

ТИПОВЫЕ ОШИБКИ В ИНФОРМАЦИОННЫХ РЕСУРСАХ ГИС, ПРАВИЛА И ОСОБЕННОСТИ ИХ ВЫЯВЛЕНИЯ В СФЕРЕ ЛЕСНОГО ХОЗЯЙСТВА (НА ПРИМЕРЕ СИСТЕМЫ «ГИС-ЛЕСПРОЕКТ»)

*Институт проблем математических машин и систем НАН Украины, Киев, Украина

Анотація. Пропонується класифікація типових правил несуперечності компонентів інформаційних ресурсів ГИС. Визначено набір правил перевірки топологічних залежностей картографічної БД лісовпорядкування. Проведено аналіз можливостей виявлення прихованих помилок інформаційних ресурсів ГИС у сфері лісового господарства та визначено напрям розвитку спеціалізованого модуля діагностики помилок на прикладі системи «ГИС-Ліспроєкт».

Ключові слова: правила несуперечності інформації, геоінформаційні системи в лісоустрої, перевірка топології.

Аннотация. Предлагается классификация типовых правил непротиворечивости компонентов информационных ресурсов ГИС. Определен набор правил проверки топологических зависимостей картографической БД лесоустройства. Проведен анализ возможностей обнаружения скрытых ошибок информационных ресурсов ГИС в сфере лесного хозяйства и определено направление развития специализированного модуля диагностики ошибок на примере системы «ГИС-Леспроект».

Ключевые слова: правила непротиворечивости информации, геоинформационные системы в лесоустройстве, проверка топологии.

Abstract. A classification of standard rules of consistency of GIS information resources components is proposed. The set of validation rules of topological dependency of forest inventory database mapping is defined. Attention is given to analysis of detecting hidden errors of GIS information resources in forestry. The direction of development of the specialized module diagnostic error through the example of “GIS-Lisoproekt” system is argued. The analysis of the possibility of detecting hidden errors of GIS information resources in forestry was done and the direction of development of the specialized module diagnostic error through the example of “GIS-Lisoproekt” system was determined.

Keywords: rules of consistency of information, GIS in forestry, topology verification.

1. Введение

Анализ мирового опыта в области создания и внедрения информационных технологий, функционирования ситуационных центров органов государственной власти показывает, что географические данные составляют более половины информационных ресурсов (ИР), используемых для поддержки принятия решений. В связи с этим наблюдается постоянный рост требований к их качеству.

Качество ИР ГИС и, соответственно, решений, принимаемых с их использованием, в значительной степени зависит от наличия или отсутствия ошибок в пространственных и/или атрибутивных компонентах геоданных. Целью предлагаемой статьи являются попытка систематизации типовых правил непротиворечивости информации в ИР ГИС (с учетом имеющегося в настоящее время методологического и программного инструментария обнаружения ошибок), описание конкретных особенностей приложения этих правил к ИР «ГИС-Леспроект» [1, 2], анализ характера «скрытых» ошибок в данных сферы лесоустройства и возможностей их выявления.

2. Классификация ошибок в ИР ГИС

2.1. Общие положения

Пути решения проблемы предотвращения, обнаружения и исправления ошибок в ИР ГИС

начинаются с систематизации и классификации типовых ошибок. Адекватная проблемно-ориентированная классификация позволяет выделить во всем многообразии ошибок наиболее важные для конкретных приложений классы, типизировать их, выработать общие подходы к борьбе с ними и соответствующий инструментарий.

Можно выделить следующий ряд самостоятельных критериев, на основе которых возможно построение непротиворечивых классификаций ошибок:

- этапы технологического процесса создания, подготовки и сопровождения данных, в процессе которых возникают ошибки;
- непосредственные источники ошибок (физические, юридические);
- структурные компоненты данных (пространственные, тематические), искажаемые ошибками, и грамматический характер ошибок;
- общие свойства данных (степень неопределенности, скорость изменения, правовые и/или экономические последствия ошибок и др.);
- специфические проблемно-ориентированные критерии.

Так, например, кафедра кадастра и основ земельного права МИИГАиК (РФ) предлагает мультикритериальную классификацию ошибок при формировании кадастровых сведений о земельных участках, основанную на трех критериях: источники ошибок (юридические, технологические), компоненты данных, правовые последствия [3]. В принятой нами классификации, ориентированной на заключительные этапы создания и подготовки ИР ГИС, иерархическая структуризация типовых ошибок основана на критериях разделения пространственных и тематических компонентов данных по грамматическим видам и уровням организации контролируемых элементов с учетом имеющегося в настоящее время готового инструментария проверки корректности топологии геоданных ГИС. На этих этапах возможно возникновение следующих ошибок:

- ошибок в отдельных компонентах, не обнаруженных на ранних этапах их подготовки и формирования;
- ошибок, возникших при внесении текущих изменений в геоданные и атрибутивные компоненты;
- искажений (возможно и злоумышленных), внесенных в процессе использования ИР ГИС для принятия определенных решений.

Уточним несколько определений, связанных с понятиями «ИР ГИС», и «ошибки».

ИР ГИС – интегрированная база данных (ИБД), включающая как пространственные, так и тематические (атрибутивные) компоненты.

Абсолютная ошибка – несоответствие представленного в ИБД свойства объекта (значения атрибута сущности) истинному (реальному) значению.

