
УДК 528.9:681.3

ББК 26.17

В 18

Т.П. Варшанина, О.А. Плисенко

Проблемы технологии автоматизированной географической диагностики территории

(Рецензирована)

Аннотация:

Обсуждаются преимущества применения объектно-ориентированной технологии и интеллектуальных систем для автоматизированной географической диагностики территорий, а также некоторые проблемы и задачи их разработки.

Ключевые слова:

Географическая диагностика, объектно-ориентированное проектирование географических систем, интеллектуальные системы геодиагностики

Объектом эколого-ресурсных исследований или геодиагностики территории являются географические системы, познание закономерностей структуры и организации которых, производится с позиции системного анализа. В качестве средства автоматизированного представления и анализа территориальных данных и знаний рассматривают Географические информационные системы (ГИС). Проектирование и организация данных в ГИС территорий традиционно осуществляется на основе методологических подходов картографии. Известно, что основанные на них ГИС обладают ограниченными возможностями системного накопления большого количества территориальных данных. На платформе картографических ГИС затруднительно моделирование широкого спектра процессов в трехмерном географическом пространстве путем имитационного воспроизведения. Такое положение приводит к тому, что, как правило, ГИС-модели регионов или ориентированы на конкретную достаточно узкую область применения, или содержат банк формально связанных моделей геосистем, отвечающих различным методологическим подходам.

На настоящее время накоплено колоссальное количество геопространственной информации, в связи с чем, возникла острая необходимость формирования методологии ее эффективной обработки путем логической структуризации в рамках системной территориальной

модели, основанной на фундаментальных знаниях географической науки. Накопление знаний в соответствие с логической структурой организации геосистем, извлечение новых знаний на основе анализа всего массива данных и исследования различных моделей предметной области представляет наиболее приоритетное направление развития автоматизированных средств поддержки географических исследований. Вследствие высокой сложности территориальных систем в процессе перехода от накопления данных к их анализу, и на его основе к выявлению новых связей и закономерностей (новых знаний), неизбежно использование методов искусственного интеллекта, таких как нейронные сети и экспертные системы, которые позволяют обобщить существующие знания и генерировать новые. Разработка инструментальных средств ориентированных на обработку знаний, включающих перечисленные методы и модель фундаментальных знаний географической науки, позволяет поддерживать необходимый уровень исследований состояния и прогноза развития природно-хозяйственных систем.

Природно-хозяйственные системы (ПХС), принадлежащие к классу сложных систем, в соответствие с универсальным принципом организации материи являются иерархически структурированными системами. Применительно к ПХС это означает, что их вертикальная и горизонтальная структура складывается из

множества взаимозависимых подсистем, которые в свою очередь также могут быть разделены на составляющие подсистемы вплоть до самого низкого уровня. В настоящее время геоинформационное проектирование сложных, очень больших систем позволяет осуществлять единственную методологию – объектно-ориентированный подход, изначально разработанный для проектирования технических систем, в том числе программных продуктов. Объектно-ориентированная технология отличается следующей идеологией, в настоящее время, в частности, реализованной в системах управления базами данных (Буч, 1999).

Объектно-ориентированное проектирование соединяет процессы объектной декомпозиции и приемы представления логической и физической, а также статической и динамической моделей проектируемой системы.

При программном проектировании сложной системы ее разделяют на все меньшие и меньшие подсистемы, каждая из которых при необходимости может быть модифицирована вне зависимости от других.

Элементарные единицы системы соответствуют объектам, выделяемым по их инвариантным характеристикам, затем объединяемым в классы объектов и далее в типы классов. Так формируются структуры иерархии объектов и классов, т.е. логическая архитектура системы. Сложные программные системы включают и другие типы иерархии, верхний уровень которой образуют иерархия модулей, описывающая отношения между физическими компонентами системы, и иерархия процессов, описывающая отношения между динамическими компонентами.

Структуры классов и объектов являются взаимозависимыми, так как каждый элемент структуры объектов представляет специфический экземпляр определенного класса. Относительно структуры классов, выделяемых по их инвариантным параметрам, размещаются общие свойства экземпляров, что оптимизирует организацию логической архитектуры большой системы и значительно уменьшает ее громоздкость.

Следует подчеркнуть следующие основные положения:

- В объектно-ориентированной системе каждый объект (экземпляр класса) идентифици-

руется по его инвариантным свойствам, отличающим его от других объектов

- Состояние объекта характеризуется его свойствами и динамическими значениями каждого из этих свойств (определяется его статическими и динамическими параметрами).

- Поведение объекта характеризуется изменением его состояния в процессе взаимодействия с другими объектами.

