

нетика. 1980. № 6. С. 36–43.

14. Бондур В.Г., Журбас В.М., Гребенюк Ю.В. Математическое моделирование турбулентных струй глубинных стоков в прибрежные акватории // Океанология. 2006. Т. 46. № 6. С. 805–820.

15. Лобанов А.А., Цветков В.Я. Пространственное моделирование // Славянский форум. 2015. № 1(7). С. 137–142.

16. Цветков В.Я. Информационное моделирование. М.: Московский государственный технический университет радиотехники, электроники и автоматики (МГТУ МИРЭА), 2015. 60 с.

17. Tsvetkov V.Ya. Spatial Information Models // European Researcher. 2013. Vol. (60). № 10-1. P. 2386–2392.

18. Заварзин Г.А. Антипод ноосферы // Вестник РАН. 2003. Т. 73. № 7. С. 627–636.

19. Гвинн М.Д., Селла Ф., Валлен К.К. Глобальная система мониторинга окружающей среды: принципы и прогресс // Комплексный глобальный мониторинг загрязнения окружающей природной среды. Труды Международного симпозиума. Л., 1980.

20. Tsvetkov V.Ya. Global Monitoring // European Researcher. 2012. Vol. (33). № 11-1. P. 1843–1851.

21. Бондур В.Г., Килер Р.Н., Старченков С.А., Рыбакова Н.И. Мониторинг загрязнений прибрежных акваторий океана с использованием многоспектральных спутниковых изображений высокого пространственного разрешения // Исследование Земли из космоса. 2006. № 6. С. 42–49.

22. Davies D. K. et al. Fire information for resource management system: archiving and distributing MODIS active fire data // Geoscience and Remote Sensing, IEEE Transactions on. 2009. Т. 47. № 1. С. 72–79.

23. Соловьев В.С., Козлов В.И., Муллаяров В.А. Дистанционный мониторинг лесных пожаров и гроз в Якутии. Якутск: Изд-во ЯНЦ СО РАН, 2009. 108 с.

Geoinformation monitoring fires

Alexandr Anatol'evich Lobanov, Ph.D., Associate Professor, Moscow State Technical University of Radio Engineering, Electronics and Automation MIREA

This article describes methods of geoinformation monitoring. Geoinformation monitoring is used for monitoring and suppression of forest fires. This article describes the space monitoring. Space monitoring is an integral part of geoinformation monitoring. This article describes a specialized information system monitoring. Article shows the details of modeling for monitoring. Integrated monitoring is the basis for monitoring flatterling fires.

Keywords: space research, monitoring, satellite monitoring, geoinformation monitoring, fires

УДК 004.8+528.06

ДОБЫЧА ДАННЫХ И ГЕОДАНЫХ

Владимир Михайлович Маркелов, соискатель,

E-mail: vmarkel123456@yandex.ru,

Московский государственный университет геодезии и картографии,

http://www.miiigaik.ru

Статья описывает новую интеллектуальную технологию – интеллектуальный анализ геоданных. Технология является развитием известной технологии Data Mining. Описана эволюция понятия геоданных. Статья показывает различие между технологиями Data Mining и GeoData Mining. Статья раскрывает понятия геоинформационное знание, пространственное знание и геознание. Статья описывает проблемы интеллектуализации анализа геоданных.

Ключевые слова: науки о Земле, геоинформатика, интеллектуальные технологии, гео-

знание, пространственное знание.

Введение

Интеллектуальный анализ данных – собирательное название технологий, используемое для обозначения совокупности технологий обнаружения в данных ранее неизвестных, практически полезных и доступных для интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности. Часто это направление обозначают Data mining [1], что сужает область исследования. Однако существует направление латентного анализа [2], которое также решает эти задачи. Оно



В.М. Маркелов

может использовать интеллектуальные методы анализа, а может обходиться без них. Коррелятивный анализ [3] также решает задачи нахождения скрытых зависимостей и связей, и построение на этой основе новых показателей, функций и систем. Импакт-анализ [4] и кластерный анализ [5] также решают задачи выявления скрытых зависимостей и связей. Все эти виды анализа могут быть дополнены и усилены методами интеллектуального анализа, но первоначально возникли как алгоритмические и информационные методы. В России в качестве эквивалента Data mining устойчиво употребляют термин «добыча данных», но не всегда его привязывают к интеллектуальной обработке.