Правило непротиворечивости данных ИБД (НП-правило) – необходимое условие истинности данных, охватываемых определенным правилом.

Относительная ошибка (противоречивость) – несоответствие представленных свойств (значений атрибутов) определенному НП-правилу.

В дальнейшем под «ошибкой» понимаются именно относительные ошибки, которые можно условно разделить на два грамматических вида:

- синтаксические ошибки, нарушающие самые общие НП-правила, относящиеся к формам представления данных и их сочетаний;
- семантические ошибки, нарушающие НП-правила, связанные со смысловым содержанием данных (свойствами объекта и/или значениями атрибутов).

Процесс обнаружения ошибок в ИБД заключается в проверке соблюдения синтаксических и семантических НП-правил для проверяемых данных и их сочетаний.

Синтаксические ошибки могут идентифицироваться и исправляться автоматически или «вручную». Идентификация и исправление семантических ошибок выполняются

вручную (как правило) с использованием дополнительных, предположительно истинных представлений данных по мажоритарному принципу (рис. 1).

Выделим следующие виды ресурсов, охватываемые НП-правилами:

- пространственные (ошибки топологии);
- атрибутивные (ошибки в значениях атрибутов и их сочетаниях);
- пространственно-атрибутивные (противоречия в значениях атрибутов и свойствах соответствующих объектов пространственных данных).

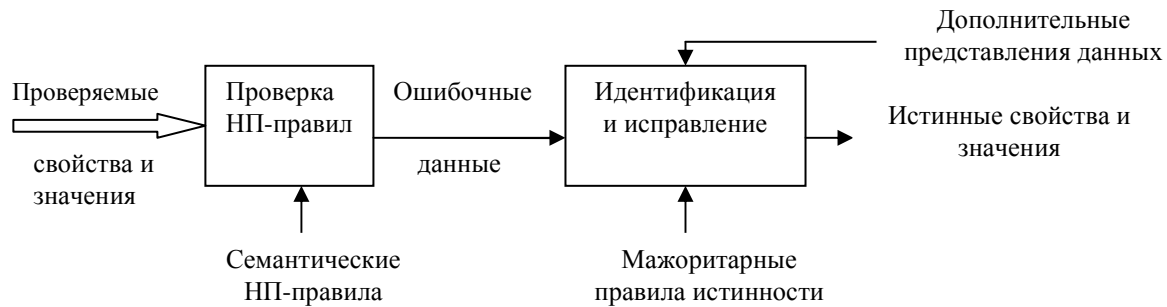


Рис. 1. Общая схема идентификации и исправления семантических ошибок

Верхние уровни классификационной схемы представлены на рис. 2.

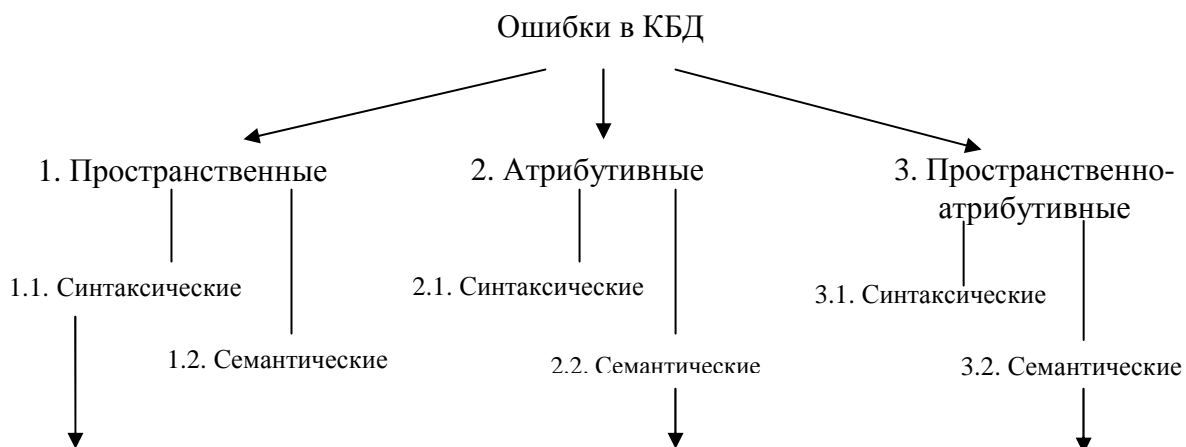


Рис. 2. Схема верхних уровней классификации

В контексте задачи общей классификации типовых ошибок в ИР ГИС наибольший интерес представляют ошибки 1.1, 2.2 (и отчасти 3.1) в связи с их многообразием, изученностью и наличием уже имеющегося методологического инструментария проверки соответствующих НП-правил. Поэтому дальнейшая декомпозиция приведенных на рис. 2 уровней и ветвей ориентирована именно на эти ошибки.

2.2. Пространственные синтаксические ошибки

Пространственные отношения связности и соседства векторных объектов – точек, линий и полигонов – в геоинформационных системах определяет топология. Топологические данные необходимы для обнаружения и исправления ошибок оцифровки (например, две линии дорог не сходятся на месте перекрестка и др.), проведения некоторых типов пространственного сетевого анализа (гидрологический, транспортный), формирования достоверной кадастровой документации в графической составляющей, при подготовке и тиражировании различных видов карт-материалов, а также в других сферах, где необходим учет пространственного размещения объектов [4].

