1. Федеральное государственное автономное образовательное учреждение высшего профессионального образования «Национальный исследовательский университет «Высшая школа экономики»

Факультет экономических наук

Образовательная программа «Экономика»

##### БАКАЛАВРСКАЯ ВЫПУСКНАЯ РАБОТА

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Название темы)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Выполнил

Студент группы № \_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Ф.И.О.)

Научный руководитель \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(должность, степень, звание, Ф.И.О.)

Консультант

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(должность, степень, звание, Ф.И.О.)

## Введение

В настоящее время сеть Интернет становится предметом бытового использования всё большего числа людей. И вместе с этим ростом числа пользователей, интернет становится все более публичным местом. Хотя история возникновения и использования интернета имеет уже довольно долгую историю, настоящий рост числа пользователей настолько большой, что возникает ощущение, что он появился совсем недавно. Этот процесс, в свою очередь, вызывает усиление государственного регулирования сети, которое приводит к появлению «цивилизованного» интернета. По этой причине тем лицам, которые желают обходить данные ограничения, приходится применять технологии анонимизации.

Одной из таких анонимизационных технологий является сеть TOR - «сеть внутри сети», которая позволяет использовать посещать web-страницы, надёжно скрывая собственный IP-адрес. Изобретение TOR настолько облегчило и повысило уровень безопасности при сокрытии своей личности, что позволяет не только эффективно обходить заинтересованным лицам налагаемые государством ограничения, но и привело к возникновению новых явлений, таких как зарождение анонимных торговых площадок, на которых возможно покупать и продавать запрещенные предметы и вещи.

Первые анонимные торговые площадки начали появляться в англоговорящих странах, однако некоторые российские торговые площадки имеют почти такую же долгую историю. Более того, согласно мнению журнала Wired[[1]](#footnote-1) в настоящее время самая долгоживущая площадка действует на территории России. Вместе с этим, в отличие от западных площадок, российские до сих пор обделены вниманием исследователей.

В данной работе поставлена цель собрать необходимые данные о функционировании российских торговых площадок, привести их к виду, удобному для анализа, а также произвести непосредственный анализ структуры и характеристик площадок. В настоящее время существует несколько активных российских площадок, однако для исследования были отобраны только две крупнейших из них. Одна из них буквально так и называется: Russian Anonymous Market Place (RAMP), имя другой – HYDRA. Более того, так как объём работ, требуемый для полноценного анализа сразу обоих площадок слишком велик, то HYDRA будет уделено второстепенное внимание. Впоследствии окажется, что основной и практически единственный предмет торговли на обоих площадках – это запрещенные психоактивные вещества. Для площадки RAMP окажется возможной оценка объёма сделок в денежном выражении. В дальнейшем поставлена цель построить индекс активности на площадке, который будет отражать динамику в объёме продаж товаров для RAMP. Далее будет протестировано наличие взаимосвязи между динамикой этого индекса и динамикой курса биткоина, в котором производятся транзакции на этой площадке.

Таким образом, ключевыми в данной работе являются следующие вопросы:

* Какие товары продаются на российских анонимных торговых площадках (в частности, на RAMP и HYDRA)? Какова их относительная структура?
* Какова географическая структура транзакций на данных площадках?
* Каковы цены на разные типы товаров?
* Каков объем сделок в денежном выражении на RAMP?
* Какой была динамика активности на RAMP за время наблюдения за площадкой?
* Есть ли какая-либо прогнозная сила у курса биткоина по отношению к рублю на эту динамику?

Для ответа на эти вопросы были собраны данные в виде непосредственного содержимого сайтов анонимных площадок и была разработана методология агрегирования информации из этого содержимого и формирования оценок.

