На правах рукописи

Маставичене Татьяна Викторовна

Совершенствование технологии ведения адресного реестра для повышения эффективности информационной системы кадастра недвижимости

Специальность 25.00.26 – «Землеустройство, кадастр и мониторинг земель»

Автореферат

диссертации на соискание ученой степени

кандидата технических наук

Москва – 2016

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный университет геодезии и картографии» (МИИГАиК) на кафедре кадастра и основ земельного права

Научный руководитель: кандидат технических наук, профессор

Максудова Людмила Георгиевна

Официальные оппоненты:

Цветков Виктор Яковлевич, доктор технических наук, профессор, ОАО «Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте», заместитель руководителя Центра перспективных фундаментальных и прикладных исследований

Юнусов Альберт Гамзатович, кандидат технических наук, федеральное государственное бюджетное образовательное учреждении высшего образования «Государственный университет по землеустройству», профессор кафедры геодезии и геоинформатики

Ведущая организация:

Открытое акционерное общество «Институт территориального планирования «Кировское архитектурное, землеустроительное проектно изыскательское предприятие» (ОАО «Кировгипрозем»)

Защита диссертации состоится «16» июня 2016 г. в 10 часов на заседании диссертационного совета Д.212.143.04 при Московском государственном университете геодезии и картографии по адресу: 105064, Москва К-64, Гороховский пер., д.4, МИИГАиК, зал заседаний Учёного совета.

С диссертацией можно ознакомиться в библиотеке МИИГАиК и на сайте по ссылке: http://miigaik.ru/nauka/dissertacionyy_sovet/dissertatsii/20160406211345-4518.pdf

Автореферат разослан и размещен на сайте www.miigaik.ru «___»_____2016 г.

Учёный секретарь
диссертационного совета Беленко Виктор Владимирович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы диссертации. В условиях глобального экономического кризиса особое значение имеет всесторонняя оптимизация налогово-бюджетной политики государства, что требует не только построения, но и продуктивного использования ГКН как фундаментальной базы экономики, одного из ключевых инструментов налогово-бюджетной политики государства и планирования развития территорий. При этом чрезвычайно важно помнить, что построение и поддержка в актуальном состоянии ГКН – трудоемкий и финансово затратный процесс, требующий активного применения современных аппаратных средств вычислительной техники, программного обеспечения и привлечения специалистов высокой квалификации.

Методологическая поддержка, разработка эффективных алгоритмов и информационных систем контроля и обработки адресной информации является важным направлением в совершенствовании информационных систем кадастра. Адресный реестр, являясь составным элементом, формирующим ГКН, непосредственным образом влияет на реализацию всего спектра стратегических целей и задач государства: от создания эффективного рыночного оборота недвижимости для формирования доходной налоговой базы до социальной защиты населения и планирования миграционной политики. Формирование и ведение адресного реестра часто сопряжено с трудностями, так как есть такие объекты налогообложения как комната, доля в праве на домовладение, постройки в составе единого объекта недвижимости, доля в праве на объект недвижимости и другое. В реальности адресный реестр должен обеспечивать корректную и однозначную фиксацию всего многообразия сделок с недвижимостью.

Степень разработанности проблемы. При работе над диссертацией были изучены коллективные труды и отдельные монографии российских авторов, посвященные математическим методам и алгоритмам, позволяющим

эффективно оперировать пространственными данными: Бугаевского Л.М., Верещака Т.В., Запрягаевой Л.А., Ивановой Л.Г., Книжникова Ю.Ф., Савиных В.П., Спиридонова А.И., Филатова В.Н., Цветкова В.Я., Максудовой Л.Г. и др. Несмотря на большое количество публикаций о значении и методах организации кадастра и структуры пространственных данных, некоторые вопросы нуждаются в более углубленном рассмотрении. Так, например, не получили до нашего времени достаточного рассмотрения недостатки в существующих способах кодирования и регистрации адресных данных.

