

ИНСТРУМЕНТЫ МОНИТОРИНГА РЫНКА НЕДВИЖИМОСТИ МОСКОВСКОГО РЕГИОНА

УДК 332.851, 339.13.017, 339.132, 338.001.36

Алексей Юрьевич Юшин, аспирант кафедры экономики городского хозяйства Московского городского университета управления Правительства Москвы (МГУУ Правительства Москвы).
Тел.: (903) 503-91-77
Эл. почта: yushin.alexey@yandex.ru

Статья посвящена мониторингу – наиболее общему и трудоемкому этапу анализа рынка недвижимости. Обозначаются основные проблемы и предлагаются инструменты для их решения.

Ключевые слова: мониторинг рынка недвижимости, источники информации о рынке недвижимости, инструменты сбора и обработки информации о рынке недвижимости.

Alexey Yu. Yushin, Post-graduate student, the Department of the Economy of municipal services, Moscow Metropolitan Governance University (MMGU).
Tel.: (903) 503-91-77
E-mail: yushin.alexey@yandex.ru

TOOLS FOR MONITORING THE REAL ESTATE MARKET IN THE MOSCOW REGION

Monitoring – the most common and time-consuming step of analyzing the real estate market. Designated the main problems and offers tools to deal with them.

Keywords: monitoring real estate market, sources of information about the real estate market, real estate valuation, tools for collecting and processing information about the real estate market.

1. Введение

Известный исследователь рынка недвижимости, Стерник Г.М., выделяет мониторинг как наиболее общий элемент анализа рынка, изначально подразумевая сбор и обработку информации [2] и позднее выделяя 4 этапа общего анализа рынка недвижимости (сбор и обработка данных, мониторинг, исследование и прогнозирование) [1]. Первые два этапа отмечаются как наиболее трудоемкие и необходимые. Здесь под мониторингом уже понимается статистическая обработка, анализ и оценка показателей. Не будем так широко трактовать это понятие и сконцентрируемся на решении задач, связанных с непрерывным сбором и обработкой информации о рынке. Попробуем обозначить типичные проблемы и предложим универсальные инструменты для их решения. Универсальность подразумевает интерфейс взаимодействия, как человеческий, так и машинный в виде интерфейса прикладного программирования для упрощения интеграции в любую информационную систему. Часто в названии таких инструментов присутствует устоявшаяся аббревиатура API (Application Programming Interface).

2. Источники исходных данных

Наиболее сложной проблемой является получения сведений о спросе. На некоторых площадках объявлений можно увидеть соответствующую рубрику, но ее нельзя рассматривать всерьез, поскольку число объявлений мало и они не всегда достоверны. Недостаточность информации можно компенсировать используя статистические инструменты контекстных рекламных систем «AdWords» и «Direct». С помощью инструмента подбора ключевых слов можно узнать количество запросов в месяц по определенным ключевым словосочетаниям, например, «куплю квартиру в ховрино». Расширив список запросов, можно собрать количественные показатели спроса, сколько было запросов, по каким типами недвижимости, в каком районе и увидеть динамику изменений. В идеале, полнота сведений о спросе могла бы революционно улучшить качество оказания услуг на рынке. Зная потребности конкретных покупателей, можно автоматически выстраивать цепочки. Сейчас эти задачи решаются профессиональными участниками рынка самостоятельно, централизованно эта информация не собирается.

С предложением ситуация обратная, исходных данных очень много. Ее источниками являются как профессиональные инструменты, например «Winner»[5] и «Cian»[4], так и множество сайтов (порталов, площадок), предоставляющих сервис по размещению объявлений, например «Из Рук в Руки», «Недвижимость и Цены», «Где Этот Дом» и т.п. Часто объявления уже содержат некоторую информацию об объектах (зданиях) из базы БТИ, год постройки, количество лифтов, планировки. Предложения рынка можно рассматривать также как источник информации о субъектах, поскольку в объявлениях часто есть ссылка на агентство. Некоторые информационные ресурсы делают акцент на социальном взаимодействии, стимулируя профессиональных участников к ведению собственных страничек, блогов.

