

Подходы к созданию распределенной системы сбора, хранения и поиска геопространственных данных

Methods for creating distributed systems of geospatial data collection, storage, and search

Ключевые слова: геопространственные данные – geospatial data; каталоги геопространственных данных – geospatial data catalogue; метаданные геопространственных данных – metadata of geospatial data; сетевентрические системы – network-centric systems; интеграционная платформа – integration platform; единое информационное пространство – common information space.

Рассматриваются особенности сетевентрических систем, повышающие эффективность процессов сбора, хранения и поиска геоданных. Разработана концептуальная модель единого информационного пространства сбора, хранения и поиска геоданных. Предложена типовая структура обрабатываемых геоданных в территориально-распределенных базах. Разработан алгоритм распределенного поиска геоданных в каталогах.

This work discusses certain features of network-centric systems that increase the efficiency of data collection, storage, and geodata information search. The concept model of the common information space for geodata collection, storage, and search is developed. A new typical structure for geodata processing in geographically distributed databases is proposed. An algorithm for distributed geodata search in catalogs is developed.

ВВЕДЕНИЕ

Высокие темпы современного развития космических средств и технологий дистанционного зондирования Земли (ДЗЗ) способствуют широкому применению геопространственных данных при решении различных социально-экономических и научных задач. Состав информации, необходимой для принятия решений государственными и муниципальными органами власти, а также силовыми министерствами и ведомствами в процессе их деятельности, все больше напоминает слоеный

КАРИН / KARIN S.

Сергей Александрович

(sergey.karin@gmail.com)

кандидат военных наук.

ФГКБОУ ВПО «Военно-космическая академия имени А.Ф. Можайского» МО РФ (ВКА им. А.Ф. Можайского), заместитель начальника кафедры оптико-электронных средств.

г. Санкт-Петербург

ДУДИН / DUDIN E.

Евгений Александрович

(DudinEA@yandex.ru)

кандидат технических наук.

ВКА им. А.Ф. Можайского, доцент кафедры оптико-электронных средств.

г. Санкт-Петербург

пирог, с постоянно увеличивающимся количеством слоев – видов используемых данных. Кроме данных о местности растет поток используемой в процессе принятия решений неструктурированной или слабоструктурированной информации, которую необходимо анализировать и учитывать.

При этом основным видом обрабатываемых данных являются геопространственные данные (геоданные). Можно утверждать, что до 80 процентов информационных ресурсов имеют координатную привязку [1].

Под геопространственными данными понимаются данные о географических объектах, которые являются формализованными цифровыми моделями материальных или идеальных (абстрактных) объектов реального или виртуального мира. Полное цифровое описание пространственного объекта и содержание пространственных данных складываются из идентификатора объекта, набора его атрибутов (свойств, характеристик) с их значениями и параметров локализации объекта в пространстве и времени (в некоторой системе пространственно-временных координат). Правила цифрового моделирования объектов реальности путем описания типов и свойств элементарных пространственных объектов, их наборов и межобъектных отношений называют моделью пространственных данных.

Уже сегодня объемы поступающих из различных источников геопространственных данных настолько колоссальны, что противоречия, возникающие между возросшими потоками информации и имеющимися организационно-техническими возможностями по их сбору, обработке, поиску и использованию, обуславливают необходимость разработки новых методов организации этих процессов. Одним из перспективных методов является использование технологий на базе концепции сетецентрических систем [2–4].

АНАЛИЗ РОЛИ, МЕСТА И ПЕРСПЕКТИВ ПРИМЕНЕНИЯ СЕТЕЦЕНТРИЧЕСКИХ СИСТЕМ В ПРОЦЕССАХ СБОРА, ХРАНЕНИЯ И ПОИСКА ГЕОПРОСТРАНСТВЕННЫХ ДАННЫХ

Главной особенностью сетецентрических систем является сетевой принцип их построения, обеспечивающий максимальное расширение возможностей получения информации в рамках предметной области, доступа к ней, ее распределения и обратной связи. При этом с технологической точки зрения сетецентрические системы предполагают наличие трех взаимосвязанных типовых элементов [5]:

- 1) информационных датчиков, обеспечивающих генерирование информации в зоне ответственности;
- 2) средств воздействия на внешнюю среду, которые обеспечивают непосредственное воздействие на контролируемые объекты в зоне ответственности;
- 3) информационно-управляющих средств, реализующих функции хранения и анализа информации, управления первыми двумя элементами, а также информирования лиц, принимающих решения.

