

УДК 528.443: 004.052.42

А. А. Шоломицкий, д-р техн. наук, проф., М. Л. Лепеш, магистр

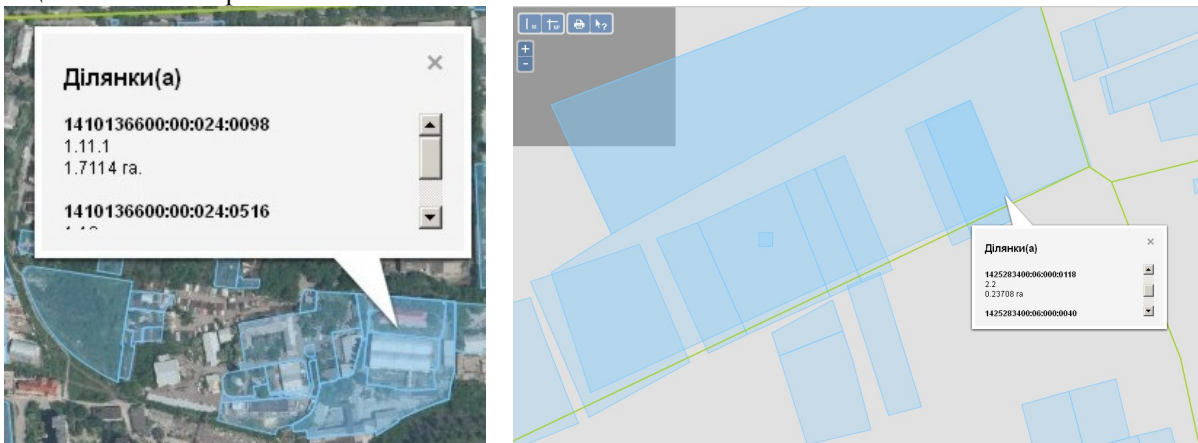
ГВУЗ «Донецкий национальный технический университет», Донецк, Украина

Использование формальных грамматик для анализа кадастровой информации

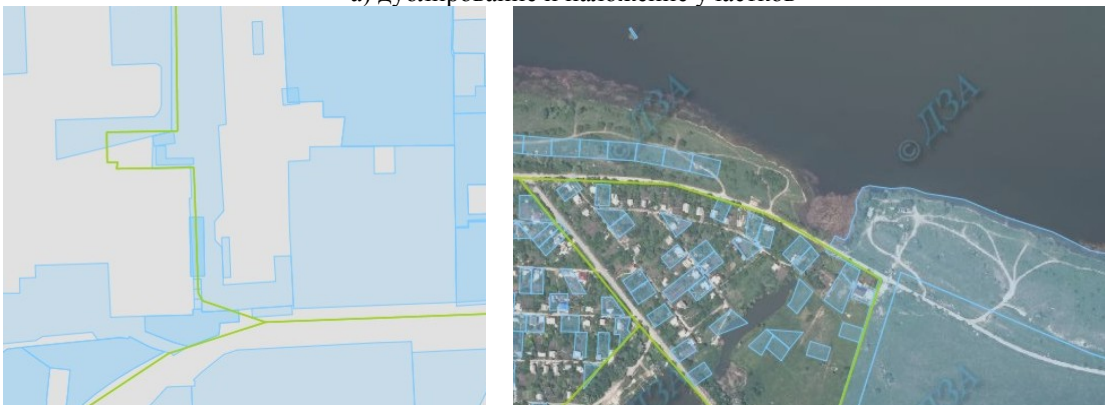
В статье рассмотрены вопросы контроля полноты кадастровой информации в автоматизированной системе государственного земельного кадастра. Предлагается для формального представления и контроля информации о земельном участке использовать контекстно-свободные грамматики. Выполнен анализ полноты информации по этому методу и предложены пути преодоления его ограничений.

Ключевые слова: объект, информация, формальная грамматика, база данных, контроль, качество.

С января 2013 года государственное агентство земельных ресурсов Украины запустило в эксплуатацию публичную кадастровую карту Украины [1]. Долгожданное событие, однако, даже беглого взгляда достаточно, чтобы увидеть огромное количество ошибок, связанных с геометрией и неправильным пространственным расположением кадастровых элементов (рис. 1). Это дублирование участков, наложение и пересечение различных границ. Ошибки в семантической информации земельных участков проверить невозможно, т. к. это закрытая информация. Однако можно предположить, что ошибки и полнота этой информации будут такими же, как и в обменных файлах, т. к. заполнение информацией кадастровой карты Украины производилось именно с помощью обменных файлов.