Важно отметить особенности программного обеспечения для профессиональных участников (агентств, риэлторов) и общую тенденцию на рынке программного обеспечения. Традиционная концепция установки программы на компьютер уходит в прошлое, ее заменяют онлайн сервисы. Главное преимущество нового подхода в том, что основные вычисления, такие как поиск объектов, осуществляются на удаленном сервере. Пользователь, открывая сайт в браузере, фактически работает в режиме «тонкого» клиента. Базы «Winner» и «Realty», можно считать исключениями, они последовательно идут по пути установки программы под Windows. Недостатки тако-

го подхода: необходимость в обновлении, технической поддержке, как со стороны агентства, так и самой базы. Встроенный инструмент экспорта объектов в Виннер, имеет ограничение и его сложно считать серьезным источником информации. Общая тенденция к раскрытию информации, появление тех же объявлений в других источниках, привела к появлению онлайн версии базы Winner («Журнал Собственник»), где обнаруживаются те же предложения, что и в программе, бесплатно, но с рекламными блоками. У «Realty» также есть сайт и поисковик объявлений, но при разработке упор был сделан на затруднении автоматического импорта, посредством некорректного кода разметки страниц. Корректное отображение такого сайта в браузере уже сложная задача, вероятно с этим связано отсутствие развития. С другой стороны «Realty» единственная база, которая официально предлагает купить архивы предложений.

База Циан [4], пожалуй, второй по значимости инструмент размещения и поиска объявлений в московском регионе. Пользуется особой популярностью среди специалистов, занимающихся арендой. Отличается строгим контролем качества, благодаря «обратной связи» и ручной проверке объявлений.

В жестких конкурентных условиях важна полнота, качество представляемой информации и удобство поиска. Мы наблюдаем соревнование, кто проще и с максимальными удобствами удовлетворит потребности профессиональных и самостоятельных участников рынка. В последнее время появилось множество «агрегаторов», альтернативных поисковых баз, информационных систем, цель которых – собрать как можно больше предложений из разных источников и предоставить дополнительные возможности, например «afu.ru» предоставляет автоматизированный сервис по размещению объявлений на большое количество площадок. Наиболее агрессивно в этом направлении продвигается «Яндекс», пытаясь реализовать свой лозунг «Найдется все», создавая онлайн площадки

для взаимодействия продавцов и покупателей в сфере товаров массового потребления, автомобилей и недвижимости. «Яндекс-недвижимость» относительно молодой проект, но у него есть серьезное конкурентное преимущество – любой поиск начинается с поисковой системы. На начальном этапе «Яндекс» сотрудничал исключительно с агрегаторами, которые быстро поняли, что получается «игра в одни ворота» [9].

3. Инструменты сбора информации

Как правило, типичный информационный ресурс старается с одной стороны затруднить автоматизированный сбор информации, а с другой, не брезгает копированием объявлений конкурентов [7]. Обозначим основные проблемы автоматизированного получения информации с онлайн ресурсов и способы их решения.

- Отсутствие четкой структуры, некорректность кода разметки страниц. Если живой посетитель легко понимает содержание страницы, то для программы небольшое изменение структуры может оказаться фатальным. Необходима адаптация к изменяющейся структуре отображаемой информации.

- Ограничения на общее количество запросов, их скорость (количество в единицу времени). Решением является использование промежуточных (проxy) серверов, сведения о которых можно получить посредством API [6].

- Наличие так называемых «капчей» (captcha), или представление контактных сведений в виде картинок. Существуют автоматизированные сервисы по распознаванию с API.

- Наиболее трудоемкий способ противодействия — использование скриптов, программно отображающих информацию на экране. Например, сайт Росреестра (содержит информацию о фактических сделках) и Официальный сайт Российской Федерации для размещения информации о проведении торгов (содержит раздел, «реализация имущества должников»). Это важный источник

информации о субъектах рынка – должниках, которые не смогли расплатиться за квартиру или иной объект недвижимости. «Разобрать» или «распарсить» (от англ. parse – разбор) подобные сайты традиционными методами, затруднительно. На этот случай есть решения в виде «безголовых браузеров» [10]. У них отсутствует графический интерфейс и для программного взаимодействия реализован API.

Можно сделать вывод: какие бы «изошренные» методы не использовались для затруднения получения информации, найдется не менее эффективное решение. Есть еще один фактор который способствует информационной открытости. По мере роста любого агентства возрастает количество одновременно продаваемых или сдаваемых объектов, что приводит к увеличению нагрузки на персонал, который вынужден добавлять один и тот же объект одновременно на сайт агентства и во множество информационных ресурсов. Как следствие, развиваются инструменты автоматического экспорта-импорта объектов. Общая схема следующая. Агентство один раз добавляет объект в свою базу, как правило на сайт, далее программно формируется файл выгрузки в формате xml, который периодически проверяет робот информационной базы и, таким образом, осуществляется перенос объявлений. Структурно, обычная html страничка и xml файл выгрузки похожи, с той разницей что структура документа xml четко формализована и специально адаптирована для описания объекта недвижимости, а значит, вероятность ошибок сведена к минимуму.