Данная работа является одним из первых исследований, затрагивающих деятельность русскоязычной части даркнета. Совсем недавно в России появился новый сегмент теневой экономики, который развивается также хорошо, если не лучше, как в остальном мире. Вместе с тем, закономерности его функционирования до сих пор мало известны и представляют большой научный интерес, поскольку не исключено, что функционирование анонимных торговых площадок наиболее близко приближено к тем предпосылкам, на которых основано большое число базовых микроэкономических моделей. С другой стороны, данный сектор выступает сильным контрастом со всей российской экономикой, если взглянуть на него через призму институциональной теории. Даркнет — особое место: с одной стороны, в нём нет никаких искусственных барьеров для начала предпринимательской деятельности, сдерживающих частную инициативу. С другой стороны, между анонимными участниками практически всегда отсутствует доверие, в связи с чем данный сектор теневой экономики мало нуждается в традиционных институтах, но сильно зависит от других, отчасти новых. Более того, даркнет является хорошим примером экономики, основанной на информационных сигналах. Всё это показывает, что изучение анонимных торговых площадок способно принести новое научное знание в разных областях экономики.

## Обзор литературы

Исследование площадок для торговли в сети TOR (которую также называют даркнетом) началось практически сразу же после их возникновения. Самый интересный предмет исследования – это площадка Silk Road, крупнейшая площадка, ставшая известная обывателю после того, как была успешно взломана американскими правоохранительными органами. Среди исследователей, Ванхоут и Бингхэм (van Hout, Bingham, 2013) были одними из первых, кто начал изучение этой площадки. Их первая работа дала обзор внутренних мотивов покупателей, которые обуславливают их выбор между традиционными способами нелегального приобретения товаров и анонимными площадками даркнета в пользу последних. Также они дали подробное описание Silk Road как площадки в целом, из которого следовало, что торговля ПАВ занимает на ней одно из ключевых мест. Через год авторы продолжили исследование (van Hout and Bingham, 2014), переведя внимание на сторону продавцов ПАВ. Анонимное интервью 10 продавцов показало, что большинство из них принимало наркотики самостоятельно, и вело свою деятельность из-за желания предоставить качественный и безопасный и надежный товар другим пользователям. Причины, по которым они выбирали даркнет, совпадало с причинами покупателей, выясненными в прошлом исследовании (van Hout and Bingham, 2013): даркнет даёт больше безопасности и больший обхват рынка по сравнению с уличной нелегальной торговлей.

Другим примером исследования, основанного на анонимном интервью является работа (Barratt, Ferris, and Winstock, 2014), в которой опрашивались лица, употреблявшие наркотики на территории США, Великобритании и Австралии. Авторов интересовало, насколько сильно среди пользователей даркнета было распространено знание о Silk Road, а также структура спроса на ПАВ и мотивы выбора в пользу анонимных площадок. Исследование показало, что только половина респондентов когда-либо слышало о Silk Road и, более того, только 11% из них когда-либо использовало её. С другой стороны, авторы не задались вопросом, какими площадками пользователи пользовались вместо Silk Road. Среди наркотиков, респонденты, использующие эту площадку, чаще всего искали MDMA, каннабиноиды и LSD.

Данные работы не используют информацию c самого Silk Road и не применяют статистических методов к большим базам данных, благодаря чему они не могут дать более конкретной информации о структуре площадки, чем информация, которую способны воспринять непосредственные пользователи. Вместе с тем, достаточно большой ряд вопросов (конкретные пропорции разных типов товаров, их средние цены, география транзакций и т.д.) довольно трудно изучать подобным методом. Другие исследователи пошли иным путем: создание офлайн-копий торговых площадок (web crawling) и работа с полученными «снимками», или «слепками», которые представляют из себя загруженные html-страницы основных разделов сайта. Работа (Christin, 2012) была практически первой, в которой использовался подобной метод. Автор делал «снимки» на протяжении восьми месяцев между концом 2011 и 2012 годами. Ключевыми вопросами в статье были типы продаваемых товаров, а также доходы продавцов и администрации сайта. На основе загруженных web-страниц были сделан вывод, что подавляющей большинство продаваемых товаров было наркотиками (из 20 наиболее продаваемых категорий на сайте 18 были различные виды ПАВ, остальные две категории – это наличные деньги и книги), и что месячный объем сделок был на уровне 1.2 млн. $, что давало администрации площадки доход в размере 92 000$. Также было показано, что среди продавцов наблюдалась высокая текучесть: большинство из них было активно на площадке только в течение 3 месяцев, за исключением 112 продавцов, которые были представлены на Silk Road на протяжении всего периода загрузки снимков площадки. Кроме этого, автор произвел подсчёт расчёт динамики ежедневного объема сделок на площадке, для чего использовалась информация о числе оставленных за период времени отзывов покупателей о качестве товара. Далее, в силу того, что отзыв оставляется не в момент покупки, а в момент получения товара, автору пришлось применять сглаживание методом скользящего среднего при подсчете динамики объема сделок.

