

20. *Ozhereleva T.A.* Systematics for information units // *European Researcher*. 2014. Vol. (86). № 11/1. Pp. 1894–1900. DOI: 10.13187/er.2014.86. 1900.
21. *Шорыгин С.М.* Элементы языка визуального моделирования // *Славянский форум*. 2014. № 2 (6). С. 171–175.
22. *Цветков В.Я.* Язык информатики // *Успехи современного естествознания*. 2014. № 7. С. 129–133.
23. *Лютый А.А.* Язык карты: сущность, система, функции. М.: ГЕОС, 2002. 2-е изд. 2002. 327 с.
24. *Майоров А.А., Цветков В.Я.* Геореференция как применение пространственных отношений в геоинформатике // *Геодезия и аэрофотосъемка*. 2012. № 3. С. 87–89.
25. *Hill L.L.* Georeferencing: the geographic association of Information. 2009. Massachusetts Institut of Technology.
26. *Бахарева Н.А.* Виртуальность и реальность в геодезическом образовании // *Славянский форум*. 2015. № 1(7). С. 24–29.
27. *Майоров А.А.* Виртуальные модели при изучении логистики // *Славянский форум*. 2015. № 1(7). С. 169–176.
28. *Савиных В.П., Цветков В.Я.* Особенности интеграции геоинформационных технологий и технологий обработки данных дистанционного зондирования // *Информационные технологии*. 1999. № 10. С. 36–40.
29. *Розенберг И.Н., Цветков В.Я.* Космическая геоинформатика: учебное пособие. М.: МГУПС (МИИТ), 2015. 72 с.
30. *Ожерельева Т.А.* Об отношении понятий информационное пространство, информационное поле, информационная среда и семантическое окружение // *Международный журнал прикладных и фундаментальных исследований*. 2014. № 10. С. 21–24.
31. *Маркелов В.М.* Пространственная информация как фактор управления // *Государственный советник*. 2013. № 4. С. 34–38.
32. *Романов И.А.* Состояние пространственной экономики // *Славянский форум*. 2013. № 1(3). С. 110–115.
33. *Tsvetkov V.Ya.* Global Monitoring // *European Researcher*. 2012. Vol. (33). № 11-1. P. 1843–1851.
34. *Розенберг И.Н.* Геоинформационный мониторинг транспортных объектов // *НАУКИ О ЗЕМЛЕ*. 2012. № 3. С. 20–25.
35. *Маркелов В.М.* Геоинформатика и логистика. LAP LAMBERT Academic Publishing, Saarbrücken. Germany, 2015. 177 с.
36. *Розенберг И.Н., Цветков В.Я.* Координатные системы в геоинформатике. М.: МГУПС, 2009. 67 с.
37. *Савиных В.П., Цветков В.Я.* Развитие методов искусственного интеллекта в геоинформатике // *Транспорт Российской Федерации*. 2010. № 5. С. 41–43.

Geomarketing research

Anrdey Alexandrovich Maiorov, rector, Moscow State University of Geodesy and Cartography (MII-GAiK), Moscow, Russia.

Shows three main groups geomarketing research: operational, tactical and strategic.

Keywords: geoinformatics, marketing, geomarketing, GIS technology, GIS, spatial economics, information technology, information technology management, information resources.

УДК 004.9

ОБНОВЛЕНИЯ ПРОСТРАНСТВЕННОЙ ИНФОРМАЦИИ

*Виктор Яковлевич Цветков, профессор, д-р техн. наук,
зам. руководителя центра фундаментальных
и перспективных исследований НИИАС,
E-mail: cvj2@mail.ru*

*Научно-исследовательский и проектно-конструкторский институт информатизации,
автоматизации и связи на железнодорожном транспорте (ОАО «НИИАС»),
<http://www.vniias.ru>*

Статья анализирует методы обновления баз геоданных. Показаны особенности таких баз пространственной информации. Описаны основные подсистемы базы геоданных. Показано значение темпоральной логики для обновления пространственной информации.

Ключевые слова: информация, пространственная информация, базы данных, базы геоданных, обновление информации, обновление баз геоданных.