Другим важнейшим отличием сетецентрической системы от систем других классов является то, что выход из строя одного или нескольких ее узлов не приводит к выходу из строя всей системы в целом. Тем самым сетецентрические системы способствуют созданию Единого информационного пространства, что, в свою очередь, позволяет повысить оперативность сбора и обработки разных видов информации, в том числе и геопространственной [5].

Проведенный анализ состава и структуры обрабатываемых геоданных показывает их высокую разнородность [1]. Вместе с тем существующая технология обработки геоданных в органах государственного и муниципального управления, а также в силовых министерствах и ведомствах во многом предполагает ручной режим работы, а также отсутствие единого информационного пространства. Все это, с учетом роста объемов

добываемых геоданных, приводит к снижению оперативности их обработки.

Таким образом, существующая технология обработки геоданных обладает следующим набором недостатков:

- 1) в процессе обмена данными между органами управления не предполагается взаимодействия автоматизированных систем;
- 2) в процессе обмена данными между органами управления не предполагается децентрализация (все информационные потоки идут через центральные органы управления);
- 3) существует необходимость заполнения и подачи заявок в неавтоматизированном режиме для получения требуемой информации;
- 4) возникает необходимость длительного ожидания обработки поданных заявок и получения требуемой информации (не менее 2 суток).

Для устранения указанных недостатков необходима разработка основанной на концепции сетецентрических систем технология обработки геопространственных данных, которая бы включала в свой состав:

- 1) модель единого информационного пространства обработки геоданных на основе концепции сетецентрических систем;
- 2) требования к составу обрабатываемых в рамках единого информационного пространства геоданных;
- 3) типовую структуру обрабатываемых геоданных;
- 4) модель (алгоритм) поиска требуемых геоданных в едином информационном пространстве.

РАЗРАБОТКА ТЕХНОЛОГИИ ОРГАНИЗАЦИИ ХРАНЕНИЯ И ПОИСКА ГЕОПРОСТРАНСТВЕННЫХ ДАННЫХ

Требования к составу обрабатываемых геоданных зависят в первую очередь от задач, решаемых пользователями в их предметной области. Способы решения этих задач во многом определяются реализованными организационно-техническими условиями. В составе последних немаловажную роль играют способы организации информационного взаимодействия.

В соответствии с этим, в рамках проведенных исследований была разработана концептуальная модель единого информационного пространства сбора, хранения и поиска геопространственных данных, которая представлена рис. 1.

Отличительной особенностью данной модели является использование территориально-распределенной базы частных каталогов геопространственных данных и организация доступа к храня-