Современные универсальные ГИС известных производителей США, Канады, России ..., такие как ArcGis, MapInfo, Панорама и другие, предоставляют обширные средства для выявления и корректировки топологических ошибок.

Одной из наиболее развитых систем в этом отношении является ArcGIS фирмы ESRI США и в основу классификации НП-правил топологии уместно положить систематизированный перечень общих топологических проверок арсенала ArcGIS [5] (табл. 1).

Таблица 1. НП-правила для топологии

НП-правило	Смысл правила	Пример
1.1.1. Полигональные объекты		
1.1.1.1. Не должны перекрываться	Полигоны не должны перекрываться внутри класса или подкласса объектов	На карте избирательных округов не должно быть никаких перекрытий
1.1.1.2. Не должны иметь пробелов	Полигоны одного класса или подкласса не должны иметь пустот внутри или между объектами	Полигоны почв не могут включать пустоты и иметь пробелы между полигонами
1.1.1.3. Не должны перекрываться с <...>	Полигоны одного класса или подкласса объектов не должны перекрываться полигонами другого класса или подкласса	Озера и земельные участки не должны перекрываться
1.1.1.4. Должны совмещаться с объектами класса <...>	Полигоны первого класса или подкласса объектов должны покрываться полигонами второго класса или подкласса объектов	Области должны быть покрыты районами
1.1.1.5. Должны совмещаться с <...>	Полигоны одного класса или подкласса объектов должны покрываться одним полигоном второго класса или подкласса объектов	Районы должны быть покрыты областями
1.1.1.6. Граница должна совпадать с <...>	Границы полигонов одного класса или подкласса объектов должны совпадать с линиями другого класса или подкласса объектов	Линии магистральных дорог образуют часть границ районов переписи
1.1.1.7. Граница площадного объекта должна совпадать с границей <...>	Границы полигонов одного класса или подкласса объектов должны совпадать с границами полигонов другого класса или подкласса объектов	Границы поселка везде совпадают с границами находящихся в нем участков, но совпадают не со всеми границами участков
• • •	• • •	• • •
1.1.1.11. Содержит точку	Полигоны одного класса пространственных объектов должны содержать хотя бы по одной точке из другого класса	Земельные участки должны содержать, по крайней мере, одну точку адресной привязки
1.1.2. Линейные объекты		
1.1.2.1. Не должны иметь висячих узлов	Конечные точки линии должны касаться любой части этой же линии или другой линии данного класса или подкласса объектов	Уличная сеть состоит из соединенных между собой линейных сегментов

1.1.2.2. Не должны иметь псевдоузлов	Конечная точка линии не может касаться конечной точки только одной линии того же класса или подкласса объектов. Конечная точка линии может касаться любой частью этой линии	В целях гидрологического анализа сегменты речной сети могут соединяться так, чтобы узлы создавались только на конечных точках или соединениях
1.1.2.3. Не должны перекрывать сами себя	Линии одного класса или подкласса объектов не должны перекрывать сами себя. Линии могут касаться, пересекать или перекрывать другие линии	В целях транспортного анализа объекты улиц и магистралей не должны налагаться сами на себя
1.1.2.4. Не должны перекрываться	Линии не должны перекрываться с любой частью другой линии этого же класса или подкласса объектов. Линии могут касаться, пересекаться друг с другом или перекрывать сами себя	Линии границ для групп участков не должны перекрываться
...
1.1.2.16. Должны состоять из одной части	Каждая линия должна состоять из одной части	Система магистралей состоит из отдельных объектов, каждый из которых состоит не более чем из одной части
1.1.3. Точечные объекты		
1.1.3.1. Должны лежать на границе	Точки одного класса или подкласса объектов должны лежать на границе полигонов другого класса или подкласса объектов	Может потребоваться, чтобы станции обслуживания инженерных коммуникаций были на границе участков
1.1.3.2. Точки должны лежать на линиях	Точки одного класса или подтипа объектов должны лежать на линиях другого класса или подкласса объектов	Станции мониторинга должны располагаться вдоль потоков
...
1.1.3.6. Должны быть полностью внутри	Точечные объекты должны располагаться внутри площадных объектов	Областные центры должны находиться внутри областей

2.3. Атрибутивные семантические ошибки

Существенные свойства объектов ГИС (и значения символьных атрибутов, определяющих эти свойства) описывают табличные данные, организованные на основе ERM-моделей реляционных СУБД. Для таких данных на сегодняшний день известно и описано в литературе большое количество конкретных методов контроля, систематизированных по-разному, на основе различных критериев. В одной из достаточно, на наш взгляд, обоснованных классификаций [6], усовершенствованной в [7], предлагается систематизация семантических ошибок ввода данных, основанная на учете следующих факторов:

- уровней организации (агрегации) защищаемых и контролируемых элементов;
- характера семантических ошибок;
- форм используемой для контроля избыточности;

– моментов введения избыточности.