- Отношения объектов основываются на операциях, которые предполагается выполнять с объектами, и на их ожидаемом поведении.

- Существенным обстоятельством является то, что отношения между объектами и классами упорядочиваются в виде стабильных промежуточных описаний, что позволяет при необходимости их редактировать.

В объектно-ориентированном анализе и проектировании представлены два типа иерархических соотношений объектов: связи и агрегация. Связи обозначают равноправные и «клиент-серверные» отношения между объектами. Агрегация описывает отношения целого и части, обуславливающие соответствующую иерархию объектов. Объект, являющийся атрибутом другого объекта (агрегата), имеет связь со своим агрегатом. Через эту связь агрегат может посылать ему сообщения.

Набор объектов, обладающих общей структурой и одинаковым поведением, представляет класс. Класс – хранит только общие свойства своих экземпляров. Между классами важны три основных типа отношений «обобщения / специализации»: «один-к-одному», «один-ко-многим», «многие-ко-многим». Структура, объединяющая множество объектов, обеспечивает их совместное целенаправленное функционирование.

Иерархия классов подразделяется на шесть типов, соответствующих основным видам взаимодействия объектов высокого уровня: ассоциирование, наследование / делегирование, агрегация, использование, инстанцирование и метаклассирование.

Определение классов и объектов – классификация, является одной из самых сложных задач объектно-ориентированного проектирования. Процесс классификации позволяет выявить инвариантные свойства объектов, помогает найти ключевые абстракции и механизмы, определить логику архитектуры системы и ее

модульной структуры. Выявление иерархии объектов системы рекомендуется производить в соответствии с подходами, сложившимися в каждой конкретной предметной области.

Основными достоинствами метода объектно-ориентированного проектирования является:

- воссоздание естественной иерархии объектов и процессов взаимодействия между ними, обеспечивающее функционирование системы как единого целого;

- эволюционное, помодульное проектирование системы;

- возможность модификации в соответствии с развитием представлений в каждой предметной области;

- оптимальное структурирование большого объема атрибутивных данных.

Географические принципы структурирования территориальной информации находят полное соответствие в идеологии объектно-ориентированного проектирования, и это не удивительно, так как логику ее организации авторы метода сверяли с логикой организации природных систем.

Анализ области решения проблемы объектно-ориентированного проектирования географических систем выявил, что существующие геоинформационные системы поддерживают модели данных на разных уровнях абстракции (логическом и физическом). Однако они зачастую описываются под общим названием, что вносит путаницу при описании класса геоинформационных систем. Такое положение сохранилось с момента возникновения ГИС, когда логический и физический уровень практически не разделялся. Примером может служить модель плоского файла, которая представляет собой хранение пространственной информации в виде растрового изображения с включенными в него атрибутами. Для такого типа моделей существуют два способа включения атрибутивной информации. Простейшим является присваивание значения атрибута каждой ячейке растра. Другим подходом, а на самом деле расширением только что описанного, является связывание каждой ячейки растра с файловой базой данных. Т.е., в такой схеме отсутствует концептуальная модель данных, не зависящая от конкретных методов их хранения, что не по-

зволяет разрабатывать модели реальных объектов геопространства.

Более высоким уровнем абстракции обладают геореляционные модели данных, в которых можно выделить концептуальный и физический уровни моделирования. Для ГИС, использующих такие модели, проблемой является кооперирование пространственной и буквенно-цифровой информации. В наиболее распространенных из них управление хранением данных выполняется двумя подсистемами: реляционной СУБД, поддерживающей атрибутивные данные; системой, обрабатывающей пространственные данные. Разработку такой модели позволяет автоматизировать аппарат реляционной алгебры. Вследствие нерешенной проблемы соединения буквенно-цифровых и пространственных данных такие системы не обеспечивают построения сложных семантических моделей, приближенных к реальности.

Дальнейшим развитием в области моделирования данных стала объектно-ориентированная технология, позволяющая рассматривать предметную область в терминах объектов, которые существуют, взаимодействуют друг с другом и могут реализовывать отношения наблюдения.

Следует отметить, что термин «объектно-ориентированная ГИС» пользователями и разработчиками трактуется по-разному, что также вносит путаницу и дает возможность некоторым компаниям, отдавая дань моде называть свои продукты объектно-ориентированными, в то время как они не соответствуют этой идеологии.

Можно выделить пять групп систем, в той или иной мере имеющих отношение к объектно-ориентированному проектированию:

1. Программные продукты, разработанные на основе объектно-ориентированных языков программирования, но не работающие с объектно-ориентированными моделями данных предметной области. К ним относятся многие из современных ГИС-продуктов.