а) дублирование и наложение участков



б) пересечение границ и систематический сдвиг

Рис. 1. Ошибки пространственного положения объектов

Для создания автоматизированной системы государственного земельного кадастра в 2003 году в качестве формата внесения информации о земельном участке был принят файл IN4 [2,3]. Однако, уже через 7 лет (в 2010 году) был предложен уже другой более полный формат XML[4,5].

Статистика говорит о низком уровне технической документации и обменных файлов их подготовки: 2009 год возвращены землеустроительным организациям 197 тыс. отчетов из 436 принятых на проверку (45%) [6].

Обменные файлы XML, в соответствии с Указом Госкомзема от 02.11.2009 № 573 формируются с целью внесения данных в Поземельную книгу и книгу записей регистрации государственных актов на право собственности на землю и на право постоянного пользования землей, договоров аренды, какие ведутся в электронном виде, и определяют набор базовых лексических и синтаксических правил для построения обменных файлов [5].

В соответствии с [5] обменный файл - это электронный документ унифицированной формы для обмена информацией, которая используется при ведении Поземельной книги и книги записей регистрации государственных актов на право собственности на землю и на право постоянного пользования землей, договоров аренды земли в электронном виде, государственного земельного кадастра и осуществлении топографо - геодезических работ, работ по землеустройству.

Таким образом, под обменным файлом понимается - унифицированная структура земельно-кадастровых данных, которая содержит результаты выполненных земельно-кадастровых работ.

Обменный файл создается в виде файла формата XML в кодировке Unicode (UTF-8). Структура обменного файла состоит из двух частей: служебной и информационной (см. рис. 2).

Служебная часть (AdditionalPart) используется для формирования и обмена информацией о реквизитах обменного файла и исполнителя работ. Служебная часть состоит из: информации об обменном файле; информации о лицах, которые сформировали, осуществили проверку (коррекцию) данных обменного файла.

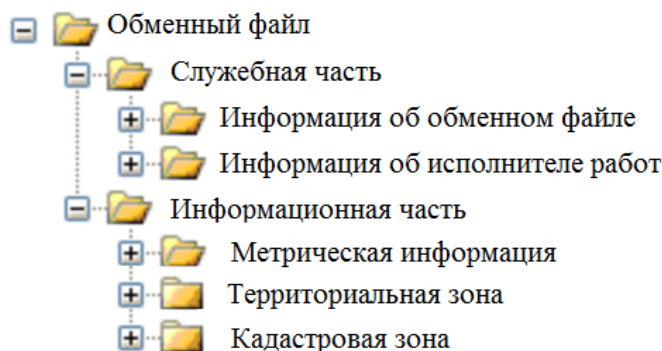


Рис. 2. Структура обменного файла

По сравнению с файлом в формате IN4, XML-файл намного сложнее [7], тогда как IN4-файл содержит в линейном виде дескрипторы, которым присваивается соответствующее значение, то XML-файл является "деревом" элементов, этот файл является структурным. XML - это стандарт построения языков разметки, который представлен текстовым форматом, который предназначен для хранения любых иерархически структурированных данных, обмена информацией между различными приложениями, с определенными синтаксическими правилами. Именно обмен информацией - главная цель создания XML. Программы, предназначенные для просмотра и обработки документа в формате XML, позволяют, не зная синтаксических правил устройства файла, создавать и изменять информацию, которую передают, например, в формате обменного файла.

Согласно требованиям к XML-файлу, он должен быть корректным (well-formed) и валидным (valid). Эти два требования являются стандартами в определении правильности XML-документа.

Первое требование - корректность (well-formed)-означает, что правильно составленный документ должен соответствовать всем общим правилам синтаксиса, установленным для XML. Если допущена ошибка в корректности файла, то XML-процессор (парсер) (программа,

которая проверяет правильность написания документа) отказывает в обработке заданного файла и классифицирует такую ситуацию как роковую ошибку, а файл не может быть XML-документом.

Второе требование – валидность (valid) – предполагает, что к XML - документу могут применяться дополнительные правила синтаксиса, согласно соответствующим стандартам данных. Такие правила хранятся в специальных файлах – схемах, где описана структура документа, все возможные атрибуты и названия элементов. Если XML - документ содержит непредвиденное значение, то XML - процессор (валидатор) сообщает об ошибке.

На основе анализа 72 948 обменных файлов IN4 из нескольких областей Украины, включающих описание 50 518 кварталов и 287 984 земельных участков, разными исследователями были выделены следующие виды ошибок [8]:

- дублирование кадастровых единиц;
- пересечение границ учетных кадастровых единиц;
- дублирование кадастрового номера;
- нарушение в географическом расположении кадастровых единиц, например, кадастровые единицы располагаются в разных координатных зонах;
- некорректное присвоение кадастрового номера участку.