Для обнаружения ошибок в сформированной БД реальное значение имеет только «естественная» избыточность, основанная, по сути, на неполном использовании всевозможных значений комбинаций символов, составляющих атрибуты и их сочетания. Естественная избыточность сопровождает данные с самых ранних этапов их формирования и кодирования и позволяет обнаруживать ошибки, внесенные практически на любом этапе технологического процесса подготовки, ввода и обработки данных. Поэтому целесообразно ограничиться двумя первыми критериями, выделив следующие уровни организации атрибутов (количественных и призначных), охватываемых НП-правилами:

- один атрибут одной записи таблицы БД;
- несколько атрибутов одной записи;
- атрибуты разных записей таблицы;
- атрибуты записей разных таблиц.

Начальный фрагмент соответствующего перечня возможных НП-правил приведен в табл. 2.

Таблица 2. НП-правила для атрибутивной информации

Правило 2.2	Смысл	Пример
2.2.1. Один атрибут записи		
2.2.1.1. Значение количественного атрибута N должно находиться в заданном диапазоне	Значение N не должно быть меньше или больше заданных N_{min}, N_{max}	Количество населения населенного пункта определенного типа должно находиться в установленном для данного типа диапазоне (город, ПГТ, ...)
2.2.1.2. Призначный атрибут Q должен принадлежать множеству разрешенных значений	Значение Q должно быть найдено в заданном справочнике	Названия областей, районов, населенных пунктов, улиц, ...
2.2.2. Несколько атрибутов одной записи		
2.2.2.1. Значения количественного атрибута N_{i1}, N_{i2} записи должны быть связаны арифметическим соотношением A	Должно выполняться равенство $N_{i1} = A(N_{i2}, N_{i3} \dots)$	Общее количество населения конкретного населенного пункта должно совпадать с количеством населения по категориям для данного пункта
2.2.2.2. Сочетание призначных атрибутов Q_{i1}, Q_{i2} должно принадлежать множеству разрешенных значений	Сочетание значений Q_{i1}, Q_{i2} должно быть найдено в заданном справочнике или соответствовать заданному логическому условию	Сочетание кодов области, района (города областного управления) и населенного пункта должно принадлежать перечню кодов административно-территориального деления Украины
...

Приведенные в табл. 2 и дальнейшие НП-правила последующих уровней организации могут быть представлены следующим обобщенным правилом:

2.2.k.j. Значения аргументов НП-правила должны удовлетворять заданному логико-арифметическому соотношению (ЛАС) [8].

Многообразие возможных ЛАС и их аргументов для каждого из уровней может быть велико в зависимости от предметной и проблемной ориентации конкретной системы.

2.4. Пространственно-атрибутивные семантические ошибки

«Перекрестные» ошибки типа 3.1, заключающиеся в нарушении соответствия между тематической и пространственной базами данных.

Типовые правила, определяющие проверки соответствия, охватывают:

3.1.1. Отсутствие разногласий в идентификационных данных.

3.1.2. Соответствие площадей полигональных объектов.

3.1.3. Соответствие длин линейных объектов.

В принципе, ошибки типа 3.1 можно считать разновидностью атрибутивных ошибок, НП-правила которых охватывают значения данных, находящиеся в базе геоданных (или вычисляемые).

3. Модуль проверки пространственных данных системы «ГИС-Леспроект»

Одним из типовых приложений ГИС к задачам государственного управления является управление лесным хозяйством. Основными направлениями развития ПО «Укрлеспроект» были определены: расширение информационного и проектного обеспечения Государственной целевой программы «Леса Украины» (2010–2015 гг.), дальнейшее развитие в лесоустройстве и лесном хозяйстве современных ГИС-технологий. Система «ГИС-Леспроект» разработана для нужд ПО «Укрлеспроект» в рамках мероприятий по выполнению указанной Госпрограммы.

Процесс ведения лесоустройства ПО «Укрлеспроект» можно разбить на ряд взаимосвязанных укрупненных этапов:

- формирование начальной тематической и картографической информации для лесничеств;
- сохранение текущей информации в локальных базах данных и формирование выходных бумажных документов;
- накопление картографических и таксационных данных на серверах баз данных по временным, территориальным и другим характеристикам и на основе этих данных генерация различного рода выходных документов (планшетов, планов, карт-схем, отчетов, справок, ...).

На рис. 3 представлена схема взаимосвязи функциональных модулей, обеспечивающих решение задач ПО «Укрлеспроект» по лесоустройству.

При формировании карт-материалов могут быть использованы различные наборы тематических слоев, такие, как границы лесничеств, кварталов, выделов, различные дороги, объекты гидрографии и др. Всего в ПО «Укрлеспроект» предусмотрено использование 255 первичных и 103 агрегированных слоев, где однотипные тематические слои для удобства формирования карт-материалов объединяются в группы. Например, газопроводы, линии электропередач, линии связи, нефтепроводы и водопроводы объединяются в один слой – коммуникаций и др.

Проведенный анализ получаемых данных показал наличие нарушений корректности топологических связей для определенных тематических слоев. С целью устранения ошибок оцифровки был определен набор соответствующих правил проверки топологии пространственных данных лесоустройства и создана дополнительная компонента системы

«ГИС-Леспроект» для проверки топологических зависимостей. Значительный объем топологических проверок вызван количеством обрабатываемых тематических слоев. Из полного списка реализованных правил проверки топологии приведем перечень применяемых проверок для тематических слоев, задействованных, преимущественно, во всех типах изготавливаемых тематических карт-материалов.