В данной работе, употребляя термины «добыча данных» или «добыча геоданных», будем иметь в виду интеллектуальный анализ данных и геоданных. Следует отметить, что геоданные [6] являются более структурированными и классифицированными, чем данные вообще. Это накладывает специфику на их интеллектуальный анализ. Необходимо также различать понятия интеллектуальный анализ [7] и интеллектуальная обработка [8].

Организация геоданных. Множество данных и геоданных собирают с помощью разных технологий и систем. Данные отражают различные качества, характеристики и свойства. Они могут иметь различные размерности разное количество значащих цифр, разное число разрядов, разную точность и т. д. Собранные данные могут храниться в виде наборов или файлов. Кроме того, при сборе данные могут организовывать связанные совокупности, называемые моделями данных. Для того чтобы разнородные данные и модели можно было обрабатывать в одной системе они должны быть упорядочены и сведены к единой информационной модели, в которой они будут дополнять друг друга.

Организацией данных называется процедура сведения разнородных данных и моделей в единую непротиворечивую информационную модель, которую в дальнейшем можно будет эффективно применять в различных технологиях анализа и управления. Эту особую информационную модель называют информационной основой. Организация данных предшествует интеллектуальному анализу данных.

Результатом организации данных является создание универсальной информационной конструкции, которая позволяет организовать хранение в базе данных. Следовательно, организация геоданных даёт возможности создания для БД и возможность их автоматизированной обработки. Геоданные для их использования должны быть классифицированы, унифицированы, интегрированы и стратифицированы.

Первым этапом организации данных является сбор информации. Он формирует так называемые первичные данные. Исходная первичная информация включает множество параметров, многие из которых дублируют друг друга. Уменьшение числа данных о реальных объектах достигается применением разных моделей, сохраняющих основные свойства объектов исследования и не содержащих второстепенных свойств.

Одной из особенностей сбора данных в геоинформатике является то, что исходные данные могут иметь не только разные размерности, но и измеряться в разных шкалах измерений. Организация геоданных направлена на объединение данных разных размерностей и шкал измерений в единую систему данных для их хранения и последу-

ющей обработки. Именно это создаёт возможность комплексного анализа данных [9], при работе с разнородными исходными данными, измеренными в разных шкалах измерений.

Поэтому следующим этапом является классификация собранной информации, которая служит основой дальнейших действий. Классификация данных позволяет соотносить различные модели и их характеристики к разным классам, подклассам и типам, что даёт возможность систематизировать исходные наборы данных и использовать свойства классов при анализе конкретных данных. Как дополнительный этап классификации гео-данных в геоинформатике присутствует процедура локализации данных [10].

После того, как данные классифицированы, осуществляется их унификация. Разнообразие технологий и методов сбора данных порождает разнообразие типов данных, которые впоследствии необходимо обрабатывать. Обрабатывать множество различных данных неудобно и неэффективно. Для упрощения процесса обработки, хранения и обмена разнородные данные приводят к единому структурному виду, который используется при последующей обработке информации. Такие данные называют унифицированными.

Процедура сведения разнородных видов и структур данных к единому виду и структуре называется унификацией. В ходе унификации данных осуществляется построение единой формы данных. После этих процедур возможно построение интегрированной модели. Унификация не создает систему данных, преобразует исходную совокупность разнородных и несогласованных данных в другую, но уже более согласованную и менее разнородную.