В.Я. Цветков

Введение. Пространственная информация хранится в базах геоданных [1–3], в банках пространственных данных [4], в хранилищах и в фондах инфраструктуры пространственных данных [5–7]. В общем случае можно говорить о системах хранения пространственной информации, включая в это понятие сайты и геопорталы [8]. Обновление пространственной информации является актуальной задачей в связи с ее оперативным изменением. В крупных городах ежегодно меняются не менее 1000 объектов, что также требует регулярного обновления пространственной информации. Одной из задач геоинформатики является получение новых знаний [9]. Базы данных пространственной информации хранят не только пространственную информацию, но и пространственные знания [10, 11] или геознания [12, 13].

Хранение пространственной информации

Пространственная информация имеет разные формы представления и организации [14]. Чаще всего это геоданные, которые создаются на основе интеграции, стратификации, и представляют собой системный информационный ресурс [15]. Этот ресурс позволяет проводить систематизацию и системный анализ пространственной информации в разных предметных областях.

Пространственная информация, организованная для хранения имеет вид пространственных информационных моделей [16, 17].

Системы хранения должны содержать топологическое описание объектов и часто имеют ассоциативную связь графической формы представления с табличной. По этой причине информация из БГД может быть визуализирована в виде карт или чертежей с помощью ГИС или САПР [18].

Поскольку геоданные являются наиболее общей и интегрированной информационной моделью в дальнейшем есть смысл говорить о базах геоданных (БГД), понимая под этим обобщенное понятие всех структурированных хранилищ систематизированной пространственной информации.

Классическое обновление информации. Как правило, информация в БГД – это описание какой-либо карты и должна соответствовать требованиям составления и представления такой карты. Процесс обновления пространственных данных сложен, так как он требует проверки корректности табличной и графической информации. Например, обновление части карты связано с анализом геометрических характеристик всех объектов карты, анализом объектов, которые частично пересекают обновляемую часть карты, а также отсутствием однозначной идентификации объекта в двух разных карт. Это приводит к тому, что технологически БГД должна включать ряд подсистем, которые не входят в состав реляционных баз данных. Это четыре основные подсистемы: контроля ввода, ввода, обновления, контроля качества информации.

Подсистема контроля вводимой информации. Данная подсистема контролирует ввод информации. Ее наличие обусловлено большими объемами пространственной информации, что затрудняет ее анализ человеком. Ввод информации выполняют при создании или обновлении базы данных. Основной задачей данной подсистемы является выявление некорректных объектов: по геометрии, по содержанию, по форматам ГБД.

Некорректным объектом по геометрии считается объект с нарушением топологии,

геометрии расположения или геометрии самого объекта. Например, возможны такие ошибки как: пересечения параллельных сторон; размещение части объекта на другом объекте; смещение объекта, деформация части объекта и т.д.

Некорректным объектом по семантике считается объект с нарушением дескриптивного описания, логического построения или с противоречивым описанием в отношении других объектов.

Некорректным объектом по формату считается объект, имеющий форматы, не соответствующие форматам базы данных. В этом случае говорят о нарушении информационного соответствия. Условие информационного соответствия требует соответствия форматов входной информации или первичных моделей форматам информационной системы, в которую эта информация вводится.

Данная подсистема используется при вводе и при обновлении информации в базе данных. По завершению работы она выводит информацию о корректных объектах и об объектах, загруженных с возможными ошибками. Например, при вводе сканерного изображения возможны аппаратные ошибки сканирования в этом случае система контроля вводя либо их исправляет, либо сигнализирует о возможных ошибках. Это подразумевает, что данная подсистема может иметь некую тестовую информацию для проверки качества.

Подсистема ввода выполняет предварительную обработку исходных данных и решает ряд задач. Подсистема ввода автоматизирует процесс ввода, что важно при больших объемах информации. Она осуществляет предобработку и загрузку данных различных форматов из допустимого перечня форматов. Технология ввода включает сжатие данных при их избыточности или дублировании. В этой подсистеме на когнитивном уровне происходит: создание новой карты, выбор темы в структуре ГБД; организация необходимых слоев; настройка соответствия геометрических типов исходных данных и ГБД; выбор атрибутивных данных; обеспечение соответствия координатных систем входной и внутренней.