Цель исследования. Разработать решения, повышающие эффективность ведения кадастра и, одновременно, позволяющие снижать финансовые затраты на администрирование кадастра или при выполнении типовых операций по массовой кадастровой оценке по итогам подробного обзора наиболее перспективных возможностей современных ГИС и СУБД, позволяющих строить эффективные информационные системы кадастра.

Задачей исследования является:

- разработать метод совершенствования ведения адресного реестра, универсально учитывающий переменную структуру адреса в информационных системах кадастра;
- предложить метод, снимающий ограничения XML-формата электронных форм кадастровых документов при работе с большими пространственными объектами в информационных системах кадастра.

Объектом исследования в работе являются информационные системы кадастра. **Предметом исследования** – конкретные реализации применяемых технологий, алгоритмов и методов в информационных системах кадастра.

Метод диссертационного исследования. В ходе работы применялись общенаучные методы структурно-функционального анализа, использовались экспертные оценки, графические и табличные приемы сравнительной визуализации данных.

Информационной основой исследования послужили официально опубликованный Классификатор адресов Российской Федерации, спецификации электронных форм документов для дистанционного предоставления кадастровых сведений.

Научная новизна работы.

1. Обоснована необходимость отказаться от реализации КЛАДР структурой из 7 таблиц в пользу иерархического справочника. В существующем КЛАДР не предусмотрено промежуточного адресообразующего элемента между районом и населенным пунктом, а именно, сельского округа или сельского поселения. Тогда как их использование необходимо в тех случаях, когда в муниципальном районе существуют несколько населенных пунктов с одинаковыми наименованиями, например, две деревни Окулово в Подосиновском районе Кировской области.

2. Разработан метод реализации корректного адресного реестра на основе иерархического справочника в СУБД Oracle Database и Microsoft SQL Server, универсально решающий проблему как необходимости, так и избыточности промежуточных адресообразующих элементов в отличие от существующей реализации КЛАДР.

3. Обоснована критика действующих в настоящее время требований к формату электронных форм XML-документов для дистанционного предоставления кадастровых сведений, которые в современных условиях не способны обеспечить развитие рационального использования и охраны земель, защиту прав собственников, землепользователей и арендаторов, справедливого формирования земельного налога и арендной платы. Используемые в нем тяжеловесные синтаксические конструкции контрпродуктивно влияют на эффективность процедур ведения кадастра, увеличивают вычислительную нагрузку на аппаратные комплексы, затрудняют контроль и своевременную коррекцию кадастровых сведений.

4. Обоснована модернизация существующего формата электронных форм XML-документов для дистанционного предоставления кадастровых сведений: расширить XSD-схему описания состава кадастровых XML-документов (словарь, модели и типы данных) с возможностью вынесения координатных данных из тела XML-документа, например, во внешний файл или сервер публикации кадастровых сведений. Кроме того, для сложных пространственных объектов, имеющих в составе более 10 тысяч точек, в спецификации формата XML-документа необходимо предусматривать возможность разделения на набор менее сложных с последующим агрегированием в целевой пространственный объект, что сделает возможным обработку пространственных объектов произвольной сложности в отличие от текущей реализации спецификации.

5. Впервые выполнено сравнение размеров межевого плана в формате XML на несколько пространственных объектов с представлением этого пространственного объекта в формате WKT и показано, что размер кадастрового XML документа в 25 раз больше представления этого пространственного объекта в формате WKT (Приложение А).

Теоретическая и практическая ценность работы.

Материалы и обобщённые примеры, содержащиеся в диссертации, применимы для аналитических сообщений, учебных пособий и специальных курсов, использованы в учебном процессе при преподавании по специальностям городской кадастр, градостроительный кадастр, земельный кадастр. Результаты работы применимы при совершенствовании и перспективном развитии информационных систем кадастра, детализации требований в технических заданиях при размещении заказов на выполнение работ для государственных и муниципальных нужд при ведении кадастра или оценки недвижимости. Применение алгоритмов, приведенных в диссертационной работе, позволит снизить финансовые и трудовые затраты на администрирование кадастра при достижении целевых требований к

быстродействию информационных систем кадастра, корректности и полноте кадастровых сведений.