ГЕОИНФОРМАТИКА

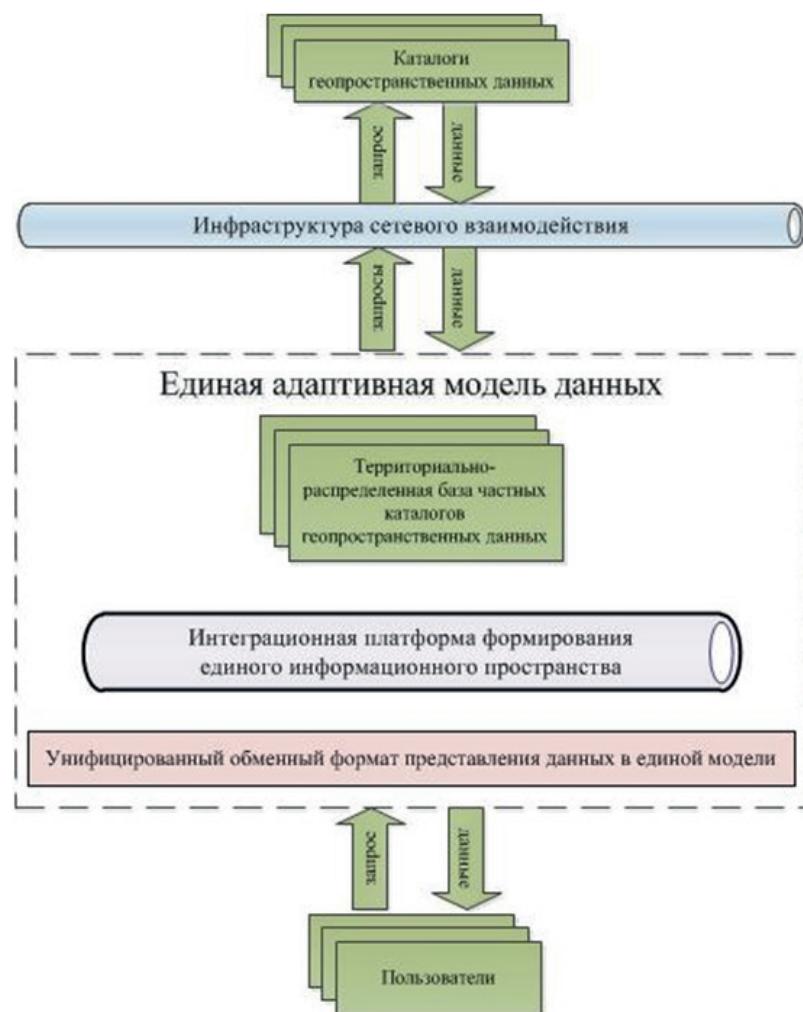


Рис. 1. Концептуальная модель единого информационного пространства сбора, хранения и поиска геопространственных данных



Рис. 2. Состав типового экземпляра геоданных, адаптированного для хранения в территориально-распределенной базе частных каталогов

щимся в каталогах данным посредством интеграционной платформы формирования единого информационного пространства [6–9]. При этом последняя играет роль связующего звена между узлами территориально-распределенной базы [1] и обеспечивает гарантированную доставку информации даже при нестабильных каналах связи [8, 9].

Информационное наполнение территориально-распределенной базы зависит от состава и структуры хранящихся в ней геоданных. Соответственно встает вопрос о разработке требований к ним.

На рис. 2 представлены требования по составу типового экземпляра геоданных, который адаптирован для хранения в территориально-распределенной базе частных каталогов.

В рамках этих требований структуру и состав метаданных предлагается реализовать с использованием ГОСТ Р 52573–2006 [10]. При этом к составу метаданных, обеспечивающих данными ГОСТ, необходимо добавить набор специфических метаданных, разбитых на две группы:

- 1) информация об объектовом составе;
- 2) информация специального назначения.

В состав первой группы должна входить информация об объектах управления (перечень которых ведется в соответствующих органах управления), тем или иным образом связанных с описываемым экземпляром геоданных.

В состав второй группы должны входить сведения, обеспечивающие формирование оптимальных поисковых запросов к каталогам геопространственных данных.

Необходимость введения этой группы метаданных обусловлена тем, что эффективность поиска может быть повышена, если при формировании поисковых запросов будет учтена специфика распределения геопространственных данных в каталогах. К примеру, многие космические аппараты ДЗЗ передают в приёмный центр данные только на определённую территорию (режим непосредственного сброса). В то же время с некоторых аппаратов (например, Электро-Л) передача данных в различные центры выполняется по регулярному расписанию.

Учёт этих особенностей при выполнении распределённого поиска геопространственных данных позволяет обеспечить эффективное разделение поисковых запросов к каталогам. Запрос к частному каталогу может содержать условия ограничения области поиска, что приведёт к значительному уменьшению нагрузки на программно-аппаратные средства центра и увеличит скорость обработки запросов.

Проведённый в рамках данного исследования анализ существующих методов распределённого поиска показал, что они не учитывают указанные особенности. Например, архитектура сервиса каталога геопортала Евросоюза, решающего задачу интеграции территориально-распределённых каталогов европейских организаций, ведающих созданием и распространением геопространственных данных, предусматривает, что сервис распределённого поиска, обрабатывающий запросы пользователей, лишь перенаправляет эти запросы всем сервисам поиска по частным каталогам, после чего объединяет результаты их выполнения и выдаёт их пользователю [11].

На рис. 3 представлены принципиальная схема состава и структуры метаданных (общих и специфических).

В состав сведений, обеспечивающих операции поиска геоданных в территориально-распределённой базе (набор специфических метаданных), должны входить:

- 1) дата последнего обновления геоданных по заданному запросу;
- 2) узел территориально-распределенной базы частных каталогов, где хранится исходный экземпляр геоданных;
- 3) права доступа в соответствии с политиками разграничения доступа.