Авторы предполагают, что большинство этих ошибок имеются и в обменных файлах формата XML и заносятся в базу данных государственной автоматизированной кадастровой системы.

Цель и задачи исследований

Авторы видят своей целью анализ полноты заполнения информацией обменных файлов формата XML, хотя идеи и методы анализа могут применяться к любой земельно-кадастровой информации, в том числе уже занесенной в базы данных.

Для анализа полноты информации о земельных участках необходимо иметь формальную модель этого объекта. Авторы предлагают использовать для этого формализм контекстно-свободных грамматик [9].

Формальная модель земельного участка

Формальная грамматика – способ описания формального языка, то есть выделения некоторого подмножества из множества всех слов некоторого конечного алфавита. Различают порождающие и распознающие (или аналитические) грамматики — первые задают правила, с помощью которых можно построить любое слово языка, а вторые позволяют по данному слову определить, входит оно в язык или нет.

Терминал или терминальный символ – объект, непосредственно присутствующий в словах языка, соответствующего грамматике, и имеющий конкретное, неизменяемое значение (обобщение понятия «слово»). Нетерминал или нетерминальный символ – объект, обозначающий какую-либо сущность языка (например: формула, арифметическое выражение, команда) и не имеющий конкретного символического значения.

Формальная грамматика определяется следующими элементами:

$$G = (S, T, N, P)$$

где T – набор терминальных символов; N – набор нетерминальных символов; P – набор правил вида: «левая часть» \rightarrow «правая часть»,

где: «левая часть» — непустая последовательность терминалов и нетерминальных символов, содержащая хотя бы один нетерминальный символ; «правая часть» — любая последовательность терминальных и нетерминальных символов;

S – стартовый или начальный символ грамматики.

Для определения сущности понятия выполняют вывод. Выводом называется последовательность строк, состоящих из терминальных и нетерминальных символов, где первой идет строка, состоящая из одного стартового нетерминального символа, а каждая последующая строка получена из предыдущей путем замены некоторой подстроки по одному (любому) из правил. Конечной строкой является строка, полностью состоящая из терминальных символов, и следовательно являющаяся выражением языка.

По иерархии Хомского, грамматики делятся на 4 типа, каждый последующий является более ограниченным подмножеством предыдущего. Мы использовали в работе тип 2 - контекстно-

свободные грамматики (левая часть состоит из одного нетерминального символа). Иногда ее еще называют КС-грамматика или бесконтекстная грамматика.

Контекстно-свободные грамматики применяются для определения структуры в грамматическом анализе. КС-грамматики находят широкое применение в информатике. Ими задается грамматическая структура большинства языков программирования, структурированных данных и т.д. Для разбора КС-грамматики достаточно автомата со стеком [9].

Набор возможных конструкций РБНФ очень невелик. Это конкатенация, выбор, условное вхождение и повторение.

Структуру начального нетерминального символа <Земельный участок> определим на основании приказа «Об утверждении Требований к структуре, содержанию и формату оформления результатов работ по землеустройству в электронном виде (обменного файла)» [5].

<Земельный участок> ::= <Служебная информация> <Информационная часть>
 <Служебная информация> ::= <Информация об обменном файле> <Информация об исполнителе работ>
 <Информационная часть> ::= <Метрическая информация> <Территориальная зона> <Кадастровая зона>

Согласно Приложения 2 [5] определим нетерминальный символ <Информация об обменном файле>.

<Информация об обменном файле> ::= <Идентификация> <Версия формата>
 <Подразделение Центра ДЗК> <Идентификатор подразделения Центра ДЗК> <Название П
 О> <Версия ПО>
 <Идентификация> ::= <Дата формирования> <Уникальный идентификатор>

В данной форме символы <Дата формирования>, <Уникальный идентификатор>, <Версия формата>, <Подразделение Центра ДЗК>, <Идентификатор подразделения Центра ДЗК>, <Название ПО>, <Версия ПО> являются терминальными символами.

Далее, в соответствии с Приложениями 3–4, 40–41 [5] сформулируем нетерминальный символ <Информация об исполнителе работ>.

<Информация об исполнителе работ> ::= <Сведения об исполнителе>
 <Описание проверок или корректировок данных>
 <Сведения об исполнителе> ::= <Наименование исполнителя> (<Код ЕГРПОУ> | <Код ГРФО>) <Лицензия> <Ответственное лицо> <Исполнитель> <Местонахождение>
 <Лицензия> ::= <Серия> <Номер> <Дата выдачи>
 <Ответственное лицо> ::= <ФИО> <Должность>
 <Исполнитель> ::= <ФИО> <Должность> <Контактная информация>
 <Контактная информация> ::= <Телефон> <Факс> <Адрес электронной почты>
 <Местонахождение> ::= <Страна> <Почтовый индекс> <Область> <Район области> <Населенный пункт> <Улица (проспект, проулок и др.)> <Номер дома> <Номер корпуса> <Номер квартиры>
 <Код страны> ≡ число (код по Классификатору Д007–96)

Далее, согласно Приложениям 5–6 Приказа определим нетерминальный символ <Метрическая информация>.

