

Николенко С.В.¹, Ширшова В.Ю.²

ГЕОИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ СОЗДАНИЯ ЭЛЕКТРОННОЙ КАРТОГРАФИЧЕСКОЙ ОСНОВЫ И ОПРЕДЕЛЕНИЯ ФАКТИЧЕСКИХ ГРАНИЦ С ПОСЛЕДУЮЩИМ ВЫЯВЛЕНИЕМ НАРУШЕНИЙ ЗЕМЛЕПОЛЬЗОВАНИЯ НА ПРИМЕРЕ НЕСКОЛЬКИХ РАЙОНОВ ТУЛЬСКОЙ ОБЛАСТИ

АННОТАЦИЯ

В статье описывается опыт работы с большим объемом растровых и векторных данных при выполнении проекта компанией «Геоскан» на территорию Тульской области в рамках программы Национальной технологической инициативы. В качестве исходных данных отделу кадастра и инвентаризации объектов недвижимости были предоставлены: 6 млн снимков с беспилотных комплексов «Геоскан», 880 тыс. кадастровых планов территорий, данные Федеральной налоговой службы и Федеральной информационной адресной системы. При оценке объема рабочих данных, учитывая необходимость их корректной и оперативной обработки, стала очевидной неизбежность использования в рабочем процессе современных технологий работы с пространственными данными. В качестве геоинформационного клиента было выбрано открытое программное обеспечение QGIS, а анализ и хранение данных были организованы средствами объектно-реляционной СУБД PostgreSQL с расширением PostGIS для обработки пространственной информации. После внедрения в работу этих технических средств значительно повысилось качество процесса ввода атрибутивной и векторной информации за счет возможности осуществления оперативного контроля над действиями, проводимыми с данными. Помимо этого существует ряд преимуществ работы с кадастровыми данными с помощью системы управления базами пространственных данных: централизованная установка системы координат, обеспечение безопасности информации, упрощение процессов объединения данных и геометрической обработки, оптимизация пространства. Таким образом, был автоматизирован процесс выявления различного рода нарушений исходного материала и возникающих в процессе обработки, оптимизирована работа и устранены некоторые ее этапы, разработана структура хранения пространственной информации с использованием открытых программных продуктов.

КЛЮЧЕВЫЕ СЛОВА: кадастр, открытые ГИС, QGIS, PostgreSQL.

¹ Санкт-Петербургский государственный университет, Университетская наб., 7/9, 199034, Санкт-Петербург, Россия; ООО «Геоскан», ул. Шателена, 26А, 194021, Санкт-Петербург, Россия, *e-mail*: s.nikolenko@geoscan.aero

² Санкт-Петербургский государственный университет, Университетская наб., 7/9, 199034, Санкт-Петербург, Россия; ООО «Геоскан», ул. Шателена, 26А, 194021, Санкт-Петербург, Россия, *e-mail*: v.shirshova@geoscan.aero

Svetlana V. Nikolenko¹, Vera Yu. Shirshova²

GEOINFORMATION SUPPORT OF CREATION OF THE SINGLE ELECTRONIC CARTOGRAPHIC BASIS OF THE UNIFORM STATE REGISTRY OF REAL ESTATE AND DETERMINATION OF ACTUAL BORDERS OF LAND USE ON AN EXAMPLE OF SEVERAL DISTRICTS OF TULA REGION

ABSTRACT

Geoscan's experience of working with raster and vector big data during the National technological initiative program's project for Tula region. This project has been done by the department of cadaster and estate objects inventorization. As initial data were 6 mln photos by Geoscans' UAV, 880 thousand cadastral plans of territories, data of The Federal Tax Service and The Federal Information Address System. For successfully, correctly and operatively processing of such amount of data is need to use modern technologies of working with spatial data. QGIS was chosen as geoinformation client-program, but analyze and storing were organized by object-relational database management system Postgresql with PostGIS extension for spatial data. This helped us made quality of attributive and vector data entering and processing up at the expense of possibility of taking control for all of actions with data. Also, such method of working has some advantages. For example, a centralized coordinate system setting, a data security, a storage optimization, a simplification of geometry processing and data consolidation. In this way, a process of defects detection of source and output data was automatized, a working schema was optimized in generally and storage structure of spatial data was developed. In our project we used open-source programs.