2. Программные продукты, использующие объектно-реляционную модель данных поддерживаемую СУБД. Класс таких моделей является промежуточным между чисто реляционными и объектно-ориентированными моделями. В основном рассматриваются объектные расширения к реляционной схеме данных, на-

пример, такие как вложение данных и присоединение процедур к реляционным полям. Вложение данных позволяет управлять векторными данными (например, значением одного вложенного атрибута может быть набор координат, представляющий полигон). Примером такой гибридной архитектуры являются ранние версии СУБД POSTGRESS.

3. Системы, использующие объектно-ориентированную модель данных, поддерживаемую СУБД. Такие системы редки и, как правило, они являются узкоспециализированными, ориентированными на конкретную предметную область. К тому же, удобный пользовательский интерфейс, позволяющий производить объектное моделирование непосредственно пользователем, пока не реализован.

4. Системы, использующие объектно-ориентированный интерфейс для работы с данными (например, ArcGis 9.1). Разработчики таких систем пытаются убедить пользователя в том, что он работает с объектами, например, с геометрическими объектами, такими как водотоки, однако не предоставляют все возможности объектно-ориентированной технологии, а именно: задание пользователем методов (а не ограничений!) поведения объекта, реализацию многоуровневого и множественного наследования.

5. Полностью объектно-ориентированные ГИС, поддерживающие эту технологию как на уровне процесса разработки интерфейса, так и на уровне моделирования данных.

Большинство промышленных ГИС можно отнести к различным комбинациям первых 3-х классов (например, ArcGIS 9, ГИС GRASS, отечественные «Панорама», ObjectLand, GeoDraw и т.д.). Что касается пятого класса, то таких систем очень мало, и все они являются исследовательскими не коммерческими разработками. Одной из лучших разработок такого класса является ГИС GOCAD, созданная в Боннском университете для моделирования геологической среды (стратиграфии, тектоники), на основе ядра GeoToolKit, работающего с пространственной СУОБД GeoStore. ГИС GOCAD поддерживает пространственно-временные модели данных (т.е. 3D/4D), а также три типа операций над ними: топологические, геометрические, семантические. В ней реализована кинематическая тектоническая модель.

Объектно-ориентированная модель применялась в этой разработке и для проектирования физической модели данных (Breunig, M., Creemers, A.B., Muller, W., Siebeck, J. 2001; Shumilov, S., Siebeck, J. 2001).

Для определения спецификаций на различные компоненты объектно-ориентированной технологии в 1989г. был создан международный, некоммерческий консорциум, объединяющий основных производителей программного обеспечения, а также представителей основных групп пользователей, OMG (Object Management Group). Разработанные этим консорциумом, архитектура CORBA(Common Object Request Broker Architecture), поддерживающая объектно-ориентированную технологию на уровне промежуточного программного обеспечения, а также набор стандартов на методы и средства объектного моделирования данных (Unified Modeling Language – UML (унифицированный язык моделирования), Model Driven Architecture – MDA (Архитектура управляемая моделью), и объектных систем управления базами данных (Object Definition Language – ODL (Язык управления объектами), Interface Definition Language – IDL (Язык определения интерфейсов), Object Query Language – OQL (Язык объектных запросов)) позволяют создавать многофункциональные, совместимые объектно-ориентированные программные продукты.

На основе анализа существующих подходов к моделированию геопространства, можно сделать вывод, что для этого случая наиболее эффективным решением будет комбинированный объектно-реляционный подход. Этот подход позволяет моделировать сложную иерархическую систему сущностей предметной области в виде объектно-ориентированной схемы базы данных, а точечные характеристики, являющиеся проявлением влияния различного рода полей (например, климатические параметры) – в виде реляционных таблиц поддерживающих непрерывный массив данных таких параметров в различные интервалы времени. Таким путем в территориальных моделях возможно воспроизведение двуединства дискретно-континуального свойства геопространства.

С помощью объектно-ориентированной технологии решается проблема объединения пространственных и буквенно-цифровых дан-

ных на концептуальном уровне. Осуществляется этот подход путем включения в структуру СУБД средств работы с пространственными данными в виде специальных модулей, предоставляющих набор базовых пространственных классов и функций, определяющих как геометрические, так и топологические операции над этими объектами. В настоящее время появились СУОРБД, обеспечивающие хранение, доступ и анализ не только атрибутивных, но и пространственных данных, например, Oracle с модулем Spatial. Такие СУОРБД позволяют разрабатывать ГИС нового поколения в виде многофункциональных интегрированных территориальных банков данных, адаптированных к решению задач управления регионами и планированию рационального природопользования.

