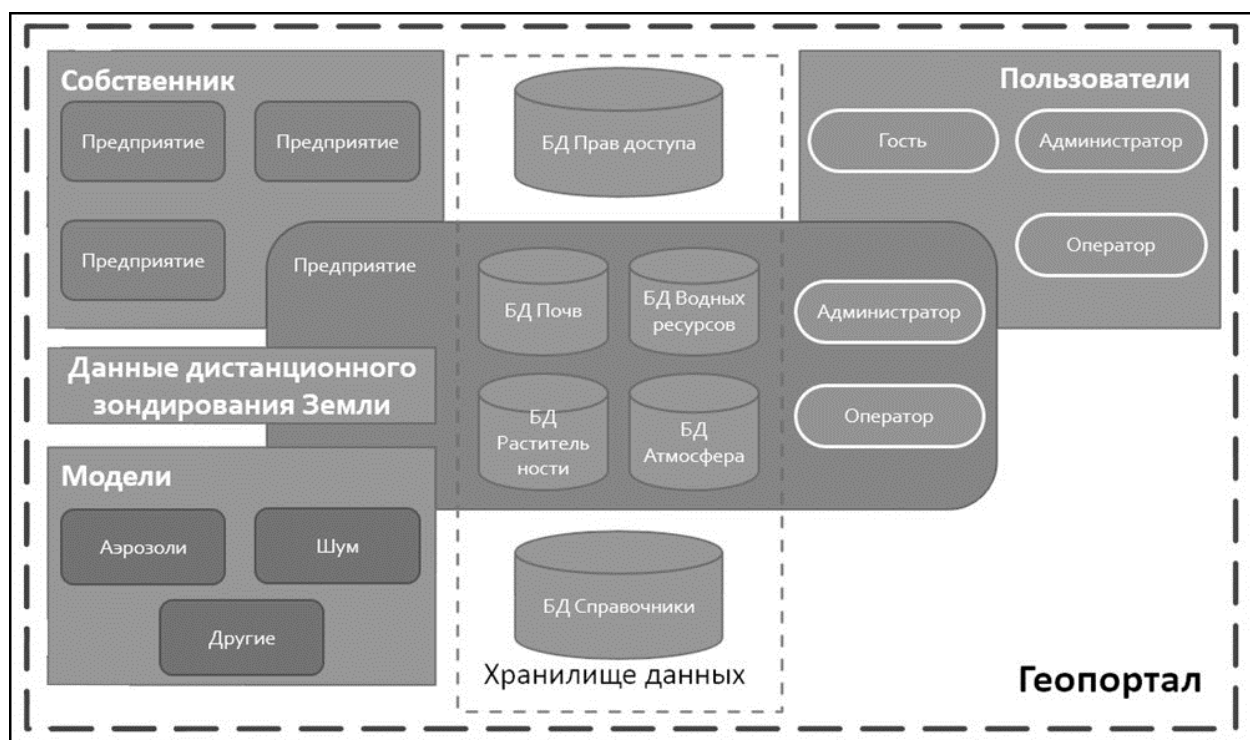


Рис. 7. Схема геопортала интегрированной информационно-вычислительной системы для динамической оценки экологического состояния угледобывающего района

Разработанная система, формируемая как геопортал, состоит из подсистем:

- 1) облачный сервис (Google App Engine);
- 2) сервис аутентификации (Google Users API);
- 3) картографический сервис (Google Map API, веб-сервер Apache);
- 4) сервис баз данных (MySQL, PostgreSQL);
- 5) расчетные сервисы (Java-сервлеты).

Общая архитектура системы показана на рис. 8.

Облачный сервис представляет собой особую клиент-серверную технологию — использование клиентом ресурсов (процессорное время, оперативная память, дисковое пространство, сетевые каналы, специализированные контроллеры, программное обеспечение и т. д.) группы серверов в сети, взаимодействующих таким образом, что для клиента вся группа выглядит как единый виртуальный сервер (рис. 9). Клиент может прозрачно для себя и с высокой гибкостью менять объемы потребляемых ресурсов в случае изменения своих потребностей (увеличивать/уменьшать мощность сервера с соответствующим изменением нагрузки на него).

Наличие нескольких источников используемых ресурсов, с одной стороны, позволяет повышать доступность системы клиент-сервер за счет возможности масштабирования при повышении нагрузки (увеличение количества используемых источников данного ресурса пропорционально увеличению потребности в нем и/или перенос работающего виртуального сервера на более мощный источник, “живая миграция”), а с другой —

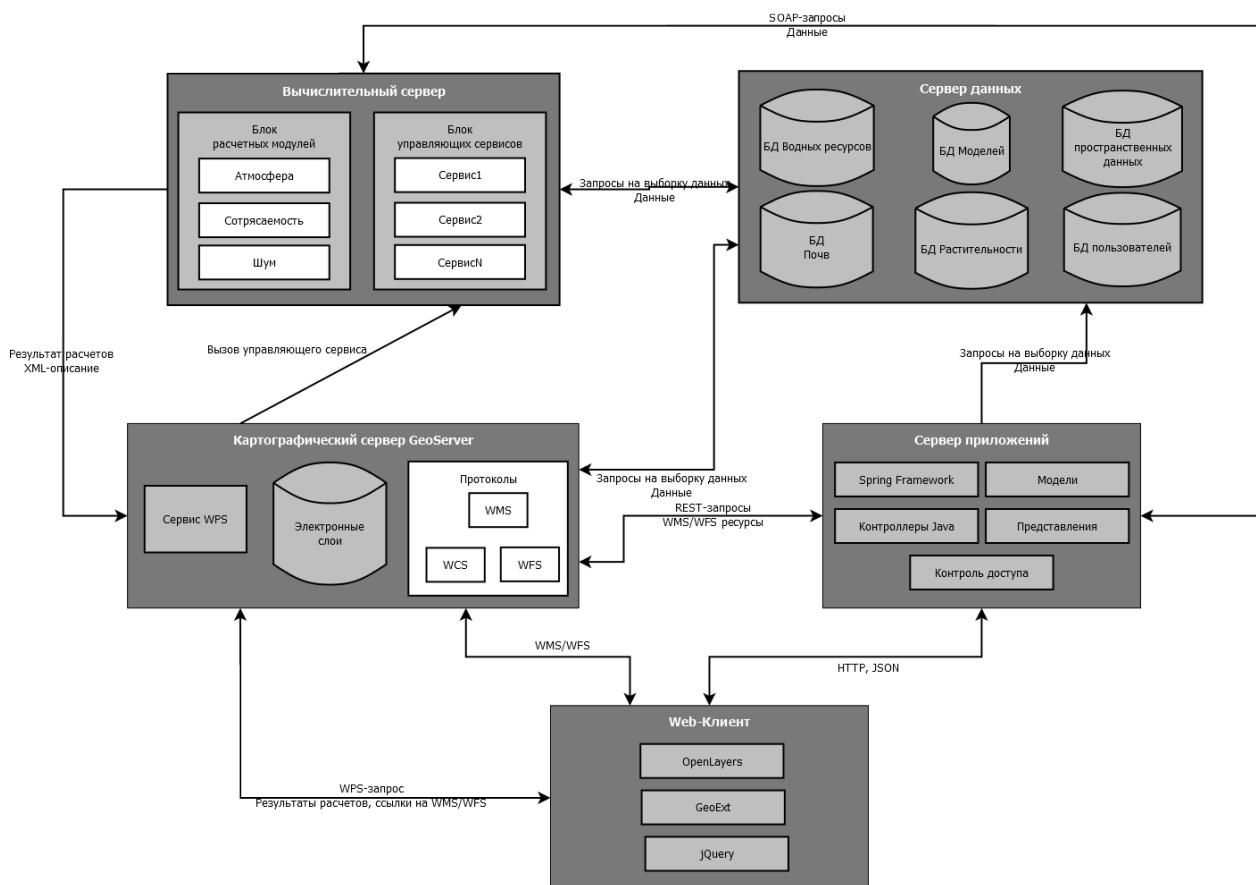


Рис. 8. Архитектура интегрированной информационно-вычислительной системы динамической оценки и прогноза геоэкологического состояния угледобывающего района