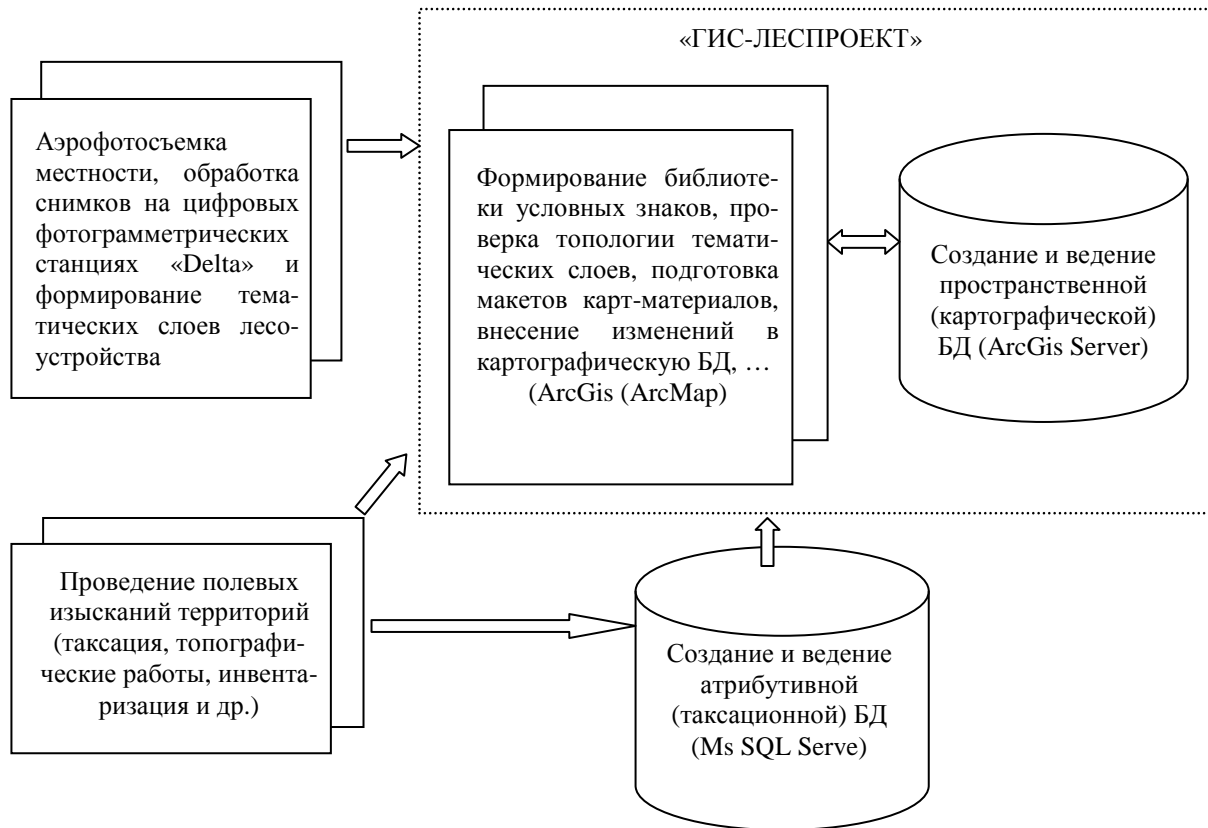


Рис. 3. Схема функциональных модулей ПО «Укргослеспроект»

Правило 1.1.1.1. Не должны перекрываться с ...

Тематические слои «Границы лесничеств», «Квартальные просеки шириной до 5 м», «Квартальные просеки по противопожарным разрывам», «Квартальные просеки на мелиоративных каналах рек», «Квартальные просеки по дорогам», «Квартальные просеки шириной от 5 м до 10 м», «Границы таксационных выделов» не должны перекрываться внутри класса объектов.

Правило 1.1.1.2. Не должны иметь пробелов.

Тематические слои «Границы лесничеств», «Квартальные просеки шириной до 5 м», «Границы таксационных выделов» не должны иметь пробелов.

Правило 1.1.1.3. Не должны перекрываться.

Полигоны тематических слоев «Автомоби́ль с искусственным покрытием», «Канал мелиорат широкий», «Карьеры», «Озера», «Противопожарный разрыв шириной более 10 м», «Реки», «Реки пересыхающие», «Ставки», «Ставки пересыхающие», «Водохранилища», «Яры» не должны перекрываться с полигонами тематического слоя «Границы таксационных выделов».

Правило 1.1.1.4. Должны совмещаться с объектами класса.

Полигоны тематического слоя «Квартальные просеки шириной до 5 м» должны покрываться полигонами слоя «Границы лесничеств».

Правило 1.1.1.5. Должны совмещаться с ...

Полигоны тематических слоев «Границы таксационных выделов», «Канал мелиорат широкий», «Противопожарный разрыв шириной более 10 м», «Озера», «Ставки», «Ставки пересыхающие», «Водохранилища» должны покрываться одним полигоном тематического слоя «Квартальные просеки шириной до 5 м».

Правило 1.1.1.6. Граница должна совпадать с ...

Границы тематических слоев: «Линии государственной границы», «Границы областей», «Границы административных районов», «Границы городских земель», «Границы государственных заповедников и национальных парков», «Указатель изменения границы между землевладельцами» должны совпадать с границами «Квартальных просек шириной до 5 м».