KEYWORDS: cadastral data, open GIS, QGIS, PostgreSQL.

ВВЕДЕНИЕ

Наличие картографической основы необходимо во многих сферах государственного управления и контроля, экономической оценки и налогообложения недвижимого имущества, реализации государственных программ, направленных на комплексное развитие различных видов недвижимости [Асаул, 2010].

Возможность получить объективные сведения о недвижимом имуществе, разрешить проблему выявления неучтенных объектов, в том числе объектов самовольной постройки, и выявить увеличение площади земельных участков за счет самозахвата земель – все эти задачи эффективно решаются с помощью геоинформационных средств обработки и анализа информации.

Вопрос организации работы с кадастровой информацией в зарубежных исследованиях [Akakba, Filali, 2017; Hanusa et al., 2018] зачастую сводится к разработке структуры хранения данных без атрибутивной и геометрической выверки. В Российской Федерации существует реестр недвижимости, хранящий и предоставляющий (по запросам) кадастровые данные, однако как международные, так и отечественные наработки направлены на актуализацию информации в базе данных без проверки нарушений землепользования. Проект «Создание геодезически точной 3D модели типового региона России на основе данных беспилотной аэрофотосъемки и технологий ГЛОНАСС», связанный с повышени-

¹ St. Petersburg State University, Universitetskaya nab., 7/9, 199034, St. Petersburg, Russia; ООО "Geoscan", Shatelena, 26A, 194021, St. Petersburg, Russia, e-mail: s.nikolenko@geoscan.aero

² St. Petersburg State University, Universitetskaya nab., 7/9, 199034, St. Petersburg, Russia; ООО "Geoscan", Shatelena, 26A, 194021, St. Petersburg, Russia, e-mail: v.shirshova@geoscan.aero

ем качества государственных услуг в сфере государственного кадастрового учета недвижимого имущества и государственной регистрации прав на недвижимое имущество и сделок с ним, выполнялся на территории Тульской области в рамках программы Национальной технологической инициативы. Работа проводилась с целью определения фактических границ землепользования, а также выявления нарушений в использовании земель.

Специфика данного вида работ заключалась в большом объеме исходных данных (около 6 млн снимков и 880 тыс. кадастровых планов территорий), требующих тщательной предварительной обработки. Работа оператора по отбору из всего объема данных необходимой информации и последующего выявления различного рода ошибок исходного материала являлась трудоемким и ресурсозатратным процессом, в котором значительную роль играл человеческий фактор. Процесс подготовки исходных материалов для одного района Тульской области мог занимать более недели. Также при участии большого количества операторов ввода и обработки пространственной информации (более 40 человек) возникла необходимость грамотной организации рабочего процесса, вследствие которого были устранены некоторые виды «ручных» проверок, такие, как проверка соблюдения топологии и сборка результатов обработки. Все вышеперечисленные тезисы являются аргументом в пользу применения геоинформационных систем.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

В качестве исходных данных были предоставлены:

– Кадастровые планы территории (КПТ). Изначально поставляются в формате XML, который необходимо сконвертировать в shp-файл. Это происходит с помощью программного обеспечения SputnikCadastre, разработанного в ООО Геоскан. При конвертации исходного XML мы получаем на выходе 3 файла: точечные, линейные и полигональные объекты отдельно. А также таблицу с объектами без геометрии в формате *.csv.;

– Ортофотопланы, созданные на основе материалов, полученных в ходе аэрофото съемки с помощью беспилотного съемочного комплекса серии Геоскан и последующей обработки в ПО Agisoft PhotoScan;

– Данные Федеральной Налоговой Службы (ФНС).

Работа была условно разделена на три этапа:

1. Выявления нарушений в использовании земельных участков в населенных пунктах.

На этом этапе проводились дешифрирование участков по ортофотопланам в населенных пунктах и выявление самовольно занятых территорий.

2. Инвентаризация земель сельскохозяйственного назначения.