Также практической целью работы является попытка ввести в профессиональное дискуссионное поле квалифицированное понимание процессов, критически важных при построении высокоэффективных информационных систем кадастра, важность опыта исследовательской конкуренции, полученного при разработке ГИС и СУБД ведущими отечественными и мировыми поставщиками ИТ-решений для разработки национальных систем кадастра недвижимости.

Результаты, выносимые на защиту.

Алгоритм реализации адресного реестра с использованием иерархического справочника.

Метод модернизации формата кадастровых XML-документов с возможностью вынесения координатных данных из тела XML-документа, например, во внешний файл или сервер публикации кадастровых сведений.

Соответствие диссертации паспорту научной специальности.

Результаты диссертационного исследования соответствуют паспорту научной специальности 25.00.26 «Землеустройство, кадастр и мониторинг земель» по пунктам областей научных исследований: 5 – «Принципы сбора, документирования, накопления, обработки и хранения сведений о земельных участках. Разработка единой методики по ведению земельного кадастра», 7 – «Информационное обеспечение Государственного земельного кадастра», 29 – «Разработка земельно-информационной системы (ЗИС) как основной части геоинформационной системы (ГИС) на основе современных информационных и геоинформационных технологий».

Эмпирическую основу исследования составила работа в Правительстве Кировской области, в филиале федерального государственного бюджетного учреждения «Федеральная кадастровая палата Федеральной службы государственной регистрации, кадастра и

картографии» по Кировской области, в институте территориального планирования «Кировское архитектурное, землеустроительное проектно-изыскательское предприятие».

Степень достоверности и апробация результатов. Достоверность работы базируется на использовании строго доказанных и корректно применяемых положениях фундаментальных и прикладных наук: дискретная математика, математический анализ, математическая статистика, теория вероятностей – и не противоречит их положениям, а предложенные в работе методы и алгоритмы строго согласуются с опытом их создания и дальнейшего совершенствования.

Основные положения диссертации докладывались на научно-технических конференциях студентов, аспирантов и молодых ученых МИИГАиК в 2012 и 2013 годах.

Публикации. По материалам диссертации опубликованы 3 статьи в журналах, рекомендованных ВАК РФ.

Объем и структура диссертационной работы. Диссертационная работа состоит из введения, трех глав, заключения, списка принятых сокращений и аббревиатур, списка используемой литературы и приложения. Основной текст изложен на 100 стр., в том числе 8 табл., 36 рис. Список использованной литературы включает 136 наименований, в том числе 109 на русском языке и 27 на иностранных языках.

Автор выражает благодарность научному руководителю – к.т.н., проф. Максудовой Л.Г., а также всем коллегам за оказанную помощь в подготовке диссертационной работы.

Основное содержание работы

Работа состоит из трех глав.

Глава 1 состоит из 3-х частей: в ней рассматривается значение информационных систем кадастра для государственного и муниципального управления, указывается на существующие недостатки при ведении адресного реестра и действующих в настоящее время XML-документов, электронных форм межевого и технического плана.

В первой части главы 1 показывается значимость кадастра и информационных систем кадастра для государственного и муниципального управления. Раскрывается состав информационных систем кадастра и роль СУБД в их архитектуре, указывается на значение СУБД в построении эффективных информационных систем кадастра.

Во второй части главы 1 подробно описывается структура Федеральной адресной системы и Классификатора адресов РФ. Указываются существующие методологические недостатки применяемых в настоящее время информационных систем кадастра. Обращается внимание на недостатки текущей реализации КЛАДР, не позволяющие корректно представлять часть адресов.