Таким образом, указанные требования, с одной стороны, обеспечивают унификацию хранящихся геоданных, а с другой – служат основой для функционирования территориально-распределённой базы частных каталогов как основы сетевентрической системы.

С учетом представленных требований предлагается следующий алгоритм поиска требуемых геоданных в территориально-распределённой базе частных каталогов, схема которого приведена на рис. 4. Данный алгоритм реализован авторами в составе автоматизированной информационной системы сбора, обработки и анализа разнородных пространственных данных [6–9].

В данном алгоритме можно выделить 3 уровня процесса поиска необходимой информации:

- 1) Уровень пользователя. На данном уровне формируются критерии для отправки поискового запроса с учетом метаданных специального назначения. Происходит формирование списка частных каталогов геопространственных данных, которые могут содержать требуемые сведения (отсекаются каталоги, которые искомые сведения точно не содержат). Осуществляется отправка поисковых запросов на сервера каталогов из сформированного списка и переход в режим ожидания ответов от них. Когда ответы от всех опрошенных серверов

ГЕОИНФОРМАТИКА



Рис. 3. Принципиальная схема состава и структуры метаданных

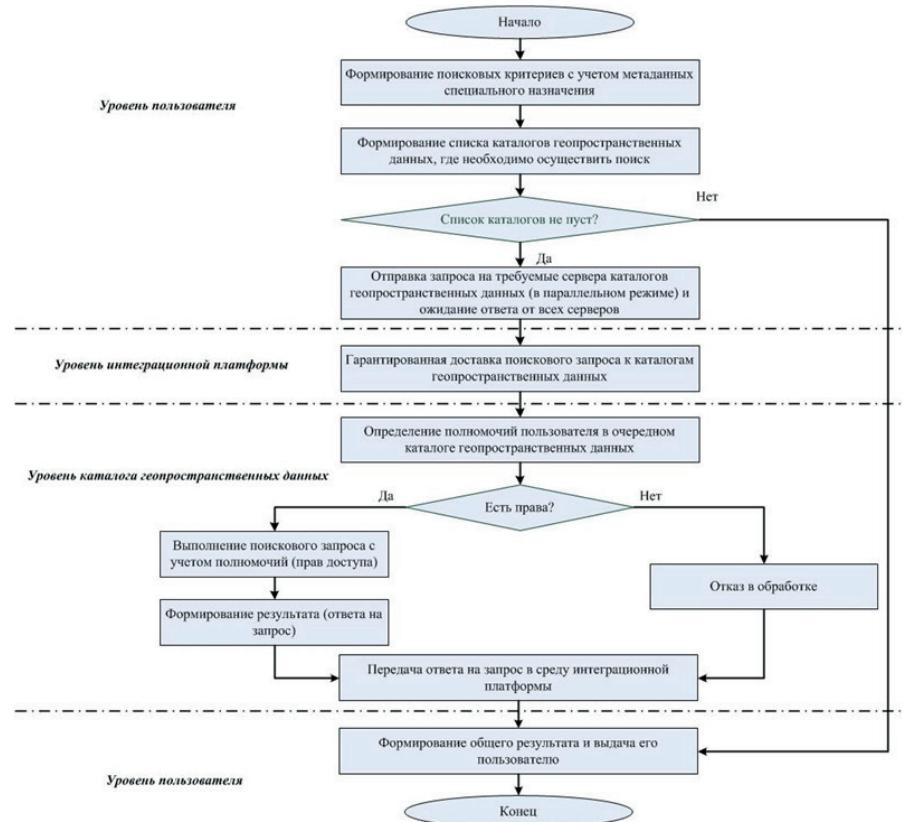


Рис. 4. Схема алгоритма поиска требуемых геоданных в территориально-распределенном банке

будут получены, происходит формирование общего результата выполнения запросов и выдача его пользователю. Фактически данный этап гарантирует, что нагрузка на инфраструктуру сетевого взаимодействия будет минимально необходимой.

2) Уровень интеграционной платформы. На данном уровне интеграционная платформа формирования единого информационного пространства обеспечивает гарантированную доставку поискового запроса на все узлы территориально-распределенной базы частных каталогов, функционирующие в рамках сетецентрической системы, а также возврат ответа от них пользователю. Гарантированность доставки обеспечивается путем применения специального программного обеспечения, реализованного авторами [6–9].