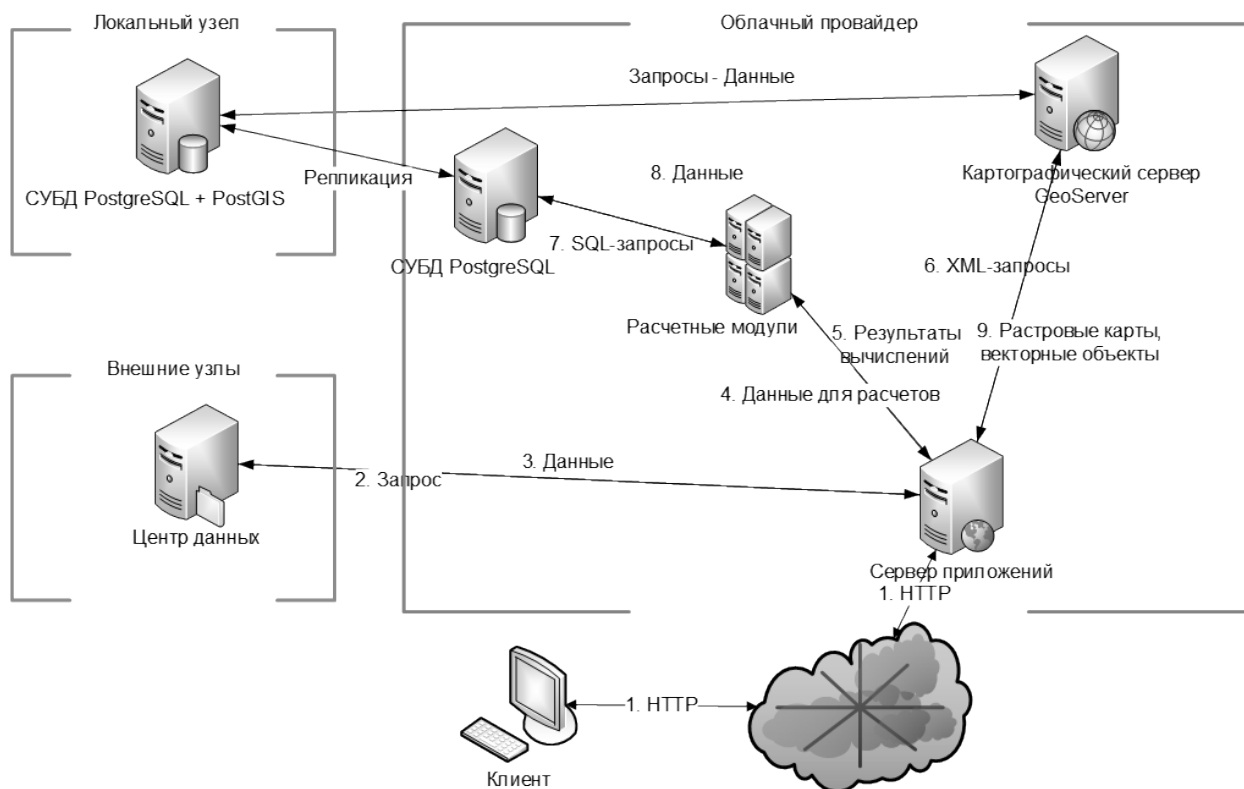


Рис. 9. Общая схема работы системы

снижает риск потери работоспособности виртуального сервера в случае выхода из строя какого-либо из серверов, входящих в группу, обслуживающую данного клиента. Вместо вышедшего из строя сервера возможно автоматическое переключении виртуального сервера к ресурсам другого (резервного) сервера.

Технологическим ядром системы служит среда Google App Engine, позволяющая создавать веб-приложения с помощью стандартных технологий Java и выполнять их на масштабируемой инфраструктуре Google. Среда Java представляет виртуальную машину Java 6 (JVM), интерфейс Java-сервлетов и поддержку стандартных интерфейсов для масштабируемого хранилища данных и служб App Engine, таких как JDO, JPA, JavaMail и JCache. Поддержка стандартов упрощает разработку приложения и делает понятным перенос приложения как из собственной среды сервлетов, так и в нее.

Свободно распространяемое программное обеспечение добавляет в IDE для проектов App Engine новые мастера проектов и конфигурации отладки. Благодаря GWT App Engine для Java существенно упрощает разработку и развертывание веб-приложений. Плагин Eclipse связан с App Engine и SDK GWT. App Engine выполняет приложения Java с помощью виртуальной машины Java 6 (JVM). SDK App Engine поддерживает Java 5 или более позднюю версию. JVM Java 6 также может использовать классы, скомпилированные с помощью предыдущих версий Java. App Engine используется для веб-приложений, которые являются связующим звеном для всех компонентов системы. В стандартной структуре каталогов WAR предоставляет классы сервлетов приложения, страницы JavaServer (JSP), статические файлы и файлы данных вместе с дескриптором развертывания (файл web.xml) и другими файлами конфигурации. App Engine выполняет запросы, вызывая сервлеты в соответствии с дескриптором развертывания.

JVM работает в защищенной среде, что позволяет изолировать приложение в целях обслуживания и безопасности. Среда позволяет удостовериться, что приложения могут выполнять только действия, не влияющие на производительность и масштабируемость других приложений.

Сервисом аутентификации предоставляется безопасный доступ к конфиденциальной информации, который обеспечивается за счет средств Google Users API. Платформа Google App Engine имеет тесную интеграцию с аккаунтами Google, которые могут позволить пользователям авторизоваться. С помощью аккаунтов Google пользователи смогут быстрее приступить к работе с приложением, не регистрируясь дополнительно на сайте, а разработчик получает возможность персонализировать приложение. Приложение может перенаправить пользователя на страницу входа и регистрации или позволить ему выйти из аккаунта. После того как пользователь проходит авторизацию, приложение может получить доступ к его адресу электронной почты и указанному псевдониму. Также можно определить, является ли пользователь администратором приложения и имеет ли он доступ в ограниченные разделы сайта, что позволит создать систему управления сайтом.