Для двумерного представления пространства такая концепция оказалась удачной и дала большой выигрыш и в скорости обработки информации, и в более компактных методах хранения, и в обеспечении многопользовательского доступа. Кроме того, визуализировать такие модели также не представляет трудности.

Другая ситуация наблюдается при построении 3D/4D моделей геопространства. Сложность представления пространственных объектов, а также различных видов отношений между ними резко возрастает, да и до настоящего времени не предлагались адекватные методы визуализации трехмерной поверхности рельефа с помощью пространственных объектов. Большинство существующих методов построения рельефа используют континуальный подход, рассматривая рельеф как поле высот (модель DEM). TIN модели выделяют объекты, но эти объекты являются геометрическими примитивами, не связанными с действительными элементарными поверхностями, образующими сложную систему рельефа земной поверхности. К тому же они являются плоскими и строятся в 2D пространстве. Известны попытки для растровых систем в качестве элементарного объекта, образующего рельеф, рассматривать воксел (трехмерный аналог пиксела). Однако воксел также представляет собой отвлеченный геометрический объект не являющийся структурообразующим для модели рельефа. На данный момент, создание трехмерной объектной модели рельефа является одной из самых важ-

ных задач, т.к. рельеф является базовой основой для моделирования различных процессов в геопространстве.

Существуют проблемы и в области разработки интеллектуальных средств обработки географических данных. Информационные интеллектуальные системы на данный момент представляют собой гибридные системы, позволяющие использовать преимущества нейросетевого подхода (быстрые параллельные вычисления, извлечение знаний из эмпирических данных, возможности самообучения и совершенствования) и традиционные экспертные системы (четкая структура причинно-следственных связей, возможность объяснения сделанного заключения, обработка символьных данных). Гибридные интеллектуальные системы позволяют более эффективно соединять формализуемые и неформализуемые знания. На данный момент в ГИС методы искусственного интеллекта применяются для решения только отдельных задач, таких как: обработка видеоизображений; преобразование растровых изображений в векторные графические модели; обработка картографической информации; обработка разнородной информации; анализ моделей ГИС; получение решений на основе геоинформации, а также в специализированных областях, например при обработке космоснимков.

Существо проблемы создания интеллектуальных систем геодиагностики заключается в отсутствии разработанной структуры базы фундаментальных знаний в области географии, являющейся интегрированной основой для интеллектуальных моделей и моделей знаний. Воссоздание логической структуры фундаментальных географических знаний территории возможно в виде объектно-реляционной геосистемной территориальной модели.

Основанием для реализации геосистемной территориальной модели являются: декомпозиция (районирование) предметной области по иерархии взаимосвязанных реальных геообъектов в результате исследования общего поведения природно-хозяйственной системы; выявление геообъектов, определение их параметрических и описательных характеристик, формирующих базу знаний региона. В географии разработаны принципы классификации / районирования иерархии объектов, как по вертикаль-

ной, так и по горизонтальной составляющим геосистем. Процедура выделения иерархии географических районов как базовых пространственных единиц территориальной модели должна включать следующие этапы: 1) установление значимых признаков и градаций их параметров, которые имеют значения, отличающие район от соседних районов; 2) составление классификации признаков; 3) проведение границ на карте.

Логическая схема базы данных состоит из следующих подсистем: литогенное основание с его рельефом, климат, гидроресурсы, почвы, биота, ландшафты, которые воссоздают иерархию объектов геокомпонентов и геосистем.

Конструирование моделей данных компонентных подсистем производится в результате: 1) пространственно-временного согласования (упорядочения данных в виртуальном территориальном пространстве во временных интервалах); 2) оптимизации структуры информации (создания подмножества параметров средоформирующего значения); 3) решения трендов созданного подмножества параметров; 4) районирования (выделения однородных в заданных интервалах параметров пространственных объектов требуемого масштабного ряда по каждому конкретному и комплексу средоформирующих показателей).

В процессе функционирования ГИС все многообразие входных данных – информация об объектах, их характеристиках, о формах и связях между объектами, различные описательные сведения – преобразуется в единую общую модель, хранимую в банке данных, представляющую собой информационное хранилище, отображающее структуру объектов реального мира и их взаимосвязи. Объекты предметной области отображаются в 3D/4D пространстве с учетом топологических, геометрических и моделируемых семантических особенностей поведения.