Для обработки по единой технологической системе и в единой информационной среде модели должны быть объединены на основе правила или метода, отвечающего требованиям оптимального хранения и обработки. Таким объединяющим методом является интеграция данных. Интеграция данных и создает систему данных вместо совокупности данных

Необходимо отметить, что геоданные образуют естественную информационную систему данных. Это обусловлено тем, что они отображают реальные объекты и явления земной поверхности, которые расположены не произвольно, а организовано и имеют объективные связи друг с другом. Можно говорить, что информация об объектах и явлениях земной поверхности образует некую систему. Отдельные модели или данные являются элементами такой системы.

Получение знаний в геоинформатике. Одной из задач геоинформатики является формирование информационных ресурсов и получение новых знаний [11] для решения различных задач. Информационные ресурсы включают различные компоненты: данные, геоданные, информацию, базы данных, знания и технологические системы. Геоданные имеют двойственное значение. Первичные геоданные представляют собой исходные измерения. Вторичные геоданные представляют собой сформированные модели различного вида. Геоданные появились через потребность общества для решения задач и понимания Земли как сложной системы.

Последние достижения в сфере информационных и интеллектуальных технологий, такие как суперкомпьютеры, интеллектуальный анализ, вычислительные методы, облачные вычисления, визуализация моделей привели технологии добыча геоданных.

Термин «геоданные» первоначально был связан с геологией и геодезией и означал специализированные данные в этих областях. С появлением геоинформатики термин «геоданные» стал использоваться как обобщение и как интегрированная совокупность данных, применяемая в разных предметных областях. По мере развития геоинформатики и понятия геоданные появлялись разные технологии обработки геоданных. Одной из таких технологий является технология «добычи геоданных».

Добыча геоданных (GeoData Mining) имеет много общего с добычей данных (Data Mining), если иметь в виду дополнительный интеллектуальный анализ, а не ограничи-

ваться получением данных. Data Mining – это процесс автоматического поиска больших объемов данных для заданных моделей [12]. В процессах поиска Data Mining использует вычислительные методы, распознавание образов и методы искусственного интеллекта. Именно поэтому технологию Data Mining относят к технологиям извлечения знаний и называют технологией интеллектуального анализа. Технология Data Mining данных может быть определена как «нетривиальное извлечение явных, ранее неизвестных и потенциально полезной информации из данных» [13] и «наука извлечения полезной информации из больших массивов данных или базы данных» [14].

Хотя этот термин обычно используется по отношению к анализу данных, тем не менее, большую роль в нем играет искусственный интеллект, что дает основание разграничивать сбор данных и Data Mining. Отметим несколько различий.

Первое различие является технологическим. При сборе данных на основе измерений получают новую форму данных. Это технологии сбора геоданных. При переходе к рассмотрению технологии «добыча геоданных», необходимо отметить применение интеллектуальных методов. В итоге такая технология направлена на получение знаний, а не просто данных.

Второе различие является структурным. Оно подчеркивает различие между «вообще» данными, применяемыми в Data Mining, и геоданными. Данные, применяемые в Data Mining, могут быть любыми, иметь любую структуру и описывать любую предметную область от биологии и медицины до физики твердого тела. Геоданные имеют вполне определенную структуру «место», время», «тема» и описывают пространственные явления, отношения и процессы.

Геоданные – тематические, пространственные и временные данные, отражающие свойства пространственных объектов, процессов и явлений, происходящих на Земле [6].

Сбор геоданных и добыча геоданных разные технологии. Первая является инструментальной, вторая интеллектуальной. Технология сбора геоданных направлена на получение измерений, которые служат основой для построения моделей. Технология добычи геоданных направлена на получение новых знаний.

Геоданные описывают пространственные отношения. Использование пространственных отношений позволяет получать новый вид знания геоинформационное знание, или геознание [15]. Геоинформационное знание – объективный, независимый от отдельного человека, набор знаний, выраженный в количественных и в качественных критериях, о географических фактах, о пространственной окружающей среде, о пространственных отношениях, полученных на основе измерений, на основе анализа, на основе обработки и интеграции информации [16].