Картографический сервис предоставляет возможность визуализировать предметную информацию на картах-подложках, включая Google. Сервис программно разделен на две части: публичную и закрытую. В первом случае используется технология Google Maps API, предоставляющая открытый интерфейс функций по работе с картами Google. Данный функционал используется для картографических расчетов и координатной привязки закрытой информации как из данного сервиса, так и из сервисов баз данных и расчетных сервисов. Технология позволяет связать карты Google с конкретным веб-сайтом посредством уникального API-ключа. Объекты прикладного интерфейса содержат методы для подгрузки сторонних карт в формате KML/KMZ. Хостинг KML/KMZ-ресурсов расположен на веб-сервере Apache.

В системе разработаны и представлены тематические карты мониторинга поверхностных и подземных вод, снегового покрова, типов ландшафтов и почв, нарушенных земель, а также редких и исчезающих видов растительности. Примеры реализации тематических карт приведены на рис. 10 и 11.

Сервис баз данных полностью базируется на компонентной модели используемых СУБД. Предлагается два открытых продукта: MySQL для хранения и доступа к текстовой информации, в том числе и расчетной, и PostgreSQL с поддержкой объектной модели PostGIS для хранения картографической информации. В системе разработаны и представлены базы данных почвенного и растительного покрова, водных ресурсов, объектов выбросов загрязняющих веществ в атмосферу.

Расчетные сервисы содержат программную логику математических моделей предметной области, реализованную на базе технологий RPC, разрабатываются в виде Java-сервлетов, что позволяет интегрировать их с любым из вышеописанных сервисов системы. Схема базы данных модельного комплекса представлена на рис. 12.

В данной системе реализованы модели оценки выбросов и распространения загрязняющих веществ в атмосферу; выпадения промышленных аэрозолей на подстилающую поверхность; распространения загрязняющих веществ и их выпадения на поверхность при промышленных взрывах с учетом реальной розы ветров; оценки качества поверхностных и подземных вод, сотрясаемости и шумового воздействия при массовых взрывах. Схема управления вычислительными модулями представлена рис. 13, а на рис. 14 приведен пример расчета по одной из моделей.