Сложившееся административно-территориальное деление Российской Федерации не позволяет корректно представить часть адресов. Например, только в Кировской области есть несколько одноименных населенных пунктов (деревень Окулово в Яхреньгском и Октябрьском с/о Подосиновского района), находящихся, согласно административно-территориальному делению, в одном муниципальном районе, но включенных в различные сельские поселения.

В третьей части главы 1 критически рассмотрен формат XML-документов электронных форм межевого и технического плана, предназначенного для дистанционного предоставления кадастровых сведений.

Выводы по главе 1

Решающую роль в эффективности информационных систем кадастра играет СУБД, функциональные возможности которой определяют эффективность информационных систем кадастра в целом. Среди наиболее востребованных функциональных возможностей современных СУБД являются высокопроизводительные и оптимально организованные инструменты работы с пространственными данными, обеспечивающие не только эффективную обработку, но и ввод-вывод пространственных данных. Для целей ведения адресного реестра в задачах кадастра СУБД должны эффективно обеспечивать работу с иерархическими структурами.

Глава 2 состоит из 4-х частей. В ней рассматриваются математические подходы и алгоритмы, реализованные в современных СУБД, позволяющие добиваться эффективности информационных систем кадастра. Предлагается использовать методы теории вероятностей для ускоренного анализа пространственных данных, показан пример реализации.

В первой части второй главы подробно рассматриваются математические основы методов индексирования пространственных данных и способы создания пространственных индексов в СУБД Oracle Database и Microsoft SQL Server. Описан механизм декомпозиции пространства карты в сеточную иерархию (рисунок 1). Раскрываются понятия минимального ограничивающего прямоугольника, тесселяции. Рассмотрены правила тесселяции.

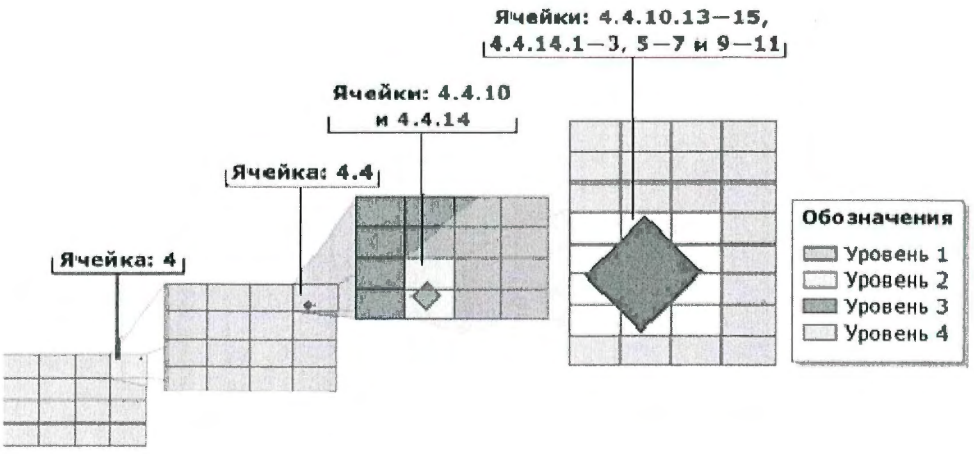

Рисунок 1. Декомпозиция пространства карты

Во второй части главы 2 посвящена рассмотрению процесса верификации и генерализации пространственных данных. Здесь перечислены требования к единообразию систем координат пространственных объектов, математическим моделям топологии пространственных объектов. Перечислены наиболее часто встречающиеся топологические отклонения. Указаны математические методы генерализации пространственных данных.

Третья часть главы 2 посвящена алгоритмам пространственного анализа и оптимизации быстродействия обработки пространственных данных. Например, может оказаться, что смежные пространственные объекты не имеют согласованных границ сопряжения, а имеют зазоры между границами или, напротив, границы имеют взаимные пересечения. Исправление подобных проблем, очевидно, не требуют графического представления на карте и могут успешно решаться использованием эффективных инструментов, реализованных для этих целей в большинстве современных СУБД (рисунок 2).