Геоинформационное знание. Геоинформационное знание представляется в виде разнообразных информационных продуктов, таких как цифровые модели, цифровые карты, каталоги координат, динамические временные модели, пространственные модели, топологические модели и т. д. Геоинформационное знание часто классифицируется как декларативное, процедурное, и конфигурационное геознание [17].

Геознания или геоинформационные знания, которые получают на основе сбора количественной информации, её обработки и анализа являются структурно согласованными в количественном и качественном отношениях. Вопрос структурной согласованности в геоинформационных знаниях присутствует всегда.

Напомним, что под информацией в кибернетике, по определению Н. Виннера, понимают ту часть знаний, которая используется для ориентирования, активного действия, управления, т. е. в целях поддержания жизнедеятельности и развития системы [18].

Интеллектуализация анализа геоданных является новым направлением в геоинформатике и сталкивается с рядом проблем. Одна из основных проблем – терминологическая. Она состоит в подмене понятия «интеллектуализация обработки» близкими категориями: компьютеризация обработки данных; автоматизация обработки геоданных; автоматизированное измерение геоданных и пр. Одной из причин является то, что близкое

к интеллектуальной обработке понятие «интеллектуальный анализ данных» трактуется в ряде отечественных источниках как автоматизированная обработка данных.

Получение пространственных знаний включает этапы построения терминологического поля, построение онтологий, извлечение знаний из фактов наблюдения (data mining) и результатов обработки пространственных данных. Пространственные знания отражают знания о пространственных объектах, и знания о пространственных и непространственных отношениях.

В настоящее время выделяют три разных типа пространственных знаний. Первый тип пространственных знаний связывают с абстрактными пространственными моделями. Этот тип знаний применяют в области математики, искусственного интеллекта и частично в области геоинформатики. Вторым типом пространственных знаний связывают с пространственными объектами и их моделями, с тем, что расположено на земной поверхности и в реальном пространстве. Этот тип знаний применяют, в первую очередь, в области геоинформатики и в науках о Земле. Во вторую в сфере искусственного интеллекта. Третьим типом пространственных знаний связывают с мышлением и представлением пространственных объектов и моделей. Этот тип знаний применяют, в первую очередь, в области психологии, образования, во вторую, в геоинформатике и искусственном интеллекте.

В научных исследованиях все шире начинают применять понятие геознания (второй тип пространственного знания), как знания связанного с пространственными отношениями. Геознание рассматривается как форма знания, связанного, в первую очередь, с пространственными отношениями на земной поверхности. Отсюда следует возможность более широкого описания геознаний.

Получение пространственных знаний на основе добычи геоданных включает этапы построения терминологического поля, построение онтологий, извлечение знаний из фактов наблюдения и результатов обработки геоданных.

Заключение. Технология «добычи геоданных» является новой интеллектуальной технологией в геоинформатике, направленной на получение новых знаний. Она дополняет и развивает технологии сбора геоданных. Технология «добычи геоданных» является развитием технологии . Data mining. Особенность технологии «добычи геоданных» является в обязательном применении интеллектуальных методов анализа информации. Это становится особо актуальным при работе с большими данными. Технология «добычи геоданных» относится к классу *smart technology*, то есть является промежуточной между информационными технологиями и интеллектуальными технологиями.