Выявление обрабатываемых земель, неиспользуемых, с разной степенью запущенности (луговая растительность, древесно-кустарниковая растительность и многолетние насаждения), также выявления нецелевого использования земель.

3. Дополнительные сведения.

а) выявление облагаемых и не облагаемых налогом территорий согласно данным ФНС;

б) анализ адресных баз Федеральной Информационной Адресной Системы (ФИАС) на предмет учета найденных путем камеральных выверок адресов.

в) ошибки исходных данных:

– дубли участков;

– пересечения границ участков.

Необходимость оперативного доступа к статистическим данным и оптимизированного процесса получения «корректных» векторных слоев на всех этапах работы привела к необходимости создания единой базы пространственных данных и использованию геоинформационных систем.

КПТ – это в первую очередь пространственные данные большого объема. Они требуют хорошо организованной инфраструктуры хранения. В противном случае использование и анализ таких данных будут крайне затруднены. Это было подтверждено путем постепенного совершенствования организации рабочего процесса.

В самом начале работы над проектом, когда на нем были задействованы 3-4 сотрудника, каждый имел набор shp-файлов, которые после завершения работы объединялись в один. На границах территорий работ двух (или более) сотрудников неизменно возникали пересечения, на устранение которых тратилось определенное количество времени. Отсутствовала возможность оперативной оценки состояния выполнения работ. Последний пункт был осложнен большой ролью человеческого фактора в контроле версий: сотрудник мог забыть передать последнюю версию своего рабочего файла вовремя. В дополнение сборные файлы результатов совместных усилий хранились локально на компьютере руководителя проекта, что далеко не лучший вариант для обеспечения сохранности информации.

Что касается анализа данных, он был возможен лишь на уровне примитивных характеристик, таких, как:

- количество объектов в слое;
- максимальное значение атрибута;
- минимальное значение атрибута;
- подсчет производных характеристик на основании уже имеющихся, с использованием несложных математических формул.

Такой небольшой набор обусловлен как минимум тем, что настольные ГИС не предназначены для подобных задач. Результат анализа в ГИС подразумевает создание какого-либо производного файла, что только усугубляет сложность структуры данных и увеличивает количество файлов, назначение которых нужно помнить и которые можно случайно удалить или потерять. Также при обновлении исходных материалов последует трудоемкий процесс обновления всех файлов, полученных на основании исходных.

Еще одна проблема – невозможность проведения анализа в пределах всего субъекта РФ целиком. Безусловно, можно создать совокупный файл из всех районов субъекта, но здесь стоит вспомнить о порядке чисел, которыми измеряется количество кадастровых участков в пределах даже одного муниципального района. Время обработки даже примитивных запросов для такого совокупного файла будет стремиться к бесконечности.

Допустим, что мы решили все же заняться геоинформационным обеспечением работ с кадастровой информацией, а также анализом полученных данных. В таком случае мы имеем полноценную базу пространственных данных, но организованную в виде файловой системы. Рационально доверить хранение нашей исходной, производной и аналитической информации системе управления базами данных (СУБД). К тому же это поможет значительно оптимизировать и автоматизировать различные процессы и этапы работы.

Система управления базами данных – это совокупность программных и языковых средств, обеспечивающих создание, управление и использование баз данных [Дейт, 2005]. Пространственные базы данных позволяют хранить геоданные и работать с ними. В качестве языковых средств в СУБД выступает структурированный язык запросов SQL (Structured Query Language). Он не отличается гибкостью, как языки программирования, имеет конечный набор команд и строгий синтаксис. Однако при умелом обращении позволяет эффективно оперировать данными [Моргунов, 2017].

Существует большое количество как проприетарных, так и свободных СУБД, имеющих расширения или же встроенный функционал для работы с пространственными данными. Выбор был сделан в пользу свободной объектно-реляционной СУБД PostgreSQL с расширением PostGIS, так как это одна из наиболее мощных современных систем управ-

ления базами данных, поддерживающая современные стандарты в области геоинформационных технологий и пространственных данных (Open Geospatial Consortium, OpenGIS) и обладающая обширным функционалом. Еще одним аспектом в пользу именно этой СУБД стала поддержка слоев PostGIS в настольной ГИС QGIS, в которой ведутся работы по определению фактических границ землепользования.