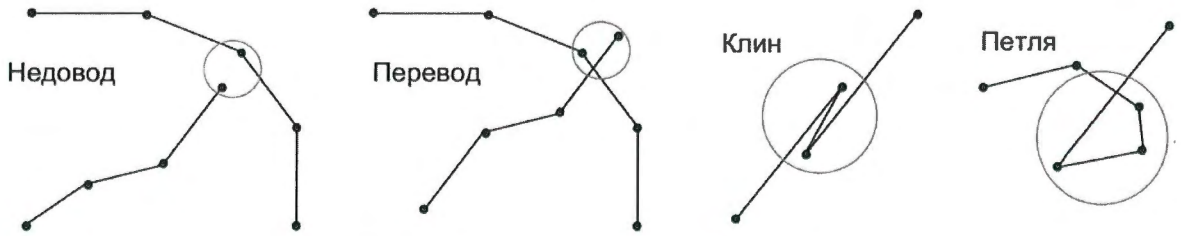

Рисунок 2. Типичные нарушения корректности топологии

Показаны математические методы: нахождения точки пересечения отрезков, определения пересечения полилиний различными методами, определения площади полигона (с указанием на особенности нахождения площади в системе координат Гаусса-Крюгера), определения нахождения точки внутри полигона или снаружи. Предлагается использовать методы теории вероятностей для анализа пространственных данных, показана реализация такого метода на примере нахождения площади произвольной

фигуры методом Монте-Карло. Указаны преимущества и недостатки применения методов теории вероятностей. Рассмотрена задача определения центральной репрезентативной точки полигона и полигональная операция – вычисление скелета полигона.

В четвертой части главы 2 адрес показывается как иерархическая структура (рисунок 4), рассматриваются методы работы с адресом, применяя математические модели, разработанные в таком разделе дискретной математики как теория графов.

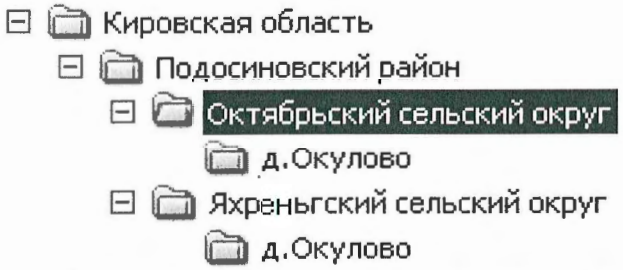

Рисунок 4. Иерархическая структура адреса

Показан метод представления иерархической структуры в таблице, имеющей два столбца: [идентификатор строки] и [ссылку на непосредственного предка]. В представленной ниже таблице 2 это столбцы [ID] и [PARENT_ID] соответственно.

Таблица 2 - Пример таблицы с иерархической структурой

Адреса			
ID	PARENT_ID	TYPE_ID	TXT
…			
1230001	NULL	1	Кировская область
1230002	1230001	2	Подосиновский район
1230003	1230002	101	Яхреньгский сельский округ
1230004	1230003	3	Деревня Окулово
…			
1240003	1230002	101	Октябрьский сельский округ
1240004	1240003	3	Деревня Окулово
…			

Развернуто показаны инструменты: структуры, типы данных, SQL-запросы, позволяющие оперировать иерархическими данными в СУБД. Уделено внимание особенностям работы с иерархическими данными в СУБД Microsoft SQL Server и Oracle Database.

Выводы по второй главе

В данной главе диссертационной работы рассмотрены теоретические основы, лежащие в основе эффективности информационных систем кадастра. В главе подробно изложены этапы и обоснованы теоретические положения, заложенные в основу эффективной обработки пространственных данных.