В ближайшее время мы будем наблюдать войну форматов, поскольку со стороны площадки объявлений гораздо проще предоставить гибкие возможности импорта (часто можно договориться о передаче объектов в формате конкурентов), чем произвольному агентству реализовывать множество отдельных выгрузок под каждую базу. Таким образом, в ближайшем будущем сбор данных о предложениях рынка будет осуществляться автома-

тически, в стандартизированном виде, непосредственно от профессиональных участников, агентств и даже риэлторов, которые, в отличие от существующих информационных ресурсов, заинтересованы в как можно более широком распространении своих эксклюзивных предложений.

4. Инструменты обработки информации

4.1. Приведение контактных данных к нормальному виду.

При обработке предложений рынка, надо быть готовым к тому что контактные сведения могут быть представлены в произвольном текстовом виде, не только в специальных полях, но и, например, в описании объекта. Некоторые доски объявлений показывают телефонные номера в виде картинок. Для решения задачи выявления номеров в стандартном виде из произвольного текста идеально подходит библиотека [8].

4.2. Вычисление стоимости объекта в валюте.

После кризиса 2008 года, большинство сделок на рынке, особенно в эконом-классе, осуществлялись в рублях. В связи с резким падением курса рубля в конце 2014 года возникают опасения, что доллар может вновь вернуться на рынок недвижимости. Участники рынка наверняка захотят зафиксировать свою прибыль или стоимость жилья в валюте. Чтобы обоснованно судить о состоянии рынка недвижимости в условиях финансовой нестабильности, необходимо учитывать не только рублевую стоимость, а также курсы валют, причем на дату создания объекта. Сайт «Центрального Банка» предоставляет удобный инструмент получения курса валют посредством html формата.

4.3. Нормализация адреса.

Недвижимость подразумевает координаты, адрес, границы, поэтому невозможно представить автоматизированную обработку первичной информации без инструментария геоинформационных сис-

тем. Как нормализовать адрес, составленный в свободной форме, т.е. разбить его на составляющие (город, район, улица, дом, метро)? Для этого подходят инструменты, предоставляемые поисковыми системами, например Google Geocoding Api [12].

4.4. Выявления дубликатов и недостоверных сведений

Одно и тоже объявление, как правило, можно обнаружить в разных информационных базах или онлайн-ресурсах. Они могут быть с разными значениями цен, в них могут быть ошибки. Иногда основной мотив – чем больше объявлений, тем больше звонков, при этом объявление может быть «уткой» или заимствовано у реального продавца. Существуют алгоритмы [3] выявления дубликатов, но в автоматическом режиме эта проблема решается лишь частично, поскольку параметры недостоверного объекта могут выглядеть корректно. Профессиональные информационные базы пытаются бороться с этим явлением, в том числе обзванивая объявления и с помощью «обратной связи».

4.5. Выявление субъектов рынка

Типичное объявление о продаже недвижимости, можно представить как граф из 2 или 3 вершин, которые обозначим: «объект», «контакт» и «агентство». Нормализация контактных сведений и выявление дубликатов позволит осуществить склейку этих простейших графов. Какой субъект рынка скрывается за контактом? Риэлтора можно вычислить по большому числу связанных объектов и часто агентство явно указано. Риэлторы могут быть в составе группы, которая является структурным элементом большинства агентств. Можно предложить простые алгоритмы поиска связанных контактов через промежуточные узлы. Похожим образом можно вычислить еще одного субъекта рынка – индивидуала, «контакт» которого связан с «агентством», но не прослеживается связь с другими «контактами». Недвижимость – не часто покупаемый или продаваемый то-

вар, поэтому можно предположить что за уникальными «контактами» скрывается собственник.

Как реализовать эти алгоритмы на практике? Относительно недавно появились специальные графовые базы данных, которые оперируют не таблицами со строго заданными полями, а гибкой структурой в виде вершин графа и связями между ними. Neo4j [11] была одной из первых созданных графовых СУБД. С тех пор это направление бурно развивается и в настоящее время таких СУБД существуют десятки. Немалую роль в этом сыграло развитие социальных сетей, где графовая постановка задачи требовала новых подходов к развитию программного обеспечения.