В этой подсистеме информационное соответствие обеспечивается на трех задачах: соответствие координатных систем входной информации и внутренней информации ГИС или ГБД; соответствие типов данных входной информации слоям в проектируемой ГБД; соответствие типов данных входной информации базовым типам данных в ГБД.

Подсистема обновления. После создания ГБД данные в ней могут вводиться дополнительной или служить основой обновления уже введенных данных. Ввод дополнительной информации может приводить к структурным изменениям базы данных. В отличие от этого при обновлении структура ГБД не изменяется, данные интегрируются в существующие слои, типы объектов и таблицы.

Процесс простого обновления выполняется в два этапа и построен с использованием известной технологии репликации [19]. На первом этапе создается буферная ГБД на основе репликации основной ГБД.

Новые данные для обновления загружаются в буферную ГБД, где происходит их анализ и сравнение с существующими, а также проверка старых данных на актуальность. На основании проведенного анализа принимается решение об обновлении информации. Данные по обновляемому объекту могут быть загружены полностью либо частично. Механизм напоминает механизм резервного копирования в различных вариантах или механизм загрузки измененного файла. Как и при вводе данных включается механизм сравнения загружаемых и обновляемых данных на предмет корректности.

После проверки данных в реплицированной базе данных она загружается в основную базу данных и происходит известный процесс оверлея, то есть наложения информации с заменой старых частей на новые. Такая двухэтапная процедура позволяет сохранить целостность основной базы данных и решать задачи информационной безопасности по предотвращению удаления или порчи ценной информации.

Подсистема контроля качества. Данная подсистема предназначена для обеспечения

качества информации, хранимой в ГБД. Подсистема контроля выявляет и предотвращает дублирование объектов и связей. Она идентифицирует различные модификации объектов. Она также проверяет данные на актуальность и сигнализирует в случае необходимости обновления. Для решения последней задачи применяют темпоральную логику.

Темпоральное обновление пространственной информации (геоданных).

Темпоральное (временное) обновление основано на использовании темпоральной логики [21], темпоральной логики и темпоральных данных [22], темпоральных баз данных [23]. В этой технологии обновления первая концепция обновления геоданных строится на применении модальной и темпоральной логики. Вторая концепция обновления [24] строится на разделении системы логического обновления данных (ЛОГ) как некоей концептуально-формальной схемы и технологии или механизма обновления. При этом надо различать программу логического обновления данных и технологические программы обновления данных. Программа логического обновления данных является надстройкой и служит индикатором необходимости обновления данных. Технологические программы в сочетании с эвристическими методами реализуют собственно сам процесс ввода информации и ее актуализацию. Такой подход обладает универсальностью, поскольку логическую надстройку можно применять к любым базам данных, которые имеют разные структуры и разные данные.

Применение программы логического обновления геоданных в базах геоданных и геоинформационных системах предъявляет требования корректности и высокой верифицируемости к системам ЛОГ. Программирование ЛОГ представляет собой прикладную область, в которой опыт существующих исследований в области верификации программ [25] может быть успешно использован для обновления БГД.

Необходимо дать различие между программами логического обновления и стандартными вычислительными программами. Стандартная программа принимает на вход некоторые данные, выполняет необходимые вычисления и выдает результат. Подобную программу можно рассматривать как реализацию некоей функции. Главная функция классической программы состоит в трансформации пространства входных данных в пространство выходных данных, что также можно интерпретировать как перевод данных из исходного состояния в конечное состояние. В терминах состояний «классическая» программа переводит результаты вычислений или действий в целевое, конечное состояние. Доминантой в этом случае является достижение конечного состояния.

В отличие от стандартной программы реактивная программа не предполагает обязательного завершения вычислений. Такая программа «реагирует» на входные воздействия, отвечая на каждое воздействие в соответствии с заданным алгоритмом или системой правил. По существу такая программа выполняет сервисные функции. При этом в терминах состояний автоматная программа переводит данные из начального состояния в некоторое иное состояние, которое может быть конечным или промежуточным. Доминантной при логическом программировании являются переходы между состояниями и механизмы осуществления этих переходов.