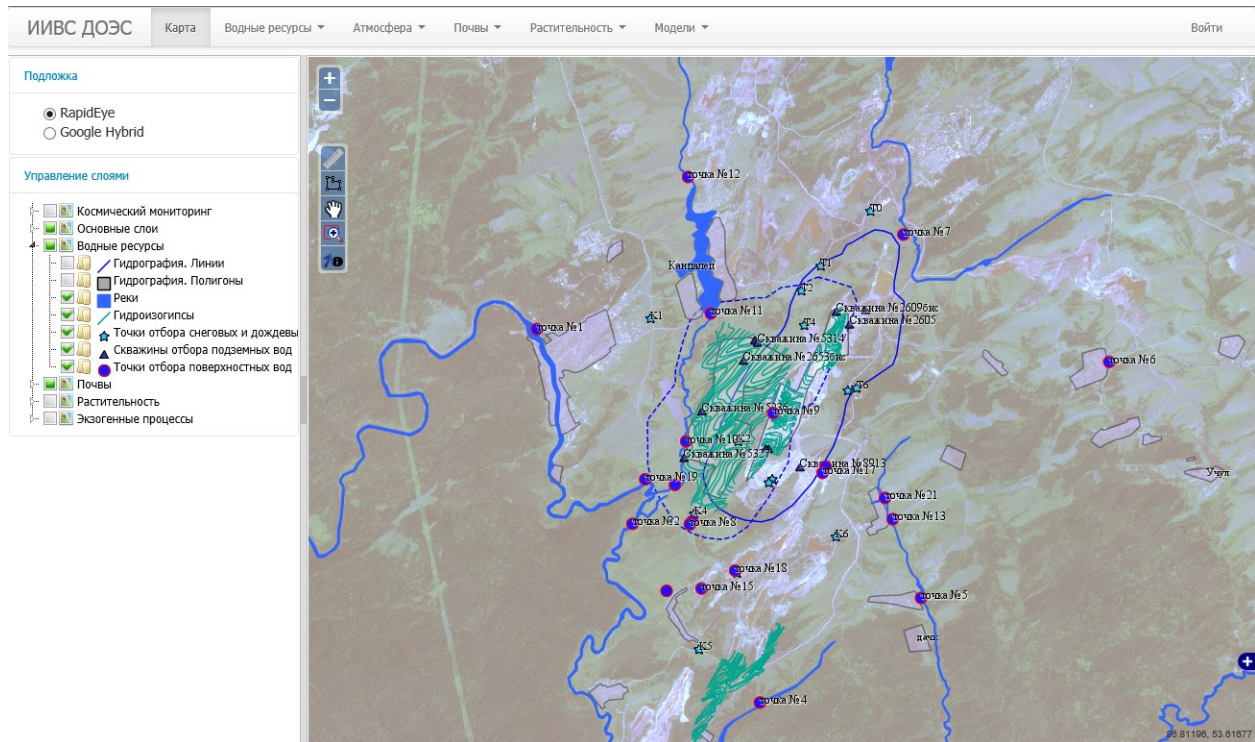


Рис. 10. Мониторинг снегового покрова, поверхностных и подземных вод

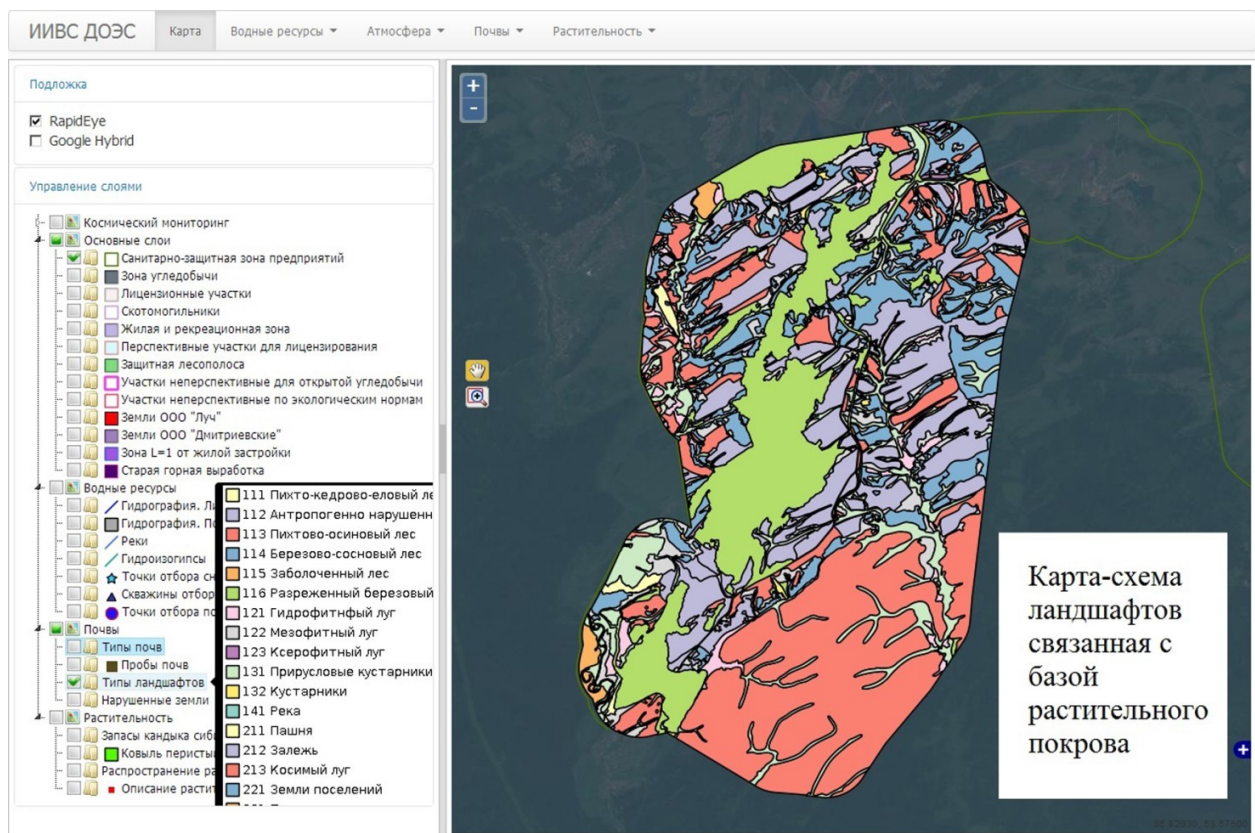


Рис. 11. Карта-схема ландшафтов, связанная с базой растительного покрова

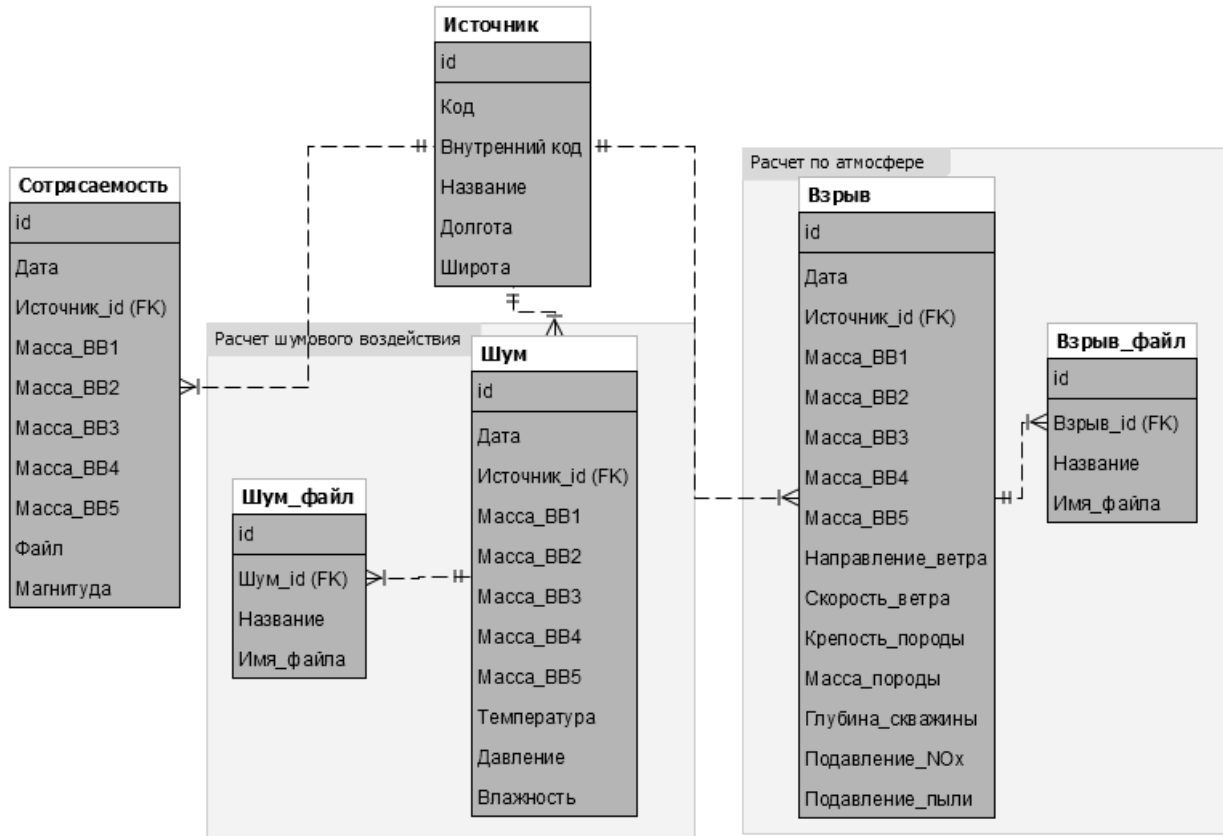


Рис. 12. Схема базы модельного комплекса

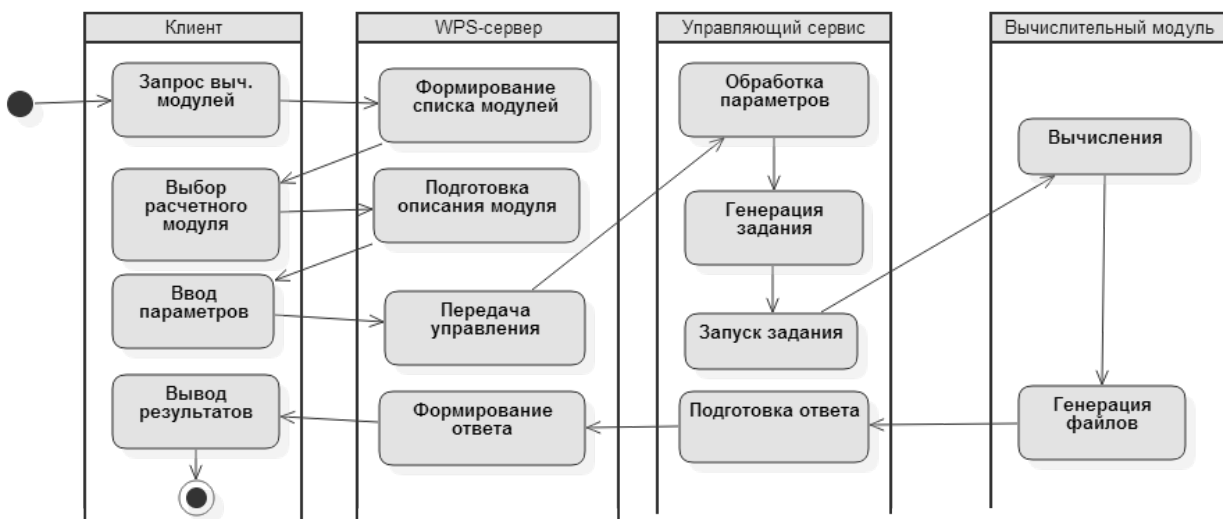


Рис. 13. Схема управления вычислительными модулями

Создан прототип интегрированной информационно-вычислительной системы динамической оценки экологического состояния для большого горнопромышленного региона (на примере Кузбасса). Система обеспечивает сбор и хранение данных по мониторингу природных ресурсов, оценку и прогноз геоэкологического состояния территории, прилегающей к горному (земельному) отводу угледобывающего (углеперерабатывающего)