Описываемая концепция предоставляет возможность разработки логической структуры программного продукта, соответствующей логической структуре системы геопространства, в результате чего расширяется область применения продукта и круг организаций – пользователей. Пользователю обеспечиваются возможности модификации состава моделируемых объектов, разработки специализированных геоинформационных систем методом адаптации, на-

полнения их знаниями о конкретных типах объектов. Таким образом, система переходит в класс полноценных экспертных систем, основанных на знаниях.

Структура геоинформационной интеллектуальной системы региона обеспечивает интеграцию геоинформационной системы, экспертной системы и нейронных сетей. Ядром системы является пространственно-временная база тематических данных, построенная на основе объектно-ориентированной технологии, и база знаний региона, соединяющая как формализуемые знания (логико-лингвистические модели, символьные вычисления), так и неформализуемые знания (в нейронных сетях).

Объекты моделируемой предметной области описываются иерархической схемой классов базы данных. Основным типом связи между классами является агрегирование. Т.е. объекты более высокого уровня иерархии объединяют в своей структуре объекты более низкого уровня. При этом классы более высокого уровня наследуют свойства и модели поведения, входящих в них классов, как правило, на основе нелинейной интегральной зависимости. Большое количество классов не позволяет описывать такую зависимость явно, используя стандартные методы. Кроме того, такая зависимость часто является не определенной и не исследованной, а потому определение ее вида и характера является отдельной задачей для экспертного оценивания. Решение данной задачи для каждого из классов предлагается с использованием нейросетевых методов, которые позволяют на основе эмпирических данных в неявном виде выделить эту зависимость и использовать для моделирования взаимодействия между классами разного уровня иерархии. Способность нейронной сети к обучению позволяет «доработать» эвристически форму зависимости для каждой обрабатываемой модели, для каждой конкретной задачи, при обучении сети на полученных опытным путем данных. Моделирование семантических зависимостей основанных на символических данных также выполняется с использованием нейронной сети. На основе сформулированных пользователем / экспертом правил, имеющих вид продукций, формируется нейронная сеть (структура, значение весов, увеличение или уменьшение количества значимых связей и т.д.)

Таким образом, происходит накопление знаний в структуре нейронной сети.

Средства работы с нейронными сетями тесно интегрированы с экспертной системой, функциями которой является формирование оптимальной нейронной сети на основе заданных продукций, а также извлечение новых знаний из нейронной сети и формирование объяснений полученных заключений при анализе данных. Для формирования оптимальной структуры нейросети используются алгоритмы эволюционных вычислений. Для извлечения знаний и формирования объяснений на естественном языке используются метод индукционных деревьев и разрабатываемые авторские алгоритмы.

Графический интерфейс для построения модели предметной области (конструктор модели исследования), позволяет пользователю строить свою иерархию классов и определять поведение, связи каждого класса явно или косвенно, на основе интегрированной библиотеки пространственно-временных классов (3D/4D измерения) и механизма нейронных сетей. Построенная таким образом модель содержит знания пользователя/эксперта о предмете исследования в числовом и символическом виде.

Таким образом, база знаний разрабатываемой системы содержит знание о конкретной прикладной задаче смоделированной экспертом. Знания представляются в виде гибридной нейронной сети, которую обучает эксперт на основе, как общих знаний, так и информации о частных случаях, известных эксперту.

ГИС обрабатывается на платформе СУОРБД Oracle 9i. Банк данных, поддерживающий общую концепцию системы, предназначен для хранения информации, организации ее обработки и основывается на системе взаи-

мосвязанных геоинформационных моделей данных подсистем, представленных иерархией пространственных объектов, относительно которых происходит – накопление / мониторинг тематической информации. Картографическая реализация результатов анализа координированных данных осуществляется средствами ArcGIS 9.1.

Ядром Территориального банка данных является информационный сервер, на платформе которого создаются банк полномасштабного ряда электронной картографической информации территории, банк базовых тематических данных по рельефу, гидрографии, почвам, климату, биоте, ландшафтами, общественно-хозяйственной сфере и база знаний региона. Информационный сервер играет роль интегратора отраслевых ГИС на единой основе тематических и топографических материалов, что обеспечивает совместимость сведений, аккумулируемых в отраслевых базах данных. Основная задача такой платформы обеспечить эффективное взаимодействие между данными различных тематических блоков, комплексный совмещенный анализ, а также работу пользователей с данными в условиях разграниченного доступа.

Такая ГИС является инструментом, обеспечивающим как геоинформационную поддержку исследований географических закономерностей организации природы и человеческой деятельности региона, так и является банком территориальных данных для геодиагностики в целях текущего управления и планирования стратегии рационального природопользования с возможностями экспертной многокритериальной оценки.