Литература

1. *Berry M.J., Linoff G.* Data mining techniques: for marketing, sales, and customer support. John Wiley & Sons, Inc., 1997.
2. *Куликова А.А.* Причинность в моделях латентно-структурного анализа и структурных уравнений // Социология: методология, методы и математическое моделирование. 2009. № 29. С. 30–44.
3. *Tsvetkov V.Ya.* Framework of Correlative Analysis // European Researcher. 2012. Vol. (23). № 6-1. P. 839–844.
4. *Ozhereleva T.A.* Impact Analysis of Education Quality Factors // European Journal of Economic Studies. 2013. Vol. (5). № 3. P. 172–176.
5. *Олдендерфер М.С., Блэифилд Р.К.* Кластерный анализ. Факторный, дискриминантный и кластерный анализ. М.: Финансы и статистика, 1989. 215 с.
6. *Савиных В.П., Цветков В.Я.* Геоданные как системный информационный ресурс // Вестник Российской Академии Наук. 2014. Т. 84. № 9. С. 826–829.
7. *Розенберг И.Н.* Инфраструктура интеллектуальных транспортных систем // Славянский форум. 2012. № 1(1). С. 242–245.
8. *Прокопчук Ю.А., Костра В.В.* Интеллектуальная обработка данных в открытых информационных системах // Проблемы программирования. 2002. № 1-2. С. 390–395.
9. *Кулаичев А.П.* Методы и средства комплексного анализа данных. М.: ИНФРА-М, 2006.

10. *Цветков В.Я.* Геоинформационные системы как системы пространственно-локализованных данных. М.: ГосНИИ ИТТ «Информика», 1999. 113 с. Номер государственной регистрации 0329900095.
11. *Маркелов В.М., Цветков В.Я.* Модели получения знаний в геоинформатике // Славянский форум. 2015. № 1 (7). С. 177–182.
12. *The Cutting Edge. An encyclopedia of Advanced Technologies.* Oxford. 2000. University Press. 360 p.
13. *W. Frawley and G. Piatetsky-Shapiro and C. Matheus,* Knowledge Discovery in Databases: An Overview. AI Magazine. Fall, 1992. P. 213–228.
14. *D. Hand, H. Mannila, P. Smyth:* Principles of Data Mining. MIT Press, Cambridge, MA, 2001.
15. *Цветков В.Я.* Пространственные знания // Международный журнал прикладных и фундаментальных исследований. 2013. № 7. С. 43–47.
16. *Савиных В.П., Цветков В.Я.* Развитие методов искусственного интеллекта в геоинформатике // Транспорт Российской Федерации. 2010. № 5. С. 41–43.
17. *Кулагин В.П., Цветков В.Я.* Геознание: представление и лингвистические аспекты // Информационные технологии. 2013. № 12. С. 2–9.
18. *Винер К.* Кибернетика, или управление и связь в животном и машине. М.: Сов. радио, 1968.

Data and geodata mining

Vladimir Mihaylovich Markelov, aspirant, Moscow State University of Geodesy and Cartography

This article describes a new smart technology - GeoData Mining. The technology is the development of technologies known Data Mining. Describes the evolution of the concept of geodata. Article shows the difference between Data Mining technology and GeoData Mining. The article reveals the concept of GIS knowledge, spatial knowledge and geoknowledge. Paper describes the problem of intellectualization analysis of geodata.

Keywords: Earth science, geoinformatics, intelligent technology, geoznanie, spatial knowledge.

УДК 004.8+528.06

ГЕОИНФОРМАЦИОННЫЙ КОСМИЧЕСКИЙ МОНИТОРИНГ

Илья Андреевич Романов, соискатель,

E-mail: ir123456aa@yandex.ru,

Научно-исследовательский институт аэрокосмического мониторинга

«АЭРОКОСМОС»,

<http://www.aerocosmos.info>

Статья описывает новую интегрированную технологию – геоинформационный космический мониторинг. Эта технология является результатом интеграции геоинформационного и космического мониторинга. Описаны ее особенности и конкретные приложения.

Ключевые слова: космические исследования, геоинформатика, мониторинг, геоданные, пространственный анализ.

Введение

Термин «мониторинг» происходит от английского monitoring в его смысловом значении как контрольное наблюдение. Первое понятие мониторинга окружающей среды относят к 1972 г. Оно трактовалось как «система повторных наблюдений одного и более элементов окружающей природной среды в пространстве и времени с определенными целями с заранее составленной программой» [1]. Дальнейшее развитие теория монито-