Рассмотрим подробнее некоторые технические преимущества работы с кадастровыми данными с помощью системы управления базами пространственных данных:

1. Система координат. Это основополагающий фактор любых пространственных данных. Кадастровые данные в РФ поставляются, как правило, в местных системах координат регионов. Параметров МСК нет в библиотеках, которые используют крупные настольные ГИС (например, библиотека proj4). Следовательно, вносить эти параметры необходимо самостоятельно. Мы сталкивались с ситуациями, когда по причинам, связанным с некорректно установленными параметрами СК, результаты векторизации оказывались смещены относительно реального положения объектов на местности. При работе через СУБД такие ситуации невозможны, так как параметры СК устанавливаются централизованно и не зависят от ПО-клиента.

2. Безопасность информации. Доступ к базе данных осуществляется с помощью логина и пароля. Пользователей можно наделять правами или же запрещать проводить им какие-либо действия с данными. Правила доступа устанавливаются для каждой конкретной таблицы. Например, можно раздать доступ на добавление новых объектов, удаление или редактирование существующих. А можно разрешить только просматривать данные, без возможности внесения изменений, или сделать таблицу невидимой для одного пользователя или группы пользователей.

3. Объединение данных. Если у двух наборов данных, хранящихся в базе, есть какой-либо общий атрибут, нам не составит труда их объединить. Сложно представить, как это можно реализовать без использования базы данных. Итак, у нас имеются следующие данные:

- таблица с земельными участками, а именно: их кадастровый номер, контур, площадь, геометрия и т. д.
- данные Федеральной налоговой службы: кадастровый номер, назначенный налог и т. д.

В данных налоговой службы не содержится информации о геометрии участков, но, объединив два набора по кадастровому номеру, мы можем получить производный слой, тем самым «визуализировать» уплату налогов. Таким образом, мы выяснили, что далеко не со всех участков, стоящих на кадастровом учете, платится налог. Не говоря уже о том, что далеко не все участки стоят на кадастровом учете.

Хорошим примером служит адресная база, с помощью которой мы находим местоположение ранее учтенных участков без установленных границ. Адресная база требует точной выверки, так как в результате получаются объединения данных из различных источников. Название какой-нибудь деревни или улицы может быть указано по-разному, так как источники не всегда соблюдают нормализацию. Для приведения всех адресов к общему виду мы сопоставляем их с данными ФИАС, которые подобно данным ФНС представляют из себя таблицы без геометрии участков. Из обоих наборов мы извлекаем уникальные значения населенных пунктов, улиц и обновляем нашу адресную базу согласно базе ФИАС.

Также имея возможность объединения таблиц, мы можем не хранить все атрибуты земельных участков для слоя, содержащего их определенные фактические границы, тем самым не увеличивая объем данных и не дублируя их. Нам достаточно иметь кадастровый номер и контур, с помощью которых мы свяжем эту таблицу с исходными кадастровыми данными и получим недостающие поля.

1. Геометрическая обработка. Как говорилось выше, исходные кадастровые данные содержат различного рода ошибки, такие, как дубли и пересечения земельных участков. Что

касается второго вида ошибок, пересечений, автоматически их можно разве что выявить. Исправление требует ручной проверки. Но процесс устранения дублей можно и нужно автоматизировать. Работая в настольной ГИС, мы можем автоматически найти дублирующиеся объекты, после чего необходимо рассмотреть каждую пару (или более) участков и, учитывая дату постановки на учет, удалить лишнюю геометрию. Работая через СУБД, мы можем не только выделить участки с идентичной геометрией, но и очистить таблицу от лишней информации, задав критерий отбора этих участков: дату постановки на учет.

Также стоит отметить, что любые геометрические операции над данными в базе данных происходят быстрее за счет механизма построения пространственных индексов.