Перед тем как Silk Road был взломан ФБР США, Aldridge и Decay-Hetu (Aldridge, Decay-Hetu, 2014) продолжили подход Christin, чтобы проверить, как озвученные в прошлой работе характеристики торговли изменились к сентябрю 2013 года. На тот момент, как утверждают авторы, месячный объем сделок вырос в 7 раз (до 7.4 млн. $), а количество продавцов удвоилось (1084 учетный записи продавцов). Более того, авторы утверждают, что большая доля совершаемых транзакций совершаются на уровне business-to-business, поскольку объемы таких сделок слишком велики для персонального потребления. Хотя конкретное значение доли таких сделок остаётся не выявленным.

После взлома Silk Road в октябре 2013 года, его реинкорнация, названная Silk Road 2 возникла спустя месяц, в ноябре. Вместе с тем, не смотря на то, что прецедент успешного взлома и деанонимизации должен был оттолкнуть пользователей от использования анонимных площадок, освещение этой темы в СМИ сделало их существование известным большему числу людей, что могло усилить рост числа их активных пользователей. В 2014 году Dolliver (Dolliver, 2015) предпринял попытку повторить работу Christin (Christin, 2012) на новых данных, чтобы выяснить как оба этих фактора сказались на популярности Silk Road 2. В итоге в работе утверждалось, что ПАВ перестали быть основным товаром на площадке (их доля среди предлагаемых товаров оказалась 19% против 53.4% у Christin (Chrtistin, 2012)). Более того, администрация сайта возможно намеренно завышала реально число товаров, выставленных на продажу, так как они хотели произвести впечатление на покупателей, как будто Silk Road 2 до сих пор является не менее популярной торговой площадкой, чем другие, которые на тот момент уже также успели появиться. Наконец, число активных продавцов, как утверждает автор, сократилось с 550 у Chrtistin (Chrtistin, 2012), до 200.

Получившиеся результаты показались довольно противоречивым для других исследователей, что вызвало череду дискуссий и критики. Munksgaard, Demant и Branwen (Munksgaard, Demant and Branwen, 2016) сделали частичную репликацию исследования (Dolliver, 2015), в которой показали что скорее всего последний имел в распоряжении только ограниченный сегмент сайта Silk Road 2, что и вызвало такие утверждения автора. Для работы были использован архив площадок даркнета (Branwen et al., 2015), который включает 9 снимков Silk Road 2 за тот же период, в котором был сделан снимок из работы (Dolliver, 2015). Несмотря на то, что все эти снимки не были до конца полными (то есть не включали в себя все html-страницы, которые планировалось загрузить с сайта, вследствие разных сбоев, нередких для даркнета), они включали в себя не менее чем 5588 активных товаров на продажу (против 1834 в (Dolliver, 2015)), при этом большинство включало более 9000 товаров. Такое несоответствие интерпретировалось как очень ограниченный снимок Silk Road 2, произведенный в работе (Dolliver, 2015). Ввиду этого, авторы предлагают более аккуратное отношение к составлению снимков, предоставление более детальной технической информации по ним, а также специальной методологии, которая на основе такой информации способна оценить качество этих снимков.

Сам архив (Branwen et al., 2015), из которого были взяты данные 9 снимков, также полезен для данной выпускной квалификационной работы, поскольку в нем содержалось описание техники произведения таких снимков.

Данная работа близка по смыслу к работе (Chrtistin, 2012), в ней также будет сделан обзор структуры российских площадок, будет оценён денежный объём совершаемых сделок и также будет построен индекс активности покупателей и продавцов на площадке. При этом, его построения использовался фильтр Калмана, а точнее его реализация в R (пакет KFAS), которая описана в статье автора программного обеспечения (Helske, 2016). С принципами работы фильтра Калмана можно ознакомиться в (Durbin J., Koopman S.J. 2012).