Анализ тщательно проработанного математического и алгоритмического аппарата, оптимизирующего работу с пространственными данными в ГИС и СУБД, позволяет сделать вывод о том, что любой недостаток в информационных системах кадастра, касающийся адресного реестра или электронного формата кадастрового документа, выглядит недопустимым, нуждается в анализе и доработке. Таким образом, выявляется:

- недостаточная эффективность решений, основанных на XML формате;
- в совершенствовании XML формата кадастровых документов находится резерв повышения продуктивности информационных систем кадастра.

Во второй части главы рассматривается работа с адресом как иерархической структурой. Подробно описаны возможности современных СУБД для работы с объектами, имеющими иерархические связи.

Глава 3 состоит из 4-х частей и содержит примеры наиболее эффективной реализации решений наиболее типовых задач, возникающих при ведении кадастра и адресного реестра.

Первая часть главы 3 продолжает рассматривать тему адреса как иерархической структуры. Даются примеры создания таблиц, предназначенных для работы с иерархическими данными. Подробно показаны примеры заполнения иерархических таблиц и SQL-запросы, позволяющие извлекать иерархические данные в требуемой форме. Делается предложение о модернизации структуры таблиц КЛАДР иерархическим

справочником. Указывается на полезную особенность справочника иерархического типа, позволяющую хранить иную, необходимую при администрировании кадастра недвижимости, информацию, даже если такая информация и не является адресной.

Вторая часть главы 3 развивает тему наиболее эффективных методов работы с пространственными данными в СУБД. Указывается на соответствие кадастровых сведений свойствам и признакам Big Data.

Термин и понятие Big Data сложился в 2008-2010 годах в результате попыток анализа феномена взрывного роста объёмов и многообразия обрабатываемых данных, происходившего в это время. Базы пространственных данных, использующиеся в информационных системах кадастра, обладают всеми признаками Big Data: имеют огромные объёмы и значительное многообразие источников получения и представления данных, имеют выраженный характер непрерывного прироста.

Существуют как успешные решения, использующие Data Mining, так и неудачный опыт применения этой технологии. Области, где применения технологии Data Mining, вероятнее всего, будут успешными, имеют следующие особенности:

- требуют решений, основанных на предметных знаниях о процессах реального мира;
- имеют непрерывно изменяющийся состав исходных данных;
- имеют доступные для контроля, необходимые и достаточные данные;
- обеспечивают высокие дивиденды от правильных решений.

В заключении части указывается на аспект, открывающий революционные перспективы в области работы с Big Data – квантовые алгоритмы, способные принципиально ускорить выполнение некоторых алгоритмических задач.

В третьей части главы 3 рассматриваются возможности для работы с пространственными данными в наиболее популярных современных СУБД. Наиболее показательные возможности работы с пространственными данными различных СУБД представлены в таблице 3.

Таблица 3 – Пространственные возможности различных СУБД

	Тип пространственного индекса	Наличие предустановленных СК, применяющихся в РФ	Функции сетевого анализа	Поддержка спецификаций OGC
Oracle Database	R-Tree	+	+	+
DB2	R-Tree	-	+	+
MySQL	MBR	-	-	+
PostgreSQL	R-Tree	-	-	+
MS SQL Server	R-Tree	-	-	+

В четвертой части главы 3 рассматривается XML-формат электронных форм документов, предназначенных для дистанционного предоставления кадастровых сведений по сравнению с WKT-форматом.

Приводятся экспертные оценки XML-формата. Даются данные сравнительного формирования межевого плана в формате XML на несколько пространственных объектов показывающие, что размер кадастрового XML документа в 25 раз больше представления этого пространственного объекта в формате WKT (таблица 4, рисунок 5).