<Метрическая информация> ::= <Система координат> <Система высот>
 <Единицы измерения> { <Координаты узлов полигона> }₁ⁿ <Описание полилиний>
 { <Закрепленные узлы> }₁^a
 n – количество узлов (n > 0)
 a – количество закрепленных узлов (a ≥ 2)
 <Система координат> ::= <SC42> | <SC42_3> | <Lokat> | <USK2000> | <WGS84> | <SC63>
 <Система высот> ::= <Baltic> | <Baltic77> | <Other>
 <Единицы измерения> ::= <M> | <Km> | <Other>
 <Координаты узлов полигона> ::= { <Уникальный номер узла> <Имя узла> <Метод определения> <X> <Y> [<H>] <MX> <MY> [<MH>] [<Описание>] }₁^m

m – количество узлов полигона

$\langle \text{Метод определения} \rangle ::= \langle \text{Геодезические измерения} \rangle | \langle \text{GPS съемка} \rangle \langle \text{Дигитализация} \rangle | \langle \text{Фотограмметрические измерения} \rangle$

$\langle \text{Описание полилиний} \rangle ::= \{ \langle \text{Уникальный номер полилинии} \rangle \langle \text{Описание узлов} \rangle$

$\langle \text{Длина} \rangle \}_{1}^k$

k – число полилиний

$\langle \text{Описание узлов} \rangle ::= \langle \text{Уникальный номер узла} \rangle_1 | \langle \text{Уникальный номер узла} \rangle_2 \dots |$

$\langle \text{Уникальный номер узла} \rangle_i \equiv \{ \langle \text{Уникальный номер узла} \rangle \}_2^i$

i – количество узлов в полилинии ($i > 2$)

$\langle \text{Закрепленные узлы} \rangle ::= \{ \langle \text{Уникальный номер узла} \rangle \}_2^j$

j – количество закрепленных узлов ($j \geq 2$)

Символы $\langle \text{SC42} \rangle$, $\langle \text{SC42}_3 \rangle$, $\langle \text{Lokal} \rangle$, $\langle \text{USK2000} \rangle$, $\langle \text{WGS84} \rangle$, $\langle \text{SC63} \rangle$, $\langle \text{Baltic} \rangle$, $\langle \text{Baltic77} \rangle$, $\langle \text{Other} \rangle$, $\langle \text{M} \rangle$, $\langle \text{Km} \rangle$, $\langle \text{Уникальный номер узла} \rangle$, $\langle \text{Имя узла} \rangle$, $\langle \text{Геодезические измерения} \rangle$, $\langle \text{GPS съемка} \rangle$, $\langle \text{Дигитализация} \rangle$, $\langle \text{Фотограмметрические измерения} \rangle$, $\langle \text{X} \rangle$, $\langle \text{Y} \rangle$, $\langle \text{H} \rangle$, $\langle \text{MX} \rangle$, $\langle \text{MY} \rangle$, $\langle \text{MH} \rangle$, $\langle \text{Уникальный номер полилинии} \rangle$, $\langle \text{Описание узлов} \rangle$, $\langle \text{Длина} \rangle$ являются терминальными символами.

Приложения 4-5, 9- 41 [5] послужили основой для рассмотрения нетерминального символа $\langle \text{Кадастровая зона} \rangle$.

$\langle \text{Кадастровая зона} \rangle ::= \langle \text{КОАТУУ} \rangle \langle \text{Номер кадастровой зоны} \rangle$

$\langle \text{Признак кадастровой единицы} \rangle \langle \text{Внешний полигон} \rangle \langle \text{Кадастровый квартал} \rangle$

$\langle \text{КОАТУУ} \rangle \equiv \text{число (10 символов)}$

$\langle \text{Номер кадастровой зоны} \rangle \equiv \text{число (2 символа)}$

$\langle \text{Признак кадастровой единицы} \rangle ::= \langle \text{Зона (кадастровая)} \rangle$

$\langle \text{Внешний полигон} \rangle ::= \langle \text{Граница} \rangle \{ \langle \text{Внутренний полигон} \rangle \}_2^c$

c – количество внутренних полигонов

$\langle \text{Кадастровый квартал} \rangle ::= \langle \text{Номер квартала} \rangle \langle \text{Информация о регионе} \rangle$