Правило 1.1.1.7. Граница площадного объекта должна совпадать с границей.

Границы полигонов тематического слоя «Границы лесничеств» Должны совпадать с границами слоя «Квартальные просеки шириной до 5 м».

Правило 1.1.2.3. Не должны перекрывать сами себя.

Линии тематических слоев линейной гидрографии («Реки узкие», «Реки пересыхающие узкие», «Ручьи», «Ручьи пересыхающие», «Каналы мелиорат шириной 3–5 м», «Канавы (рвы) шириной от 1 м до 3 м»), коммуникаций («Газопровод», «Трубопроводы», «Нефтепровод», «Линии связи», «Водопровод») не должны перекрывать сами себя.

Правило 1.1.3.1. Должны лежать на границе.

Точки тематического слоя «Копцы» должны лежать на границе слоя «Границы лесничеств».

4. Направления развития модуля диагностики ошибок

В настоящее время диагностика целостности компонентов ИР ГИС "ГИС-Леспроект" практически ограничивается контролем синтаксических ошибок топологии. Представляется целесообразным (и закономерным) расширение области диагностики в направлении охвата семантических ошибок 1.2, 2.2 и 3.2.

4.1. Семантические ошибки топологии

В отличие от синтаксических топологических ошибок, разложенных по классификационным "полочкам" и сравнительно легко диагностируемых, семантические ошибки явным образом не видны, и их обнаружение и исправление связаны со значительно большей трудоемкостью, в частности, и "ручной". Большой частью семантические ошибки топологии можно рассматривать как ошибки позиционной точности данных, которая определяется как величина отклонения измерения данных о местоположении (обычно координат) от истинного значения; ошибки регистрации и определения контрольных точек, преобразования координат, особенно когда не известна проекция исходного документа; ошибки обработки данных, генерализации и проблемы интерпретации; ошибки перевода векторных данных в растровый формат [4].

В основе ИР нынешних ГИС-технологий управления лесоустройством во многом лежат лесоустроительные планшеты, ранее выполняемые ручным способом. Ручные технологии изначально не могли обеспечить нужную точность, а многократное воспроизводство лесоустроительных планшетов привело к накоплению и углублению ошибок в положениях опорных линий (до 60 м и более) и угловой дезориентации геодезических данных (до 12–18 градусов) [4]. Типовые проявления общих ошибок: отклонения направлений квартальных просек, отклонения относительного взаимоположения лесных массивов от их действительного положения на местности и т.п.

Лесные геоданные в той форме, в которой они хранятся в базе геоданных ПО "Укр-гослеспроект", по точности иногда приближаются к кадастровым данным (а именно, для

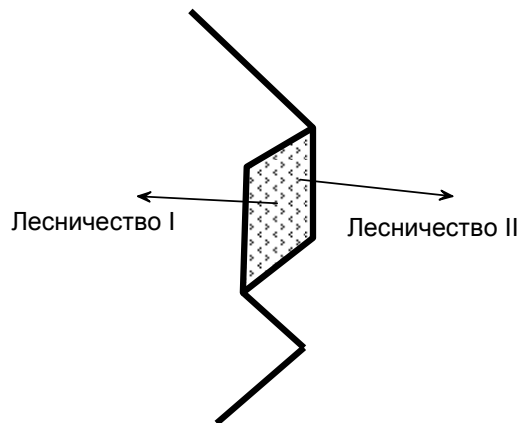


Рис. 4. Пример ошибок границ лесничеств

внешних границ лесничеств) и часто для границ лесных кварталов, так как внешние границы лесничеств определяются по геодезическим точкам, а лесные кварталы ограничены на местности просеками или иными естественными или искусственными рубежами. Границы же объектов внутри лесных кварталов менее точны как из-за затрудненности наблюдений на местности, так и нечеткой, в сущности, природы лесных объектов. Хотя и для границ лесничеств возможны ошибки, пример которых показан на рис. 4.

Кардинальным способом обнаружения и исправления неточностей является ручная сводка между собой границ топологических объектов лесоустройства с выполнением геодезических работ (дополнительный геодезический контроль). Учитывая нечеткую природу

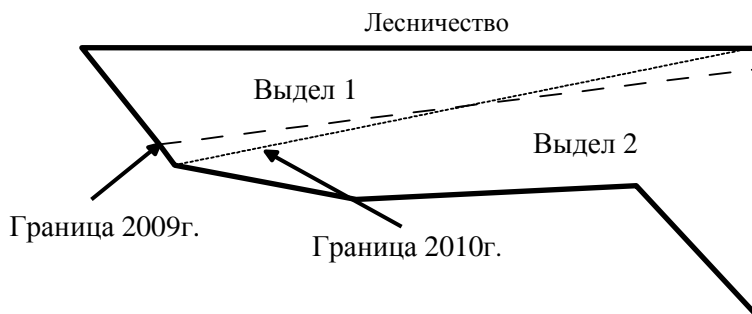


Рис. 5. Пример "подозрительного изменения" в 2009–2010 гг.