Программа логического обновления геоданных (ПЛОГ) представляет собой конечный автомат [24], который использует множество состояний геоданных в базе данных и переходов между ними. Система логического обновления геоданных является «реагирующей» системой. Она может моделироваться как управляемый дискретный автомат, имеющий конечное множество входов (актуальности), подключенных связей актуализации к набору геоданных, и множество выходов, подключенных к технологическим программам обновления. ПЛОГ контролирует состояния входов и вырабатывает определенные последовательности программно заданных действий, отражающихся в изменении выходов. ПЛОГ предназначена для работы в режиме реального времени.

Темпоральные характеристики геоданных. Временные характеристики всегда входят в геоданные, однако форма их описания может быть различной: явной и неяв-

ной. Именно от описания геоданных зависит свойство темпоральность как возможность использования для геоданных темпоральной логики и темпоральных моделей. Можно определить темпоральность геоданных как свойство, заключающееся в явной форме выражения временных характеристик, делающих возможной применение теории темпоральных данных [26] и темпоральных баз данных [27].

Для построения темпоральных данных необходимо ввести понятие «тип времени». Типом времени, который рассматривается в темпоральных БД, является время операции (транзакционное время). В СУБД процедуре записи в БД можно сопоставить промежуток времени обновления – между моментами добавления записи и ее удаления из БД. При этом операция обновления понимается как составная операция удаления старой записи и добавления новой.

Говоря о типах времени геоданных, необходимо ввести понятие гранулярности времени. Гранулярность времени [12] показывает, насколько близкие моменты на оси времени все еще будут отличимыми друг от друга. В оптических пространственных данных это напоминает пространственное разрешение. Например, для данных о наблюдениях за деформациями инженерных сооружений достаточно использования интервалов времени фиксации по неделям или по 10 дням, а для времени мониторинга пожаров по данным дистанционного зондирования может быть использовано разбиение интервала наблюдения по часам. С каждым типом времени может быть связан некий дескриптор-календарь, который определяет диапазоны значений, гранулярность времени и прочие характеристики.

Второй особенностью является разделение интервальное и точечное представление данных или фактов. Оно напоминает нечеткое и четкое представление данных, однако, в нашем случае это только четкие значения. Точечное представление функции определяет ее значение в каждой точке, описываемой ее аргументами. Интервальное значение функции задается одинаковым для интервала аргументов. В случае точечного представления базы данных субъект получает срез всех фактов на какой-то конкретный момент времени. При интервальном представлении БД задается определенный факт и временные периоды его истинности. Для сравнения, обычная «четкая» реляционная модель опирается на точечное представление для актуального состояния данных.

Со временем фиксации факта и со временем операции связывается домен «Время», который может дискретным или непрерывным. Для представления времени в базе геоданных применяется домен с конечными и дискретными значениями. Как правило, значения времени в домене упорядочены. Момент времени актуализации геоданных в БД, является временем выполнения операции.

Для описания объектов модели геоданных может быть использована нотация методологии информационного проектирования [29]. В этой нотации сущность на ER-диаграмме представляется прямоугольником с именем в верхней части. В прямоугольнике перечисляются атрибуты сущности; при этом атрибуты, составляющие уникальный идентификатор сущности, подчеркиваются. Для атрибутов может быть задан домен. Такое представление геоданных не очень удобно для двумерной реляционной модели. Именно поэтому сдерживается применение темпоральных методов в геоинформатике.

Обработка временных данных требует критериев поиска по диапазону, что приводит к использованию соединений таблиц, которые строятся на отношениях неравенств. Механизм выполнения таких соединений является проблемой в реляционных БД. Кроме того, для темпоральных данных часто требуется соединять таблицы на основе перекрытия диапазонов дат, что приводит к нарушению условия однозначности. В языке SQL нет операции, позволяющей непосредственно задать такое соединение, и приходится применять процедурный подход к реализации таких запросов.

Таким образом, темпоральная модель геоданных отличается от традиционных моделей геоданных наличием формы представления времени, которое фиксируется в базе

геоданных с помощью так называемых временных меток (time stamp). Временные метки можно рассматривать как дополнительные атрибуты, связанные с фиксацией показаний времени. Исследование временных меток предметной области является одной из главных задач при обновлении геоданных, поскольку временные метки могут иметь несколько значений, т.е. неоднозначную интерпретацию в рамках предметной области.