3) Уровень каталога геопространственных данных. На данном уровне обеспечивается выполнение поисковых запросов на всех функционирующих узлах территориально-распределенной базы частных каталогов геопространственных данных и отправка результатов потребителю. Отправка осуществляется посредством интеграционной платформы.

ВЫВОДЫ

Предложенный состав геоданных, а также алгоритм их поиска с использованием территориально-распределенной базы частных каталогов геопространственных данных предполагает снижение временных затрат на поиск информации в процессе разработки отчетных информационных документов и, в свою очередь, значительное сокращение времени, прошедшего с момента поступления данных в распределенную систему до момента непосредственного представления обработанных результатов потребителю.

Реализация данных подходов позволит повысить оперативность и обоснованность принятия требуемых управленческих решений, что особенно важно в условиях быстременяющейся обстановки, когда оперативная доставка информации решает исход событий.

В настоящее время коллективом разработчиков, возглавляемым авторами, разработан прототип автоматизированной информационной системы (АИС) сбора, обработки и анализа разнородных пространственных данных, которая реализует представленные в настоящей статье подходы и которая легла с основу ряда опытно-конструкторских работ [6–9, 12]. Использование данной АИС позволило значительно повысить качество выполнения работ, снизить сроки выполнения ОКР, а также обеспечить полное выполнение требований технических заданий.

Литература

1. Карин, С.А. Интеграция в едином информационном пространстве разнородных геопространственных данных / С.А. Карин // Информационно-управляющие системы. – 2012. – №2. – С. 89–94.
2. Mica R.Endsley, Daniel J. Garland, Situation awareness: analysis and measurement, Lawrence Erlbaum Associates, 2000, ISBN 0805821341, 9780805821345.
3. Ерёменко, Е.Н. Неогеография: особенности и возможности / Е.Н. Ерёменко // Материалы конференции «НеогеографияXXI-2008» IX Международного Форума «Высокие технологии XXI века». – Москва, 22–25 апреля 2008 года. – С. 170.
4. Ерёменко, Е.Н. Неогеография и Situational Awareness / Е.Н. Ерёменко // Материалы конференции «НеогеографияXXI-2009» X Международного Форума «Высокие технологии XXI века». – Москва, 21–24 апреля 2009 года. – С. 434–436.
5. Карин, С.А. Разработка подходов к реализации распределенного банка геопространственных данных с учетом особенностей сетецентрических систем / С.А. Карин, Р.А. Зелянин // Труды Военно-космической академии. Выпуск 635. – СПб.: ВКА им. А.Ф. Можайского, 2012. – С. 28–34.
6. Программная платформа поддержки создания и обеспечения функционирования прикладного программного и лингвистического обеспечения типовых территориально-распределенных защищенных автоматизированных систем сбора и обработки разнородной информации (DynamicDocs) / С.А. Карин [и др.]. – РОСПАТЕНТ. Свидетельство № 2012614987 от 04.06.2012.
7. Карин, С.А. Программный комплекс динамической рубрикации информационных ресурсов в типовых территориально-распределенных защищенных автоматизированных системах сбора и обработки разнородной информации (DynamicRubrics) / С.А. Карин. – РОСПАТЕНТ. Свидетельство № 2013660371 от 31.10.2013.
8. Карин, С.А. Программный комплекс формирования единого информационного пространства в типовых территориально-распределенных защищенных автоматизированных системах сбора и обработки разнородной информации (DynamicInteractor) / С.А. Карин. – РОСПАТЕНТ. Свидетельство № 2014614562 от 29.04.2014.
9. Карин, С.А. Программный комплекс мониторинга функционирования узлов типовых территориально-распределенных защищенных автоматизированных системах сбора и обработки разнородной информации (DynamicDBMonitor) / С.А. Карин. – РОСПАТЕНТ. Свидетельство № 2014615492 от 28.05.2014.
10. ГОСТ Р 52573–2006. Географическая информация. Метаданные.
11. Ioannis Kanellopoulos. The INSPIRE EU GeoPortal / European Commission Joint Research Centre Institute for Environment and Sustainability. – Режим доступа: <http://www.ec-gis.org>, свободный. – Загл. с экрана.
12. Карин, С.А., Инновационная платформа для создания специализированных приложений обработки пространственных данных / С.А. Карин // Труды Военно-космической академии. Выпуск 645. – СПб.: ВКА им. А.Ф. Можайского, 2014. – С. 28–34.