Таблица 4 - Зависимость размера файла кадастрового XML документа и структуры пространственного объекта в формате WKT от количества точек составляющих пространственный объект

Кол-во точек пространственного объекта	**Размер файла XML, КВ**	**Размер файла WKT, КВ**	**Отношение размера XML к размеру WKT**
127	75	4	18,75
1 135	622	25	24,88
6 065	3 535	141	25,07
8 374	4 785	192	24,92
10 115	5 725	231	24,78
17 644	10 143	406	24,98
24 581	14 323	568	25,22
245 844	143 202	5 670	25,26
270 428	157 522	6 237	25,26
295 013	171 842	6 804	25,26
319 598	186 162	7 370	25,26
368 767	214 802	8 504	25,26
417 937	243 442	9 638	25,26
467 106	272 082	10 772	25,26
516 275	300 721	11 906	25,26
565 445	329 361	13 040	25,26
При попытке сформировать XML файл с количеством точек более 600 тыс. происходит сбой.			

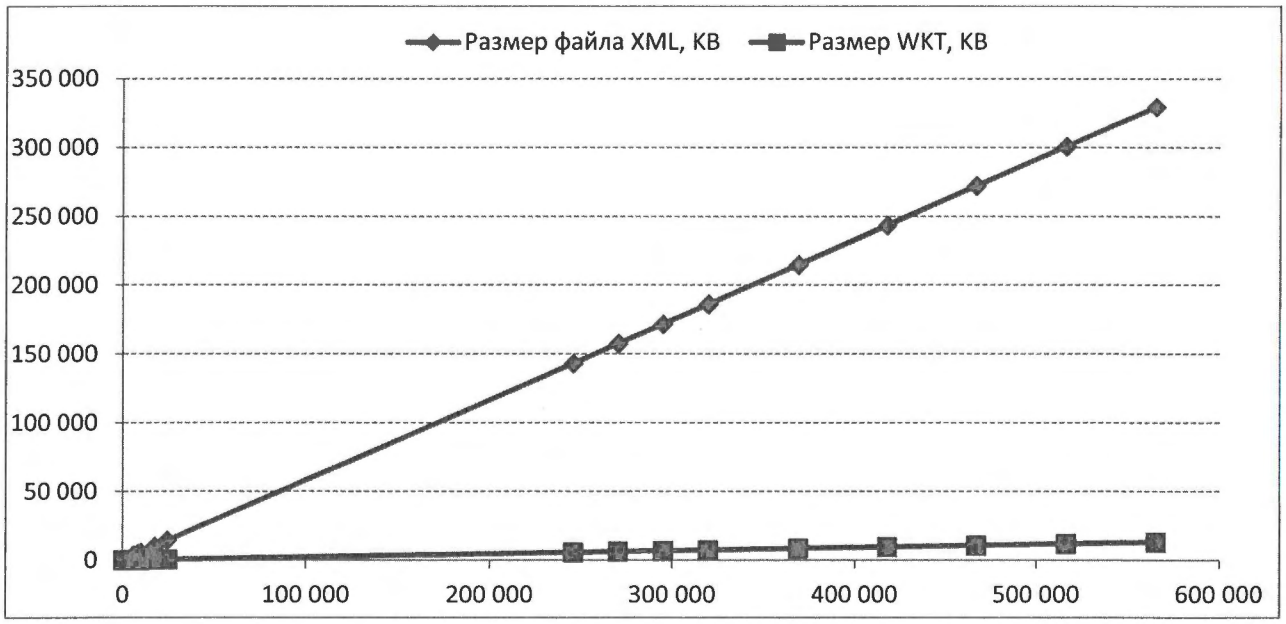

Рисунок 5. Графическая иллюстрация зависимости размера файла кадастрового XML документа и структуры пространственного объекта в формате WKT от количества точек составляющих пространственный объект

Выводы по третьей главе

Современные СУБД имеют эффективные инструменты работы с иерархическими данными. Это предоставляет новые возможности в реализации объектно-реляционного способа решения задачи ведения адресного реестра, учитывая при этом особенности формирования адреса в Российской Федерации.