объектов лесного хозяйства, а также возможность необоснованно значительных затрат, естественно такую сводку производить в первую очередь для сомнительных топологических объектов. Для обнаружения таких объектов представляется полезным ретроспективный анализ пространственно-временных зависимостей между границами объектов, установленных в разные моменты времени (учитывая принятый в лесоустройстве интервал обновления информации – один год). Пример того, как ретроспективный анализ в топологии может вызвать сомнение в правильности, показан на рис. 5.

Решение задачи контроля семантических ошибок в атрибутивных данных включает два этапа:

4.2. Семантические ошибки в атрибутивных данных

1) анализ смысловых связей атрибутов и установление ЛАС (логико-арифметические соотношения), определяющих НП-правила;

2) организация базы правил непротиворечивости (БПНП).

Уточним сначала смысл и формы задания ЛАС для соответствующего НП-правила. Конкретное ЛАС, охватывающее кортеж из K , N и Q – атрибутов, может быть выражено

через K – местный предикат $P[(N, Q)_{j_1}, \dots, (N, Q)_{j_k}]$, для которого справедливо равенство

$P = 1$, если элементы кортежа находятся в определенном отношении $R_{ЛАС}$.

Первый этап заключается в поиске атрибутов – одной записи, разных записей, разных таблиц – связанных известным отношением $R_{ЛАС}$, и формулировании соответствующих НП-правил в терминах ЛАС. Формулировка должна включать локализирующую, адресную часть (имена таблиц, доменов, записей) и операционную часть (логические операции AND, OR, \in , ... и арифметические операции +, -, *, ...). Масштабы этого этапа иллюстрируют следующие цифры, характеризующие атрибутивную БД системы «ГИС-

Леспроект». Количество таблиц – более 40, справочников – более 100, среднее количество атрибутов в таблице – 12–15. Важное значение имеет ранжирование ЛАС, основанное на оценке ожидаемой результативности.

Второй этап заключается в выборе канонической формы представления ЛАС в БПНП (например, продукция, точнее, ее условная часть – IF...) и подготовка логической структуры БПНП.

Логическая структура БПНП определяет для проектантов (а впоследствии и для администратора системы) в наглядном виде правила непротиворечивости и атрибуты БД, с которыми они связаны. В [8] для подобного назначения предлагается графово-логическая модель, в которой узлы графа представляют собой укрупненные структурные составляющие БД (типы данных, заголовки таблиц, атрибутов), а обозначенные связи объединяют элементы (атрибуты, домены), входящие в предикаты, которые составляют определенное НП-правило. Составной частью модели является спецификация, содержащая соответствующий предикат с сокращенной адресной (локализирующей) частью. Таким образом, "графическая" составляющая отражает факт наличия смысловых связей элементов БД в смысле взаимозависимости их значений, а "логическая" - описывает сущность этих взаимозависимостей. По сути, такая модель подобна графовым концептуальным ER-моделям. Принципиальная разница заключается в том, что в ER-диаграммах отображаются бинарные отношения типа "входит в", "выходит из", "первичный – порожденный" и т.п., а в упомянутой модели – К-арные отношения "значения элементов соответствуют НП-правилу с номером обозначенной связи".

Собственно говоря, НП-правила (и организацию их проверки) можно рассматривать как определенное обобщение стандартных операторов типа CONSTRAINT CHECK SQL-команды CREATE, реализующих табличные ограничения на целостность данных, которые проверяются при вводе информации в БД. Обобщение направлено на:

- расширение области определения составляющих кортежей на различные структуры ИБ;
- расширение типов логических условий и операций контрольных проверок;
- расширение возможностей планирования процессов выполнения контрольных проверок (по сравнению, например, с механизмами триггеров).

4.3. Семантические пространственно-атрибутивные ошибки

Проверка соответствия картографической БД и таксационной БД (ТБД) возможна по следующим параметрам (атрибутам):

- 1) номеров кварталов;
- 2) номеров выделов в кварталах;
- 3) площадей кварталов и выделов;
- 4) площадей полигональных и линейных выделов (для которых существенна площадь) в разрезе кварталов;
- 5) категорий земель;
- б) длин линейных выделов.

Для п.п. 1,2 проверка выполняется прямым перебором и сравнением номеров кварталов и соответствующих им выделов в КДБ и ТБД.

Для п.п. 3,4 проверки возможны для таких тематических слоев, как: «Границы лесничеств», «Квартальные просеки шириной до 5 м», «Границы таксационных выделов», «Автомобильные дороги с искусственным покрытием», «Овраги», «Карьеры», «Оползни действующие», «Озера», «Водохранилища», «Ставки», «Ставки пересыхающие», «Канал мелиоративный широкий», «Противопожарный разрыв шириной более 10 м», «Реки», «Реки пересыхающие», «Автомобильные дороги без покрытия», «Усовершенствованные грунтовые

дороги», «Автомобильные дороги с покрытием», «Автомобильные дороги с покрытием, которые строятся».

Проверка п. 5 выполняется на основании имеющейся в КДБ информации относительно топологических кодов выделов объектов лесоустройства и категорий земель, хранящихся повывдельно в ТБД.