В ходе работы необходимо было также получить представление о работе площадок, для чего были использованы статьи из русскоязычных СМИ, которые можно увидеть в списке использованной литературы.

## Что из себя представляют анонимные торговые площадки

Анонимные торговые площадки возникли как альтернатива традиционному способу приобретать и продавать запрещенный товар, в котором личная безопасность каждой стороны имеет более высокий уровень, поскольку участники рынка не знают друг друга и поэтому не могут оказывать содействие правоохранительным органам в случае, если были ими пойманы. Более того, по сравнению с традиционными способами, охват рынка (а, следовательно, и выбор на нём) на анонимных площадках гораздо шире, поскольку продавец может, благодаря сохранению анонимности предлагать свой товар практически неограниченному числу покупателей. Покупатели, в свою очередь, по тем же причинам гораздо легче получают доступ к информации о предлагаемом товаре у разных продавцов. По этим причинам, анонимные торговые площадки обладают рядом преимуществ, что привело к их быстрому распространению после возникновения.

Любую транзакцию можно разделить на 4 этапа: поиск контрагента; переговоры и заключение сделки; перечисление денег продавцу; доставка товара покупателю. И на каждом этапе необходимо обеспечить приемлемый уровень анонимности.

Способ обеспечения анонимности на 1-ом этапе сделки основан на технологии TOR. TOR – является закрытой сетью серверов, которые перенаправляют запросы пользователей из обычного интернета. Каждый запрос, проходя через эту сеть, выходит из неё с измененным ip-адресом – таким образом установить компьютер, с которого этот запрос был отправлен, становится практически невозможным. Конечно, само использование сети не является гарантией анонимности. Если пользователь будет вести себя не осторожно, то риск раскрытия анонимности становится довольно большим. Однако в целом, для рядового пользователя этого достаточно, чтобы обеспечить себе безопасность, в то время как для более значимых участников торговли доступны дополнительные средства, такие как VPN-каналы, специализированные операционные системы и так далее. Более того, желающие иметь полностью безопасный доступ в интернет, могут использовать отдельный компьютер и не оставлять на нём личной информации. В конце концов, TOR – не единственное средство, однако оно достаточно удобное для рядового пользователя, поскольку здесь не требуется глубоких знаний и серьёзных затрат.

При успешном завершении поиска контрагента, следующим этапом сделки являются переговоры о цене, количестве и других условиях. Здесь возникает проблема, вызванная тем, что несмотря на анонимность, каждая сторона должна иметь гарантии, что общение происходит в приватном режиме, когда условия сделки не становятся известными кому-либо кроме участников переговоров. То есть необходимо передавать сообщение абсолютно неизвестному лицу таким образом, чтобы доступ к нему имело только оно. Для этого используются специальные каналы связи, например, общение по любому открытому каналу (электронная почта, skype и т.д.) с применением GPG-ключей для шифрования информации. Используется пара связанных ключей, один из которых известен только владельцу и позволяет дешифровать сообщение, второй – открытый ключ, который пересылается предполагаемому собеседнику и необходим для того, чтобы собеседник мог зашифровать отправляемое сообщение. Благодаря такому способу, установление контакта и осуществление связи возможно в условиях, когда каждое сообщение между двумя собеседниками, включая момент отправки друг другу открытых ключей, открыто для других пользователей интернета.

Третий этап – это перечисление денег продавцу. Для обеспечения анонимности используется платежная система и одноименная валюта Bitcoin (BTC). BTC возникла так же недавно, как и технология TOR[[2]](#footnote-2). Это удобный способ производить денежные транзакции через интернет, также полностью соблюдая свою анонимность. Данная платежная система не имеет централизованного регулятора, имеющего власть заморозить транзакцию или конфисковать средства. Регистрация денежного кошелька в ней – проста и не требует никаких личных данных.