$\langle \text{Внешний полигон} \rangle \langle \text{Участок} \rangle$

$\langle \text{Номер квартала} \rangle \equiv \text{число (3 символа)}$

$\langle \text{Участок} \rangle ::= \langle \text{Местоположение} \rangle \langle \text{Категория и целевое назначение} \rangle$

$\langle \text{Форма собственности} \rangle \langle \text{Метрическая информация участка, ограничения} \rangle$

$\langle \text{Собственники или пользователи} \rangle \langle \text{Право пользования участком} \rangle \langle \text{Реквизиты}$

$\langle \text{документации} \rangle \langle \text{Государственные акт} \rangle \langle \text{Денежная оценка} \rangle \langle \text{Аренда} \rangle \langle \text{Субаренда} \rangle$

$\langle \text{Ограничения} \rangle \langle \text{Угодья} \rangle \langle \text{Смежники} \rangle$

Такие фразы были составлены для всех элементов информации о земельных участках, регламентированных [5].

Проверка формальной модели

Для проверки информации о земельных участках была создана программа, которая выполняет анализ правильности фраз КС-грамматики на корректность и полноту заполнения. Исходными данными для проверки наличия ошибок, содержащихся в обменных файлах, стали обменные файлы в XML-формате созданные в ходе паевания земель сельскохозяйственного назначения. Обменные файлы были созданы для 1487 земельных участков, входящих в состав пяти кадастровых зон.

При добавлении обменных файлов в базу программно выполняется проверка на отсутствие данных входящих в нетерминальные символы. Результаты анализа записывались в базу данных для систематизации, обобщения и статистического анализа.

Очень часто незаполненными остаются нетерминальные символы: $\langle \text{Аренда земельного участка} \rangle$, $\langle \text{Субаренда земельного участка} \rangle$, $\langle \text{Ограничения земельного участка} \rangle$. Всегда указываются нетерминальные символы: $\langle \text{Категория и целевое назначение} \rangle$, $\langle \text{Форма собственности} \rangle$. Практически полностью заполняется информация по следующим нетерминальным символам: $\langle \text{Метрическая информация} \rangle$, $\langle \text{Кадастровый квартал} \rangle$ и $\langle \text{Угодья земельного участка} \rangle$. Общая гистограмма заполнения информации в обменных файлах приведена на рис. 3.

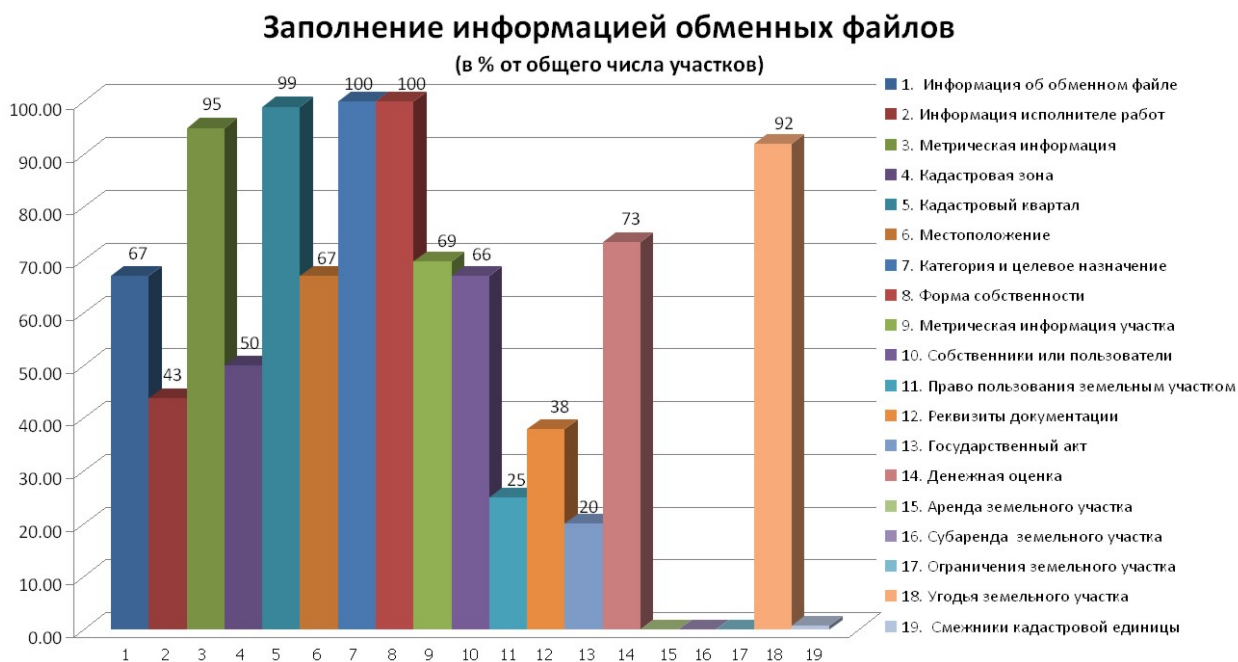


Рис. 3. Гистограмма заполнения информации в обменных файлах

Детальный анализ ошибок паспортных данных приведен на рис 4.