5. Заключение

Медленно, но верно, рынок недвижимости России движется в сторону информационной открытости. Не малую роль в этом играет развитие инструментов, облегчающих сбор и обработку информации. Универсальность этих инструментов позволяет с легкостью задействовать труд многих разработчиков и ученых для решения частных проблем, связанных с мониторингом. В общем, мониторинг остается затратным этапом анализа, поскольку требует специфических навыков, применения разнообразных инструментов, причем непрерывно и в течении длительного времени. Однако, по мнению автора, соблюдены необходимые и достаточные условия для создания комплексного инструмента мониторинга, который бы позволил посредством интерфейса прикладного программирования получать нормализованные сведения о рынке недвижимости.

Литература

1. Стерник Г.М., Стерник С.Г. Анализ рынка недвижимости для профессионалов. – М.: Изд. «Экономика», 2009. – стр. 250–255.
2. Методология анализа рынка недвижимости URL: <http://realtymarket.ru/metodi-eskie-materiali/METODOLOGIYA-ANALIZA-RINKA-NEDVIJIMOSTI.html> (дата обращения 23.12.2014).

3. Стерник С.Г., Стерник Г.М. Основные положения методологии сбора и обработки информации о рынке недвижимости в интересах оценочной деятельности (Доклад на XII Международном конгрессе оценщиков «Информационно-аналитическое обеспечение индивидуальной и массовой оценки недвижимости» 16 декабря 2010 г., г. Москва).

4. База Циан URL: <http://www.cian.ru> (дата обращения 23.12.2014).

5. База Winner URL: <http://www.baza-winner.ru> (дата обращения 20.12.2014).

6. Прокси-листы URL: <http://hideme.ru/proxy-list/> (дата обращения 20.12.2014).

7. Статья «Виннер пойман на воровстве», URL: <http://www.realto.ru/journal/articles/winner-pojman-na-vorovstve> (дата обращения 23.12.2014).

8. Библиотека libphonenumber URL: <https://github.com/googlei18n/libphonenumber> (дата обращения 23.12.2014).

9. Я.Недвижимость тестирует карточки объектов. Трафик упадёт? URL: <http://roem.ru/2014/08/11/>

yan104825/?c (дата обращения 23.12.2014).

10. Браузер Phantom URL: <http://phantomjs.org> (дата обращения 23.12.2014).

11. Графовая база данных Neo4j URL: <http://neo4j.com/> (дата обращения 23.12.2014).

12. Google Geocoding Api URL: <https://developers.google.com/maps/documentation/geocoding/?hl=ru> (дата обращения 23.12.2014).

References

1. Sternik G.M., Sternik S.G. Real Estate Market Analysis for Professionals. –М.: Экономика, 2009, – р. 250–255.

2. The methodology of the analysis of real estate market. URL: <http://realtymarket.ru/metodi-eskie-materiali/METODOLOGIYA-ANALIZA-RINKA-NEDVIJIMOSTI.html> (last visited 23.12.2014).

3. Sternik S.G., Sternik G.M. Main provisions of the methodology for collecting and processing information about the real estate market in the interest of evaluation activity (Report on the XII International Congress of Appraisers “Information and analytical

support of individual and mass valuation of real estate,” December 16, 2010, Moscow).

4. Cian URL: <http://www.cian.ru> (last visited 23.12.2014).

5. Baza Winner URL: <http://www.baza-winner.ru> (last visited 23.12.2014).

6. Proxy-lists URL: <http://hideme.ru/proxy-list/> (last visited 23.12.2014).

7. Article «Winner caught stealing» URL: <http://www.realto.ru/journal/articles/winner-pojman-na-vorovstve> (last visited 04.05.2014).

8. Libphonenumber Library URL: <https://github.com/googlei18n/libphonenumber> (last visited 23.12.2014).

9. Yandex-Realty testing card facilities. Traffic will fall? URL: <http://roem.ru/2014/08/11/yan104825/?c> (last visited 23.12.2014).

10. Phantom Browser URL: <http://phantomjs.org> (last visited 23.12.2014).

11. Graph database Neo4j URL: <http://neo4j.com/> (last visited 23.12.2014).

12. Google Geocoding Api URL: <https://developers.google.com/maps/documentation/geocoding/> (last visited 23.12.2014).