В задачах, связанных с массовой загрузкой или выгрузкой пространственных данных в СУБД, целесообразно использовать WKT-нотацию консорциума OGC. Это позволяет повышать производительность процедур, облегчает поиск ошибок, так как WKT-формат обладает оптимальной лаконичностью и максимально адаптирован для машинной обработки и визуального чтения.

Совокупность кадастровых сведений, содержащих пространственные и атрибутивные данные, обладают основными признаками Big Data: имеют огромные объёмы данных, многообразие источников получения и представления данных в сочетании с непрерывным приростом, пополнением и обновлением. В самом ближайшем будущем это потребует разработки нового инструментария, основанного на нетрадиционных методах анализа и обработки, например, статистических методов.

Применение тяжеловесных синтаксических конструкций XML-документов, предназначенных для дистанционной передачи кадастровых сведений в регламенте АИС ГКН, контрпродуктивно влияет на эффективность процедур ведения кадастра. Использование WKT-нотации консорциума OGC позволит оптимизировать документооборот в кадастровых процедурах.

Комплексно рассматривая задачу повышения эффективности и оптимизации при ведении кадастра, приходится признать, что, концентрируясь только на применении даже самых совершенных программных и аппаратных средств, радикально решить задачу повышения

эффективности нельзя. Решающее значение все равно будет лежать на рационально продуманной теоретической и методологической базе, охватывающей как юридическую, так и топографо-геодезическую основу земельного кадастра в полном соответствии с реалиями современного государства.

В **заключении** кратко повторяются выводы по главам, подводится общий итог диссертационной работы, даются предложения о направлениях по дальнейшему усовершенствованию обработки различных видов кадастровой информации и развитию архитектуры информационных систем кадастра.

Оперативный доступ к актуальной кадастровой информации позволяет принимать наиболее эффективные управленческие решения на уровне государства, региона или муниципального образования, при необходимости воздействовать на рынок земли и недвижимости, обеспечивать гарантии при совершении сделок, планировать социальную и миграционную политику государства.

Без вовлечения кадастровой информации во все стороны экономической жизни невозможно обеспечить стабильный экономический рост и благосостояние. При этом важно помнить, что кадастр сам по себе не является абсолютной ценностью. Финансовые и трудовые затраты, связанные с администрированием кадастровой информации должны быть посильны хозяйствующим субъектам. Поэтому так важно для ведения кадастровой деятельности использовать наиболее оптимальный передовой инструментарий.

Целесообразно разработать систему требований при размещении заказов на выполнение кадастровых работ или оценки недвижимости для государственных и муниципальных нужд. Детализировать требования в технических заданиях таким образом, чтобы снижать финансовые и трудовые затраты при достижении заданных требований качества. Применять наиболее

совершенные инструменты и технологии. Активно вводить в профессиональное дискуссионное поле квалифицированное понимание процессов, критически важных при построении высокоэффективных информационных систем кадастра, помнить о важности опыта исследовательской конкуренции, полученного при разработке ГИС и СУБД ведущими отечественными и мировыми поставщиками IT-решений.

Список публикаций по теме диссертации

1 Маставичене, Т.В. Возможности современных баз данных для снижения стоимости проведения массовой кадастровой оценки / Маставичене Т.В. // 2014 г, Известия высших учебных заведений. Геодезия и аэрофотосъемка. – 2014. – № 5. – С. 88-91.

2 Маставичене, Т.В. Подходы к реализации адресного реестра с применением инструментария современных СУБД / Маставичене Т.В. // 2015 г, Известия высших учебных заведений. Геодезия и аэрофотосъемка. – 2015. – № 4. – С. 132-134.

3 Маставичене, Т.В. Влияние XML-формата на эффективность информационных систем кадастра / Маставичене Т.В. // 2015 г, Известия высших учебных заведений. Геодезия и аэрофотосъемка. – 2016. – № 1. – С. 41-43.