Метрическая информация передает измерительные характеристики объекта (земельного участка). Она отражает свойство участка располагаться в определенной части пространства и занимать некоторую его часть.

Диаграмма распределение ошибок в паспортных данных (в % от общего количества участков)

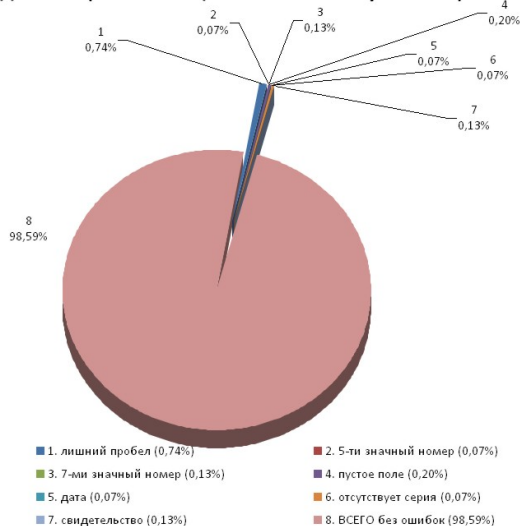


Диаграмма распределение видов ошибок в паспортных данных (в % от общего числа ошибок)

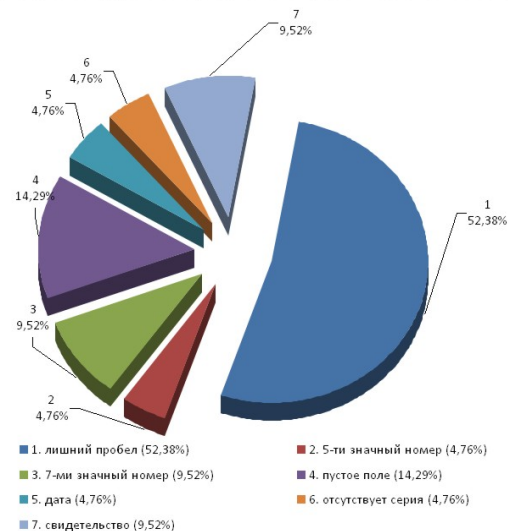


Рис. 4. Количество и вид ошибок в паспортных данных

Метрическая информация в обменном файле содержит координатные и некоторые (числовые) атрибутивные данные. По этим данным вычисляются площади и выполняется проверка на пересечение границ земельных участков.

При анализе пересечения контуров, необходимо определить, что можно считать пересечением. При принятии обменных файлов в Донецкой области допустимым является пересечение (наложение границ) не более 1 см. Эта цифра ни чем не аргументирована. Более того она противоречит указу Госкомзема «Про затвердження Інструкції про встановлення (відновлення) меж земельних ділянок в натурі (на місцевості) та їх закріплення межовими

знаками» [10]. Средняя квадратическая ошибка местоположения межевого знака относительно ближайших пунктов государственной геодезической сети, геодезических сетей сгущения, городских геодезических сетей не должна превышать:

- 1) в городах Киеве, Севастополе, городах - областных центрах и городах областного подчинения – 0,1 м;
- 2) в других городах и поселках – 0,2 м;
- 3) в селах – 0,3 м;
- 4) за границами населенных пунктов для земельных участков площадью до 10 га – 0,5 м, а площадью 10 га и больше – 2,5 м [10].

Можно сделать вывод, что допустимая величина наложения границ земельных участков должна зависеть от точности геодезических работ. На практике установлено, что пересечение границ участков не редкость. Часто пересечения границ земельных участков объясняется тем, что каталоги координат границ земельных участков в правоустанавливающих документах (до 2003 г.) составлены преимущественно в условных системах координат. Очень сложно установить их границы на местности в принятой в настоящее время системе координат. Вторая, и основная причина, это низкое качество геодезических работ при выполнении кадастровых съемок.

При анализе метрической информации программно контролировалась площадь земельного участка. Так, в программе был выполнен перерасчет площадей земельных участков по координатам, и сравнение площади, указанной в обменном файле и вычисленной по координатам. Гистограмма на рис. 5 иллюстрирует наличие обменных файлов с отклонением в площади земельного участка.