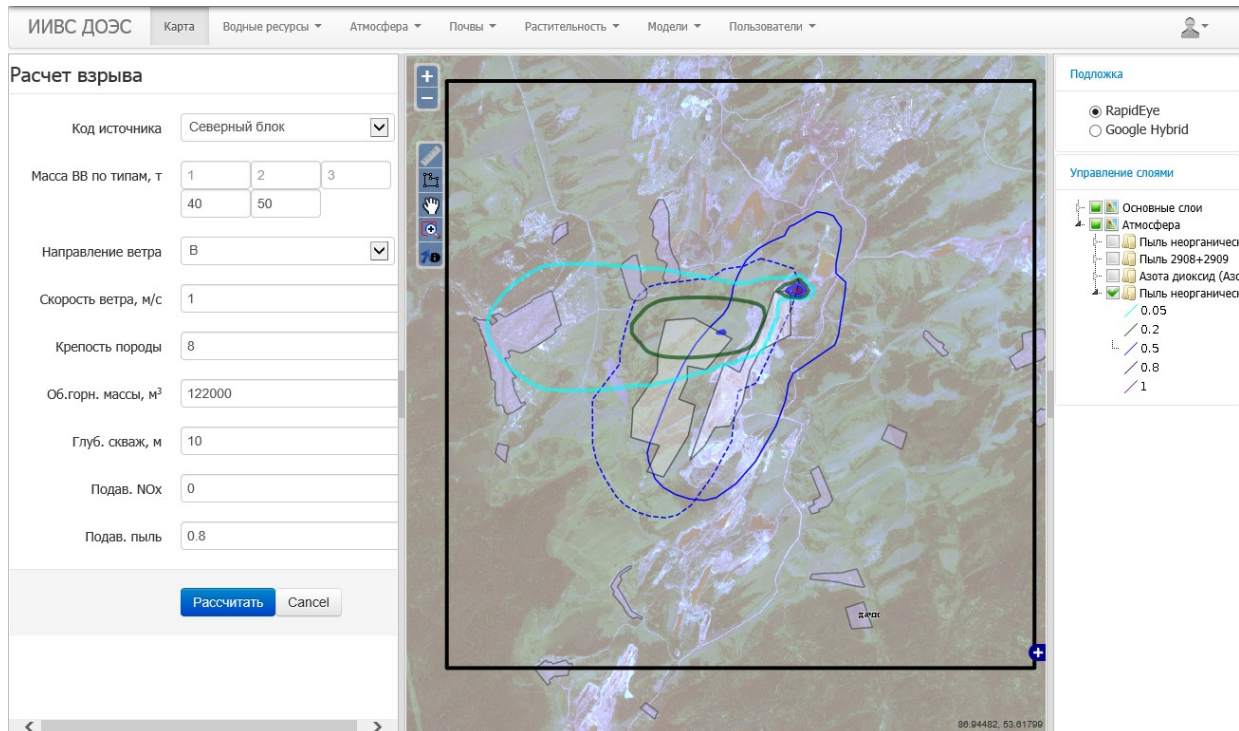


Рис. 14. Пример расчета распространения и оседания пыли при массовом взрыве

предприятия, и угледобывающего района, на территории которого могут размещаться предприятия различного профиля. Она интегрирует данные как наземного, так и дистанционного мониторинга и обеспечивает их информационную поддержку. Разработанная ГИС может быть использована как стандарт информационного обеспечения мониторинга экологического состояния окружающей среды в угледобывающем (углеперерабатывающем) районе.

Сейсмические приложения. По аналогичной схеме разработан веб-сервис как элемент комплексной интеллектуальной ГИС сейсмодетекционного мониторинга на основе облачной технологии Google App Engine в интеграции с сервисами IRIS Data Management Center. Веб-сервис предназначен для классификации сейсмических явлений и реализует алгоритмы сегментации и сравнительного анализа нестационарных процессов на основе эвристического блока энтропийного анализа. Предложен оригинальный подход к многокомпонентной модели ГИС, в которой изменения интерфейса, расчеты модели и визуализация результатов происходят в едином контексте, позволяя инициировать различные варианты начальных условий, получая стек результатов в одной сессии. Технология асинхронных вызовов удаленных сервлетов, содержащих программную логику расчета математической модели с передачей матрицы значений таймсерий, позволила выполнять http-запросы к IRIS WS без перезагрузки графического интерфейса. Данный подход снял проблему сохранения состояния переменных на стороне как сервера, так и клиента.

Использованный в рамках проекта алгоритм расчета траекторий миграции центров сейсмоэнерговыведения позволил перейти от анализа единичных событий к оценке сейсмической обстановки в регионе в целом, что в свою очередь позволило осуществлять сравнение различных территорий. Энтропийный подход дает возможность уверенно классифицировать сейсмические события различной физической природы — от зем-

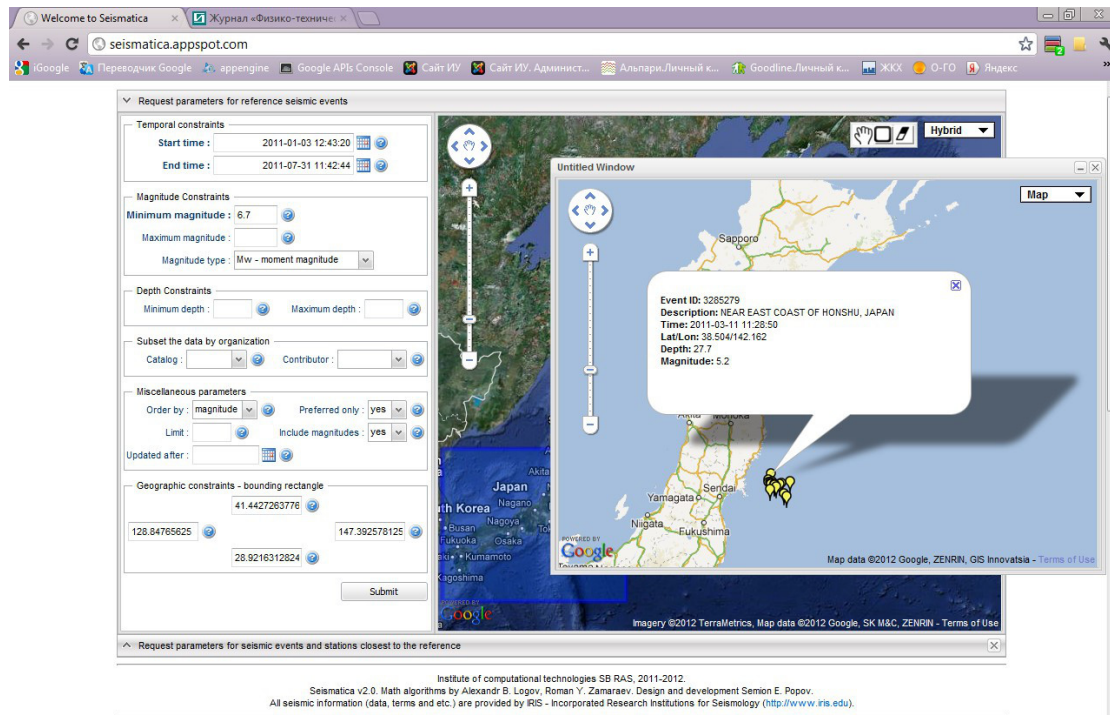


Рис. 15. Пример работы блока для классификации высокоэнергетических событий на базе IRIS ДМС с использованием облачных технологий