При обновлении геоданных применяют следующие интерпретации временных меток. Время фиксации события или факта (Valid time), т.е. так называемая действительная дата – это временная метка, которая представляет время события или состояния предметной области как уже было указано в предыдущем разделе. Например, это дата подписи на документе, показания времени, снятые с контролирующих приборов, дата отгрузки проданного товара со склада и т.д.

Время операции (Transaction time) – это временная метка, представляющая время, когда была выполнена операция хозяйственной деятельности организации, как уже было указано в предыдущем разделе. Время сбора данных (Capture time) – это временная метка, представляющая время, когда данные были извлечены или собраны из источника данных (возможно, некоторой внешней БД). Время актуализации данных (Apply time) – это временная метка, связанная со временем загрузки данных в БГД.

Время, определяемое пользователем (User-defined time) – это временная метка, представляющая момент или моменты времени, которые пользователь намерен хранить в атрибуте сущности, но непосредственно не связанные с фиксацией временной зависимости в модели данных.

Таким образом, данный механизм позволяет реализовать технологию обновления, основанную на автоматическом анализе актуальности пространственной информации.

Заключение. Технология обновления пространственной информации связана с технологиями организации пространственной информации с построением пространственных моделей, построением классических реляционных баз данных, построением ассоциативной связи графики и табличной информации, с проведением топологического и семантического анализа. Технология обновления пространственной информации может быть связана с темпоральной логикой и разбиением программы обновления на логическую индикаторную часть и технологическую реализующую часть. Кроме рассмотренных в данной статье подходов для обновления пространственной информации можно применять интеллектуальные процедуры [30] и процедуры с использованием паттернов [4, 18] как особой формы хранения визуальной информации, отражающей пространственную информацию в виде электронной карты.

Литература

1. *Розенберг И.Н.* Геоинформационные базы данных в информационном обеспечении центров управления перевозками МПС: сб. докладов: Информационные технологии на железнодорожном транспорте «ИНФОТРАНС – 2001». Сочи, 2001. С. 170–176.
2. *Ванин А.Г., Цветков В.Я.* Средства организации баз данных в автоматизированных фотограмметрических технологиях // Геодезия и картография. 1985. № 7. С. 35–37.
3. *Розенберг Е.И., Розенберг И.Н., Сазонов Н.В., Цветков В.Я.* Система формирования мобильной базы данных для обслуживания участков железной дороги. Патент на полезную модель. RU93189U1. Заявка: 2009144856/22, срок действия с 03.12.2009.
4. *Цветков В.Я., Лобанов А.А., Матчин В.Т., Железняков В.А.* Обновление банков данных пространственной информации // Информатизация образования и науки. 2015. № 1 (25). С. 128–136.
5. *Майоров А.А.* Новые системы хранения пространственной информации // Перспективы науки и образования. 2013. № 5. С. 25–31.
6. *Савиных В.П., Соловьев И.В., Цветков В.Я.* Развитие национальной инфраструктуры пространственных данных на основе развития картографо-геодезического фонда Российской Федерации // Геодезия и аэрофотосъемка. 2011. № 5. С. 85–91.
7. *Осокин С.А.* Геопортал «Дата+» // Пространственные данные. 2008. № 2. С. 26.
8. *Матчин В.Т.* Состояние и развитие инфраструктуры пространственных данных // Образовательные ресурсы и технологии. 2015. № 1(9). С. 137–144.
9. *Hill Linda L.* Georeferencing: The Geographic Associations of Information – MIT Press Cambridge, Massachusetts, London, England, 2009. 272 p.
10. *Цветков В.Я.* Пространственные знания // Международный журнал прикладных и фунда-

ментальных исследований. 2013. № 7. С. 43–47.

11. *Benjamin Kuipers*. Modeling Spatial Knowledge (1978) // *Cognitive Science*. № 2. P. 129–153.

12. *Кулагин В.П., Цветков В.Я.* Геознание: представление и лингвистические аспекты // Информационные технологии. 2013. № 12. С. 2–9.