**Диаграмма отклонения площади земельных участков
(в % от общего количества участков)**

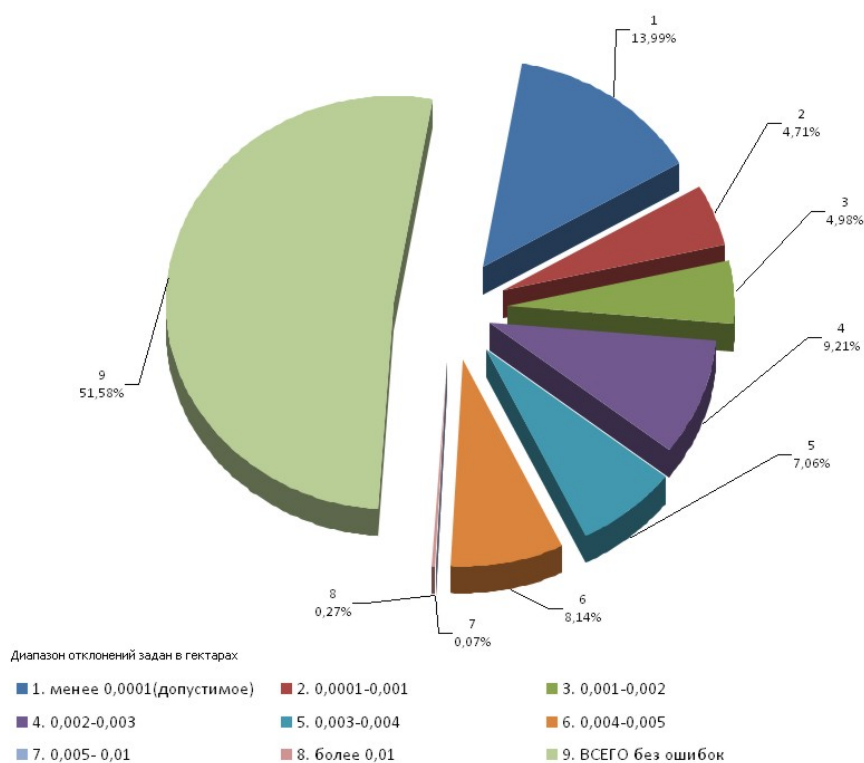


Рис. 5. Гистограмма отклонений площадей земельных участков

Анализ обменных файлов в формате XML показал, что число ошибок в файлах этого формата уменьшилось по сравнению с так называемым форматом IN4 [8, 11]. Новый формат обменного файла позволяет преодолеть многие ограничения, присущие обменному файлу формата IN4. Однако и новый формат может содержать ошибки и пропуски информации, без которой

функционирование автоматизированной системы государственного земельного кадастра будет неэффективной.

Выводы

Анализ полноты заполнения базы данных автоматизированной системы государственного земельного кадастра показал, что формализм контекстно-свободных грамматик позволяет выполнить анализ полноты информации в обменных файлах и в информационных системах, везде, где можно формально описать объекты и их характеристики с помощью выражений контекстно-свободных грамматик.

Однако, этот же анализ показывает, что не во всех случаях достаточно возможностей КС-грамматик, в некоторых случаях информация о земельных участках и его атрибутах зависит от контекста информации. Например, на рис. 3 можно сделать вывод, что информация по элементам 15, 16, 17 пропущена при заполнении обменных файлов, однако на самом деле в этой группе обменных файлов нет участков с этой информацией. Поэтому для более корректного анализа заполнения информации следует использовать и более высокий уровень интеллектуальной обработки данных и знаний – уровень экспертных систем [12].

В экспертной системе можно задать систему правил, которая будет учитывать контекстную зависимость элементов информации между собой. Кроме того, экспертные системы можно эффективно использовать не только для контроля полноты информации, но и для контроля правильности информации. Для этого в экспертную систему можно заложить правила создания SQL-запросов к базам данных других государственных учреждений. Например паспортные данные о владельце(ах) земельного участка можно проверять по базам данных министерства внутренних дел, данные о юридических лицах можно проверять по базам данных областных статистических управлений, данные о физических лицах по базе данных ИНН и т. д.

Конечно статистика анализа обменных файлов в XML-формате не такая представительная, как была проведена различными исследователями с форматом обменных файлов IN4 [8, 11], и может содержать как субъективные ошибки исполнителей и представлять их квалификационный уровень. Одно очевидно – использование КС-грамматик в сочетании с экспертными системами могут значительно улучшить качество информации в земельно-кадастровых системах и повысить эффективность их применения.