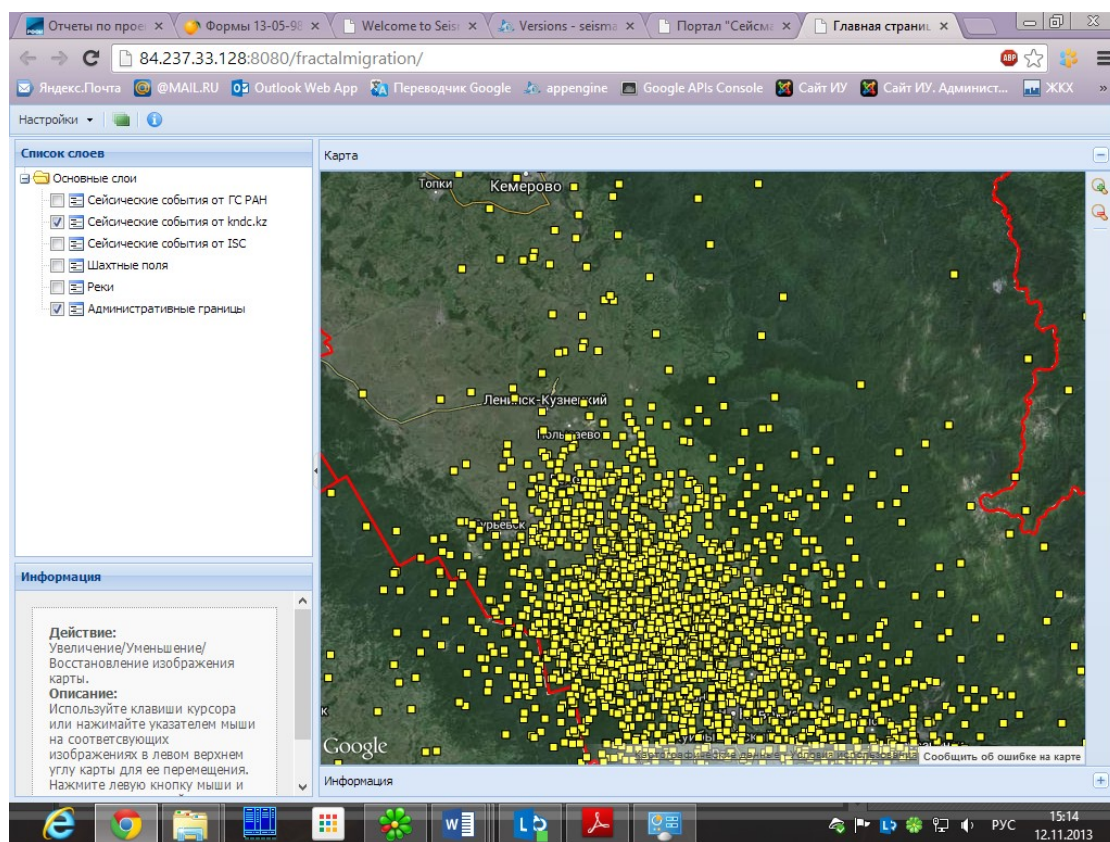


Рис. 16. Сервис расчета фрактальных размерностей траекторий миграции центров сейсмоэнерговыведения для Кузбасса

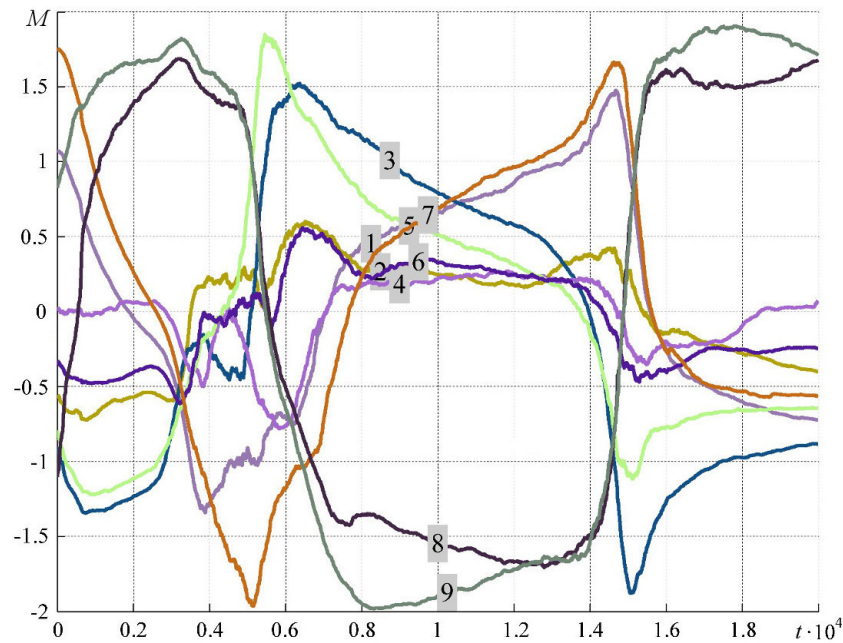


Рис. 17. Работа эвристического блока анализа различных сейсмических событий (землетрясений и промышленных взрывов)

летрясений до промышленных взрывов и внезапных выбросов. Применение облачных технологий повысило надежность сервиса и скорость расчетов в несколько раз. Работа облачных сервисов показана на рис. 15–17.

Приведенный на рис. 17 пример показывает, что разработанные характеристические функции сейсмических событий, представленные во временной области, позволяют оценивать их подобие в энтропийной метрике по фазам накопления информации. При анализе геодинамической ситуации на территории Кемеровской области показана возможность разделения зарегистрированных сейсмических событий на различных горнодобывающих предприятиях на природные и техногенные.

Разработанная геоинформационная система сейсмодеформационного мониторинга, включающая элементы геопорталов и интеллектуальных ГИС, является расширяемой, к ней могут быть подключены различные базы данных (в том числе и базы знаний) и другие расчетные модели в силу универсальности принятых проектных решений. В настоящее время ведутся работы по созданию комплекса программ расчета смещений поверхности на основе данных радарной интерферометрии.

Заключение

Проведенный анализ геоинформационных систем позволил выявить их основные особенности и направления будущего развития. В первую очередь оно будет связано с широким распространением геопорталов, облачных сервисов и развитием систем сервисно-ориентированной архитектуры, которые позволят создавать распределенные ГИС различной направленности. Интеграция геоинформационных систем с быстро развивающимися системами дистанционного зондирования Земли резко увеличит возможности современных ГИС, позволяя в режиме реального времени актуализировать пространственную информацию, особенно в области принятия важных решений. Несомненно,

одним из перспективных направлений создания ГИС будет использование искусственного интеллекта, особенно мультиагентных систем, применение которых в ГИС еще только начинается, но имеющиеся решения уже говорят о том, что здесь можно получить существенно важные новые, как научные, так и практические, результаты.