2. Оптимизация пространства. Если какой-либо набор данных, будь то подсчет статистики, расчет площади и количества samozaxvatov, геометрический слой с выявленными пересечениями и т. д., является полностью производным от иной таблицы или нескольких таблиц, его можно создать в виде представления. Представление, на языке баз данных, – это таблица, являющаяся результатом выполнения SQL-запроса и генерируемая в процессе ее вызова. Одним словом, представление не занимает физического места на диске, всегда хранится в оперативной памяти. Более того, данные в такой таблице всегда оперативно меняются вслед за изменениями данных, на которых она основана, что является большим плюсом, так как исчезает необходимость постоянного обновления всех производных таблиц.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ИХ ОБСУЖДЕНИЕ

Технологическая схема создания электронной картографической основы представлена на рис. 1.

Из данных об участках без геометрии было выявлено около 36 тыс. ранее учтенных участков.

Объемы исходных данных составили порядка 880 тыс. земельных участков.

Всего было отвекторизовано более 880 тыс. участков.



Рис. 1. Блок-схема процессов работы

Fig. 1. Flowchart of work processes

Выявлено более 60 тыс. самовольно занятых земель у участков, прошедших кадастровый учет.

В исходных данных было установлено 67 тыс. пересечений границ участков и более 16 тыс. дублирующихся земельных участков.

Инвентаризация земель сельскохозяйственного назначения выявила 175 случаев нецелевого использования, а также позволила установить, что немногим более 23 тыс. участков используются, а около 25 тыс. являются заброшенными.

ВЫВОДЫ

За все время выполнения проекта технология решения задач и организация рабочего процесса неоднократно подвергались изменениям в целях нахождения максимально оптимизированного и универсального способа решения. Был проделан долгий путь: от «ручной» обработки до использования системы управления базами пространственных данных. Качество результатов обработки любых данных большого объема заметно возрастает при минимизации влияния человеческого фактора. Таким образом, стоит рассмотреть вариант работы с пространственными данными с помощью СУБД, если:

1. Ведется работа с объемными массивами географической и сопровождающей информации, а также с числом сотрудников больше одного;
2. Важна сохранность исходной информации и результатов работы, нужен контроль доступа к данным;
3. Необходима возможность оперативного обновления данных и доступа к ним;
4. Важно, чтобы в процессе анализа и обработки данных ничто не было случайно пропущено или не принято во внимание.

Опытным путем выяснили, что использование системы управления базами данных – это лучшее решение для автоматизации и оптимизации всех этапов работы, экономии времени и достижения максимальной эффективности использования данных.

СПИСОК ЛИТЕРАТУРЫ

1. *Асаул А.Н.* Экономика недвижимости. 2-е изд. СПб.: Питер, 2010. 624 с.
2. *Дейт К. Дж.* Введение в системы баз данных. 8-е изд. М.: Вильямс, 2005. 1328 с.
3. *Моргунов Е.П.* Язык SQL. Базовый курс. М., 2017. 257 с.
4. *Akakba A., Filali A.* Object-Relational Modelling and Establishment of a Generic Database for the Management and Monitoring of Urban Planning Permissions in the City of El-Eulma (Algeria). *Journal of Settlements and Spatial Planning*, 2017. V. 8. P. 139–145.
5. *Hanusa P., Peška-Siwika A., Szewczyk R.* Spatial analysis of the accuracy of the cadastral parcel boundaries, *Computers and Electronics in Agriculture*. January, 2018. V. 144. P. 9–15.

REFERENCES

1. *Akakba A., Filali A.* Object-Relational Modelling and Establishment of a Generic Database for the Management and Monitoring of Urban Planning Permissions in the City of El-Eulma (Algeria), *Journal of Settlements and Spatial Planning*, 2017. V. 8. P. 139–145.
2. *Asaul A.N.* The economy of real estate. 2nd ed. SPb.: Piter, 2007. 624 p. (in Russian).
3. *Date C.J.* An Introduction to Database Systems. M.: Williams, 2005. 1328 p. (in Russian).
4. *Hanusa P., Peška-Siwika A., Szewczyk R.* Spatial analysis of the accuracy of the cadastral parcel boundaries, *Computers and Electronics in Agriculture*. January, 2018. V. 144. P. 9–15.
5. *Morgunov E.P.* The SQL language. Basic course. Moscow, 2017. 257 p. (in Russian).