Библиографический список

1. Публічна кадастрова карта України [карта] [Електронний ресурс]. – Режим доступа: <http://map.dazru.gov.ua/kadastrova-karta>. – 10.01.2013.
2. Указ Президента Украины "О мерах по созданию единой системы государственной регистрации земельных участков, недвижимого имущества и прав на них в составе государственного земельного кадастра" от 17.02.2003 № 134/2003 [Электронный ресурс]. – Режим доступа: <http://zakon2.rada.gov.ua/laws/show/134/2003>. – 18.11.2011.
3. Постановление Кабинета Министров Украины "О создании единой системы государственной регистрации земельных участков, недвижимого имущества и прав на них в составе государственного земельного кадастра" от 17.07.2003 № 1088 [Электронный ресурс]. – Режим доступа: <http://zakon1.rada.gov.ua/laws/show/1088-2003-%D0%BF>. – 18.11.2011.
4. Постановление Кабинета Министров Украины "Об утверждении порядков ведения Поземельной книги и Книги записей о государственной регистрации государственных актов на право собственности на земельный участок и на право постоянного пользования земельным участком, договоров аренды земли" от 09.09.2009 № 1021 [Электронный ресурс]. – Режим доступа: <http://zakon2.rada.gov.ua/laws/show/1021-2009-%D0%BF>. – 18.11.2011.
5. Приказ Государственного комитета Украины по земельным ресурсам «Об утверждении Требований к структуре, содержанию и формату оформления результатов работ по землеустройству в электронном виде» от 02.11.2009г. № 573 [Электронный ресурс]. – Режим доступа: <http://zakon2.rada.gov.ua/laws/main/nreg.z0157-10>. – 18.09.2012.
6. Сохнич С. А. Состояние и проблемы создания автоматизированной системы ведения государственного земельного кадастра [Электронный ресурс] / С. А. Сохнич. - Режим доступа: http://www.nbu.gov.ua/portal/Chem_Biol/Vldau/APK/2010_1/files/10ssslcc.pdf. - 20.05.2012.
7. Малицкий А. XML - новый формат обменного файла для земельного кадастра / А. Малицкий // Сучасні досягнення геодезичної науки та виробництва: Збірник наукових праць. – 2011. – Вип. 2 (22). – С. 229–231.

8. Германова К. О. Вдосконалення технологи створення та редагування індексно- кадастрової карти / К. О. Германова, І. Л. Шубш // Наукові праці Донецького національного технічного університету Сер. «Гірничо-геологічна». – 2009. – Вип. 9 (143). – С. 58–68.
9. Горбатов А. В. Фундаментальные основы дискретной математики. Информационная математика / А. В. Горбатов. – М.: Наука, Физматлит, 2000. – 545 с.
10. Указ Государственного комитета Украины по земельным ресурсам «Про затвердження Інструкції про встановлення (відновлення) меж земельних ділянок в натурі (на місцевості) та їх закріплення межовими знаками» от 18.05.2010 №376 [Электронный ресурс]. - Режим доступа: <http://zakon2.rada.gov.ua/laws/show/z0391-10.> - 18.09.2012.
11. Кияшко Е. А. Исследование кадастровой информации в базах данных [Электронный ресурс] / Е. А. Кияшко. - Режим доступа: <http://www.masters.donntu.edu.ua/2010/igg/kiyashko/diss/index.htm>. - 18.11.2011.
12. Тейз А. Логический подход к искусственному интеллекту: от модальной логики к логике баз данных / А. Тейз., П. Грибомон; пер. с франц. – М.: Мир, 1998. – 494 с.

Надійшла до редакції 10.04.2013

А. А. Шоломицький, М. Л. Лепеш

ДВНЗ «Донецький національний технічний університет», Донецьк, Україна

Використання формальних граматик для аналізу кадастрової інформації

В статті розглянуто питання контролю повноти кадастрової інформації в автоматизованій системі державного земельного кадастру. Пропонується для формального представлення і контролю інформації по земельних ділянках використовувати контекстно-вільні грамматики. Виконано аналіз повноти інформації за цим методом і пропонуються шляхи подолання його обмежень.

Ключові слова: об'єкт, інформація, формальна граMATика, база даних, контроль, якість.

A. Sholomytskyi, M. Lepesh

Donetsk National Technical University, Donetsk, Ukraine

Using formal grammar for cadastral information analysis

The article analyzes metrical and semantic errors in information in public cadastral system of Ukraine, causes of their appearance and issues of complete cadastral information control. We propose to use context-free grammars for formal presentation of information about the land. The formalization of information about the land parcel in the context-free grammar form was done and software to analyse sentences and accumulate statistics in the database was created. We also analyze completeness of information in exchange files using this method. We suggest using expert systems that allows analyzing context-dependent information to overcome the limitations. For the control of correctness of information in public cadastral system we proposed to use queries to databases of other government organizations.

Keywords: object, information, formal grammar, database, control, quality.