Список литературы / References

- [1] **Hanks, R.R.** Encyclopedia of geography terms, themes, and concepts. Santa Barbara, California: ABC-CLIO, 2011. 405 p.
- [2] **Star, J., Estes, J.** Geographic information systems: Englewood Cliffs. An introduction. New Jersey: Prentice-Hall, 1990. 303 p.
- [3] **Duecker, K.J.** Geographic information systems and computer-aided mapping // Journal of the American Planning Association. 1987. Vol. 53. P. 383–390.
- [4] **Dangermond, J.** Introduction and overview of GIS // Paper presented at Geographic Information Systems Seminar: Data sharing – Myth or reality. Ontario: Ministry of Natural Resources, 1988. P. 128–145.
- [5] **Shekar, S., Hiong, H.** Enciclopedia of GIS. New York: Springer, 2008. 1370 p.
- [6] **Баранов Ю.Б., Берлянт А.М., Капралов Е.Г., Кошкарёв А.В., Серапинас Б.Б., Филиппов Ю.А.** Геоинформатика. Толковый словарь основных терминов. М.: ГИС-Ассоциация, 1999. 204 с.
Baranov, Yu.B., Berlyant, A.M., Kapralov, E.G., Koshkarev, A.V., Serapinas, B.B., Filippov, Yu.A. Geoinformatics. Explanatory dictionary. Moscow: GIS-assotsiatsiya, 1999. 204 p. (in Russ.)
- [7] **Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W.** Geographic Information Systems and Science. New York: John Wiley & Sons, 1999. 432 p.
- [8] **Korte, G.** The GIS Book. 5th Updated edition. New York: OnWord Press, 2000. 645 p.
- [9] **Bocher, E., Neteler, M.** Geospatial free and open source software in the 21st Century // Proceedings of the first Open Source Geospatial Research Symposium, OGRS 2009. Berlin: Springer-Verlag, 2012. 261 p.
- [10] **Кошкарёв А.В.** Понятия и термины геоинформатики и ее окружения: учебно-справочное пособие. М.: ИГЕМ РАН, 2000. 76 с.
Koshkarev, A.V. Concepts and terms of geoinformatics and its surroundings: uchebno-spravochnoe posobie. Moscow: IGEM RAN, 2000. 76 p. (in Russ.)
- [11] **Осокин С.А.** Теоретические основы и методика создания локальной инфраструктуры пространственных данных: Автореф. дис. канд. географ. наук. М.: МГУ, 2010. 24 с.
Osokin, S.A. Theoretical foundations and a technique for design of a local spatial data infrastructure: Ph.D. Dissertetation. Moscow: MGU, 2010. 24 p. (in Russ.)
- [12] **Leszek, L., Maciej, R.** Geoinformation metadata in INSPIRE and SDI. Berlin; Heidelberg: Springer-Verlag, 2011. 168 p.
- [13] **Дубинин М.Ю., Рыков Д.А.** Открытые настольные ГИС: обзор текущей ситуации // Информационный бюллетень ГИС-Ассоциации. 2009. № 5(72). С. 20–27.
Dubinin, M.Yu., Rykov, D.A. Open desktop GIS: an overview of the current situation // Informatsionnyy byulleten' GIS-Assotsiatsii. 2009. Vol. 5(72). P. 20–27. (in Russ.)
- [14] **Balram, Sh., Dragicevic, S.** Colloborative geographic information systems. Hershey, Pensilvania: Idea Group Publishing, 2006. 364 p.
- [15] **Scott, D.** GIS for web developers: Adding where to your web applications. Raleigh, North Carolina: The Pragmatic Bookshelf, 2007. 260 p.

- [16] **Parashar, M., Li, X., Chandra, S.** Advanced computational infrastructures for parallel and distributed applications. Hoboken, New Jersey, USA: Willey-Interscience, 2009. 518 p.
- [17] Data Mining in Grid Computing Environments / Ed. Werner Dubitzky. Oxford: Wiley-Blackwell, 2008. 288 p.
- [18] **Reese, G.** Cloud application architectures: Building applications and infrastructure in the cloud. Sebastopol, CA: O'Reilly, 2009. 204 p.
- [19] **Yee, R.** Pro Web 2.0 Mashups: Remixing data and web services. New York: Apress, 2008. 603 p.
- [20] **Кошкарёв А.В.** Эффективное управление пространственными метаданными и геосервисами в инфраструктурах пространственных данных // Пространственные данные. 2008. № 1. С. 28–35. (Адрес доступа: www.gisa.ru/44539.html).
- Koshkarev, A.V.** Effective management of spatial metadata and geo-services in spatial data infrastructure // Prostranstvennyye Dannye. 2008. No. 1. P. 28–35. (Available at: www.gisa.ru/44539.html). (in Russ.)
- [21] **Кошкарёв А.В.** Геопортал как инструмент управления пространственными данными и геосервисами // Пространственные данные. 2008. № 2. Адрес доступа: <http://gisa.ru/45968.html>.
- Koshkarev, A.V.** Geportal as a tool for managing spatial data and geoservices // Prostranstvennyye Dannye. 2008. No. 2. Available at: www.gisa.ru/45968.html (in Russ.)
- [22] **Якубайлик О.Э.** Методы построения прикладных геоинформационных систем на основе картографических веб-сервисов геопортала // Труды Международной конференции “Математические и информационные технологии, МИТ-2011”. Косовская Митровица, Сербия: Альфа университет, 2011.
- Yakubaylik, O.E.** Methods for constructing of applications based on geographic information systems mapping using Web services Geportal // Sbornik trudov Mezhdunarodnoy Konferentsii “Mathematical and Informational Technologies, MIT-2011”. Kosovskaya Mitrovitsa, Serbiya: Alfa Universitet, 2011. (in Russ.)
- [23] **Buaya, R., Broberg, J., Goscinski, A.M.** CLOUD COMPUTING: Principles and paradigms. Hoboken, New Jersey: John Wiley & Sons, 2011. 637 p.
- [24] **Fiore, S., Aloisio, G.** Grid and cloud database management. New York: Springer, 2011. P. 279–306. DOI: 10.1007/978-3-642-20045-8_14.
- [25] **Mahmood, Z., Hill, R.** Cloud computing for enterprise architectures. New York: Springer Science & Business Media, 2011. 343 p.
- [26] **Bell, M.** Service-oriented modeling (SOA): Service analysis, design, and architecture. New Jersey: John Wiley & Sons; 2008. 366 p.
- [27] **Keller, R., Baru, Ch.** Geoinformatics: Cyberinfrastructure for the Solid Earth Sciences. Cambridge: University Press, 2011. 374 p.
- [28] **Inmon, W.H.** Building the data warehouse. Third Edition. New York: Willey & Sons, 2002. 424 p.
- [29] **Mohanty, S., Jagadeesh, M., Srivatsa, H.** Big Data imperatives. New York: Apress, 2013. 311 p.
- [30] **Patil, DJ.** Building data science teams. USA: O'REILLY Media Incorporated, 2011. 18 p.
- [31] **Barlow, M.** Real-time big data analytics: Emerging architecture. USA: O'Reilly Media Incorporated, 2011. 26 p.
- [32] **Guo, Sh.** Hadoop operations and cluster management cookbook. Birmingham: Packt Publishing Ltd, 2013. 368 p.