13. *Розенберг И.Н., Вознесенская М.Е.* Геознания и геореференция // Вестник Московского государственного областного педагогического университета. 2010. № 2. С. 116–118.

14. *Бородко А.В.* и др. Геодезия, картография, геоинформатика, кадастр: энциклопедия; в 2 т. / А.В. Бородко, Л.М. Бугаевский, Т.В. Верещака, Л.А. Запругаева, Л.Г. Иванова, Ю.Ф. Книжников, В.П. Савиных, А.И. Спиридонов, В.Н. Филатов, В.Я. Цветков. М.: Картоцентр-геодезиздат, 2008. Т. I А-М.

15. *Савиных В.П., Цветков В.Я.* Геоданные как системный информационный ресурс // Вестник Российской Академии Наук, 2014. Т. 84. № 9. С. 826–829. DOI: 10.7868/S0869587314090278.

16. ГОСТ Р. 52055-2003 Геоинформационное картографирование. Пространственные модели местности. М.: Госстандарт, 2003.

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

18. *Соловьев И.В., Железняков В.А.* Алгоритмы обновления электронных карт по данным дистанционного зондирования Земли // Геодезия и картография. 2014. № 4. С. 13–18.

19. *Карави А.С.* Модели тиражирования библиографических баз данных // Формирование современной информационно-библиотечной среды: сб. науч. тр. Новосибирск: ГПНТБ СОРАН, 2004. С. 180–197.

20. *Васильева К.А., Кузьмин Е.В.* Верификация автоматных программ с использованием LTL // Моделирование и анализ информационных систем. 2007. Т. 14. № 1. С. 31–43.

21. *Цветков В.Я.* Применение темпоральной логики для обновления информационных конструкций // Славянский форум. 2015. № 1(7). С. 286–292.

22. *Bowen K.A.* Model theory for modal logic: Kripke models for modal predicate calculi. Kluwer Academic Pub, 1979. Т. 127.

23. *Clarke E.M., Emerson E.A., Sistla A.P.* Automatic verification of finite-state concurrent systems using temporal logic specifications // *ACM Transactions on Programming Languages and Systems (TOPLAS)*. 1986. Т. 8. № 2. С. 244–263.

24. *Цветков В.Я., Матчин В.Т.* Концептуальное построение программы обновления геоданных // Перспективы науки и образования. 2015. № 1. С. 30–38.

25. *Вельдер С.Э., Шальто А.А.* Верификация автоматных моделей методом редуцированного графа переходов // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2009. № 6(64). С. 66–73.

26. *Date C. J., Darwen H., Lorentzos N. A.* Temporal data and the relational model: a detailed investigation into the application of interval and relation theory to the problem of temporal database management. Morgan Kaufmann, 2003.

27. *Костенко Б.Б., Кузнецов С.Д.* История и актуальные проблемы темпоральных баз данных // Труды Института системного программирования РАН. 2007. Т. 13. № 2.

28. *Суворова А.В., Пащенко А.Е., Тулупьева Т.В.* Оценка характеристик сверхкороткого временного ряда по гранулярным данным о рекордных интервалах между событиями // Труды СПИ-ИРАН. 2010. Т. 1. № 12. С. 170–181.

29. *Соловьев И.В., Майоров А.А.* Проектирование информационных систем. Фундаментальный курс: учеб. пособие для высшей школы / под ред. В.П. Савиных. М.: Академический проект, 2009. 398 с.

30. *Цветков В.Я., Железняков В.А.* Интеллектуальное обновление данных в банке данных земель сельскохозяйственного назначения // НАУКИ О ЗЕМЛЕ. 2012. № 2. С. 73–79.

Principles for updates geodatabases

Viktor Yakovlevich Tsvetkov, Professor, Doctor of Technical Sciences.

Center fundamental and advanced research, the deputy head.

Research and Design Institute of design information, automation and communication on railway transport, Moscow

The article analyzes the methods of updating geodatabases article describes the features of such databases of spatial information. This article describes the major subsystems of the geodatabase. The article shows the importance of temporal logic to update the spatial information.

Keywords: information, spatial information, databases, geodatabases, updating of information, update a Geodatabase.