Проверка п. 6 возможна, например, для таких линейных тематических слоев, как «Квартальные просеки шириной от 5 м до 10 м», «Квартальные просеки по противопожарным разрывам», «Квартальные просеки по мелиоративным каналам, рекам», «Квартальные просеки с нечеткими границами», «Квартальные просеки по дорогам», «Газопроводы», «Трубопроводы», «Нефтепроводы», «Линии связи», «Водопроводы», «Узкоколейные железные дороги», «Лесные дороги до 5 м», «Дамбы и искусственные валы шириной более 3 м», «Улучшенные лесные дороги», «Постоянные пешеходные тропы», «Хозяйственный визирь», «Реки узкие», «Реки пересыхающие узкие», «Ручьи», «Ручьи пересыхающие», «Каналы мелиоративные шириной 3-5м», «Канавы (рвы) шириной от 1 до 3 м», «Хребты», «Экскурсионный маршрут», «Научно-познавательный маршрут», «Эколого-познавательный маршрут», «Туристический маршрут», «Водный маршрут», «Мосты», «Противопожарный разрыв шириной до 10 м», «Таксационный визирь». Для выполнения данной проверки необходима соответствующая информация в ТБД.

5. Заключение

Приведенная в разд. 1 классификация ошибок носит общий «скелетный» характер и для конкретных приложений нуждается в развитии и конкретизации. В первую очередь это относится к семантическим ошибкам. Но и перечень синтаксических ошибок топологии, приведенный в табл. 1, не носит исчерпывающего характера. Так, например, в кадастровой системе учета недвижимости [3] выделяются ошибки, которые могут быть выявлены с помощью средств ArcGis, например, наложения границ участков (правило 1.1), пересечения границ участков (правило 2.4), контур не должен иметь всяких узлов (правило 2.1), «недолет» или «перелет» характерной точки границы земельного участка (правило 3.2), земельный участок находится на территории другого кадастрового квартала (правило 1.4). Однако применение правил топологии, имеющихся в ArcGis, недостаточно для устранения всех имеющихся ошибок оцифровки кадастровых геоданных. Так, например, не будут выявлены разрывы границ или их нестыковка (чересполосица), разворот границ и др. Для выявления и устранения такого рода ошибок, а также для автоматизации процесса проверки топологии целесообразно создание дополнительных проблемно-ориентированных модулей проверки, применяемых к штатным возможностям полнофункциональных ГИС. Пример подобного модуля описан в [10].

Особенности геоданных в управлении лесоустройством (нечеткая «морфология» лесных объектов, исторически сложившаяся низкая точность их представления) таковы, что семантические ошибки топологии (в первую очередь, ошибки точности) могут быть предметом повышенного внимания. Хотя для представления лесных объектов потенциально нужна меньшая точность, чем, например, для кадастровой недвижимости, количество и значимость факторов влияния на точность здесь больше. В связи с этим представляется возможным и перспективным диагностика семантических ошибок топологии на основе нечетких правил, определяющих правдоподобность границ и характер объектов лесоустройства. Важная роль в создании таких правил отводится авторами пространственно-временным отношениям между объектами лесоустройства.

СПИСОК ЛІТЕРАТУРИ

1. Практичне застосування ГІС-технологій для планово-картографічного супроводження лісовпорядкування (на прикладі системи «ГІС-Ліспроєкт») / Б.О. Білецький, В.А. Литвинов, В.П. Беспалов [та ін.] // Математичні машини і системи. – 2013. – № 3. – С. 76 – 86.
2. Майстренко С.Я. Система «ГІС-ЛІСПРОЕКТ» як прототип геоінформаційної складової кадастрової системи / С.Я. Майстренко // Математичні машини і системи. – 2015. – № 3. – С. 93 – 99.
3. Овчинникова А. Классификация кадастровых ошибок [Электронный ресурс] / А. Овчинникова. – Режим доступа: <http://www.zemvest.ru/jurnal/archiv-jurnala/9-2013/07/>.
4. Тикунов В.С. Геоинформатика / Тикунов В.С. – М., 2005. – 480 с.
5. ArcGis 10. Geodatabase topology rules and topology error fixes [Электронный ресурс]. – Режим доступа: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Geodatabase_topology_rules_and_topology_error_fixes/001t000000sp000000/.
6. Литвинов В.А. Контроль достоверности и восстановления информации в человеко-машинных системах / В.А. Литвинов, В.В. Крамаренко. – Киев: Техніка, 1986. – 200 с.
7. Алгоритми і моделі автоматичної ідентифікації та корекції типових помилок користувача на основі природної надмірності / Г.Є. Кузьменко, В.А. Литвинов, В.І. Ходак [та ін.] // Математичні машини і системи. – 2004. – № 2. – С. 134 – 148.
8. Задача повышения качества информационной базы агентно-ориентированных многоуровневых интеллектуализованных СППР / Г.Е. Кузьменко, В.А. Литвинов, Ю.Г. Пилипенко [и др.] // Математичні машини і системи. – 2002. – № 3. – С. 53 – 61.
9. Ott T. Time-Integrative Geographic Information Systems: Management and Analysis of Spatio-Temporal Data / T. Ott, F. Swiaczny. – Berlin; Heidelberg; New York; Barcelona; Hong Kong; Springer, 2001. – 234 с.
10. Федорова Н. Модуль проверки пространственных данных на наличие топологических ошибок [Электронный ресурс] / Н. Федорова. – Режим доступа: <http://www.itpgrad.ru/node/1575>.

Стаття надійшла до редакції 23.02.2016