- [33] **Perera, S.** Thilina Gunarathne Hadoop MapReduce Cookbook. Birmingham: Packt Publishing Ltd, 2013, 300 p.
- [34] **Aji, A., Wang, F., Vo, H., Lee, R., Liu, Q., Zhang, X., Saltz, J.** Hadoop-GIS: A high performance spatial data warehousing system over MapReduce // The Proceedings of the VLDB Endowment. 2013. Vol. 6(11). P. 1009–1020. (Available at: www.vldb.org/pvldb/vol6/p1009-aji.pdf).
- [35] Remote sensing-advanced techniques and platforms / Ed. B. Escalante-Ramirez. Rijeka, Croatia: InTech Europe, 2012. 462 p.
- [36] Integration GIS and remote sensing / Ed. V. Mesev. Chichester: Wiley, 2007. 296 p.
- [37] Transforming remote sensing data into information and applications / Eds Steering Committee on Space Applications and Commercialization, Space Studies Board, Ocean Studies Board, Division on Engineering and Physical Sciences, Division on Earth and Life Studies, National Research Council: 2001. 90 p. Available at: <http://www.nap.edu/catalog/10257.html>
- [38] **Бычков И.В.** Корпоративная интеллектуальная технология обработки пространственно-распределенных данных в задачах управления регионом: дис. ... д.т.н. Кемерово, 2003. 240 с.
Buchkov, I.V. Corporate intellectual processing technology for spatially distributed data management problems in the region: The thesis for the degree of Doctor of Technical Sciences. Kemerovo, 2003. 240 p. (In Russ.)
- [39] **Попович В.В., Потапычев С.Н., Панькин А.В., Шойда С.С., Воронин М.Н.** Интеллектуальная ГИС в системах мониторинга // Труды СПИИРАН 2006. № 3, т. 1. С. 172–184. Адрес доступа: <http://www.proceedings.spiiras.nw.ru>
Popovich, V.V., Potapuchev, S.N., Pankin, A.V., Shayda, S.S., Voronin, M.N. Intellectual GIS in systems of monitoring // SPIIRAS Proceedings. 2006. No. 3, vol. 1. 172–184. Available at: <http://www.proceedings.spiiras.nw.ru> (in Russ.)
- [40] **Самодумкин С.А.** Интеллектуальные геоинформационные системы // Матер. Междунар. науч.-техн. конф. “Открытые семантические технологии проектирования интеллектуальных систем”. Минск, 2011. С. 521–525.
Samodumkin, S.A. Intelligent geographic information systems // Proceedings of the International Scientific Conference “Open Semantic Technologies of Intelligent Systems”. Minsk, 2011. P. 521–525. (in Russ.)
- [41] **Куссуль Н.Н., Шелестов А.Ю., Скакун С.В., Кравченко А.Н.** Интеллектуальные вычисления в задачах обработки данных наблюдения Земли. Киев: Наукова думка, 2007. 196 с.
Kussul', N.N., Shelestov, A.Yu., Skakun, S.V., Kravchenko, A.N. Intelligent computing problems for processing of Earth observation data. Kiev: Naukova dumka, 2007. 196 p. (in Russ.)
- [42] Geographic data mining and knowledge discovery. Second Edition / Eds H.J. Miller, J. Han. London: Press Taylor & Francis Group, 2009. 461 p.
- [43] **Fisher, R., Breckon, T.P., Dawson-Howe, K., Fitzgibbon, A., Robertson, C., Trucco, E., Williams, K.I.** Dictionary of computer vision and image processing. Second edition. New Jersey: John Wiley & Sons, 2013. 382 p.
- [44] **Ponniah, P.** Data warehousing fundamentals for IT professionals. New York: John Wiley & Sons, 2010. 601 p.
- [45] **Геловани В.А., Башлыков А.А., Бритков В.Б., Вязилов Е.Д.** Интеллектуальные системы поддержки принятия решений в нештатных ситуациях с использованием информации о состоянии природной среды. М.: Институт системного анализа РАН, 2001. 304 с.

- Gelovani, V.A., Bashlykov, A.A., Britkov, V.B., Vyazilov, E.D.** Intelligent decision support systems in emergency situations that uses information on the state of the environment. Moscow: Institut Sistemnogo Analiza RAN, 2001. 304 p. (in Russ.)
- [46] **Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M.** Agent-based models of geographical systems. Springer, 2012. 747 p.
- [47] **Тoffоли Т., Марголуc Н.** Машины клеточных автоматов. М.: Мир, 1991. 280 с.
Toffoli, T., Margolus, N. Machinery of cellular automata. Moscow: Mir, 1991. 280 p. (in Russ.)
- [48] **Фон Нейман Дж.** Теория самовоспроизводящихся автоматов. М.: Мир, 1971. 384 с.
Von Neumann, J. The theory of self-reproducing automata. Moscow: Mir, 1971. 384 p. (in Russ.)
- [49] **Poli, R., Healy, M., Kameas, A.** Theory and applications of ontology: Computer applications. New York: Springer, 2010. 585 p.
- [50] **Asch, K., Brodaric, B., Laxton, J.L., Robida, F.** An international initiative for data harmonization in geology: 10th EC GI & GIS Workshop, ESDI State of the Art. Warsaw, 23–25 June, 2004. 9 p.

Поступила в редакцию 17 сентября 2015 г.

GIS today: current state, perspectives, solutions

SHOKIN, YURIY I.¹, POTAPOV, VADIM P.^{2,*}

¹Institute of Computational Technologies SB RAS, Novosibirsk, 630090, Russia

²Institute of Computational Technologies SB RAS, Kemerovo Branch, Russia

*Corresponding author: Potapov, Vadim P., e-mail: potapov@ict.sbras.ru

The current state and prospects for development of geoinformation systems based on their retrospective analysis are both considered. The analysis of the main components of the open source and commercial software systems is done; their advantages and disadvantages are examined.

Based on the known solutions, the design for geoportals technology is considered and a new approach to implementation that extends the functional and computational capabilities of such systems is offered. Integration and examples of GIS software products which are connected with remote sensing systems are described.

A new class of GIS that uses the elements of artificial intelligence is analyzed. The intelligent GIS (IGIS) which allow to progress to new levels of processing and analysis of geo-information, both in local and distributed versions, are considered. For the ever-growing volumes of information, it is encouraged to pay attention to the construction of the specific GIS technology that uses Big Data concept (BIG DATA) and its implementation to the so-called HADOOP clusters. To analyze the latest approaches, we present the examples of specific GIS and conclude on their relevance in establishing of such systems.

Keywords: GIS, RSD, intelligent geographic information systems, metadata, BIG DATA, data mining, data warehouse, cloud computing, SOA, geoportals, multi-agent systems, ontology, knowledge base.

Received 17 September 2015