И, наконец, последний этап сделки – передача товара покупателю. Распространено два способа анонимной передачи товара. Первый заключается в использовании почтовой системы. Покупатель указывает почтовый адрес (не свой), на который приходит товар продавца. Товар упакован в таком виде, чтобы обеспечить низкую вероятность его обнаружения в замаскированной посылке. В случае, если в ходе досмотра почтовая организация его обнаружила, то правоохранительным органам будет необходимо доказать, что получатель знал о содержимом посылки, для того чтобы заставить его нести правовую ответственность. Как правило, покупатель ожидает, что при грамотном и осторожном поведении во время «приёма» посылки, он сможет её избежать. Второй способ заключается в том, что продавец выбирает немноголюдное место и прячет в нём свой товар. После поступления оплаты он сообщает покупателю место, где находится товар. После этого покупатель отправляется в это место, находит там товар и подбирает его. Как видно, при первом и втором способе продавец находится в большей безопасности чем покупатель, поскольку продавец имеет возможность деанонимизировать или даже поймать покупателя на последнем этапе совершения сделки. Мотив на такой поступок у него может быть, например, если он будет пойман правоохранительными органами и пойдет на сделку со следствием. Поэтому почти полную гарантию анонимности на анонимных торговых площадках имеет на самом деле только продавец, но не покупатель. С другой стороны, на традиционных рынках покупатель ещё больше зависим от продавца, поскольку здесь покупатель практически сразу, с первого этапа совершения сделки предоставляет продавцу исчерпывающую информацию о себе, в то время как на анонимных площадках последнему нужно ещё приложить какие-то усилия, чтобы его вычислить. В конце концов, употребление ПАВ является административным правонарушением, в то время как сбыт – уголовным.

Следующей особенностью анонимных рынков является невозможность принудить контрагента исполнить свои обязательства. Если оплата товара производится раньше его вручения покупателю, то покупатель не защищен от недобросовестного продавца, захотевшего не отдавать ему товар и оставить деньги себе.

Сложилось два основных способа решить эту проблему. Во-первых, каждая сторона имеет репутацию на анонимном рынке, даже несмотря на то, что постоянно остаётся инкогнито. Каждый участник рынка стремится выбирать контрагента, который имеет продолжительную историю сделок, к которым не было претензий у прошлых партнеров по сделке. Как правило, сначала производится оплата товара, поэтому в большинстве случаев покупатель находится в зависимом положении от продавца, а не наоборот. Однако благодаря механизму репутационных потерь он не захочет обмануть покупателя, поскольку в таком случае покупатель даст сигнал другим участникам рынка, что продавец оказался недобросовестным и продавец лишится части будущей выручки. Разумеется, такой способ не даёт гарантий от обмана. Тем не менее, он весьма надежен: чем больше успешных сделок заключил продавец, тем он популярнее. Если популярность продавца очень высока, то даже единственный неприятный случай может сильно повлиять на его операционную деятельность.

Второй способ предотвращения конфликтных ситуаций – это арбитражная система. Некоторые участники площадок играют особую роль и называют себя гарантами. Покупатель и продавец обращаются к гаранту, чтобы он был арбитром в их сделке. Они ознакомляют его с условиями сделки[[3]](#footnote-3), а покупатель перечисляет ему денежные средства, которые предназначены для оплаты товара или услуги продавца. После передачи товара или исполнения услуги продавец ожидает от гаранта перечисления вручённых ему денежных средств. Если у покупателя не возникло претензий к качеству товара или услуги, то он сообщает гаранту, чтобы тот перечислил оплату. Если претензии есть, то конфликт обязан решить гарант. Если он считает, что претензии обоснованы, то он возвращает деньги обратно покупателю; в противном случае он отправляет их продавцу. Разумеется, часть средств покупателя остаётся у гаранта в качестве комиссии за его работу. Данный способ ещё более надежен для покупателя, чем механизм репутационных потерь, хотя и здесь он не имеет гарантий от недобросовестного поведения продавца, потому что гарант будет исходить из того, что любая сторона может вводить его в заблуждение. Более того, гарант может некоторым образом обеспечивать безопасность покупателя в тот момент, когда он принимает товар продавца: если после обозначенного времени покупатель не выходит на связь, то гарант сообщает об этом другим участникам рынка, принося очень большой вред репутации продавца.

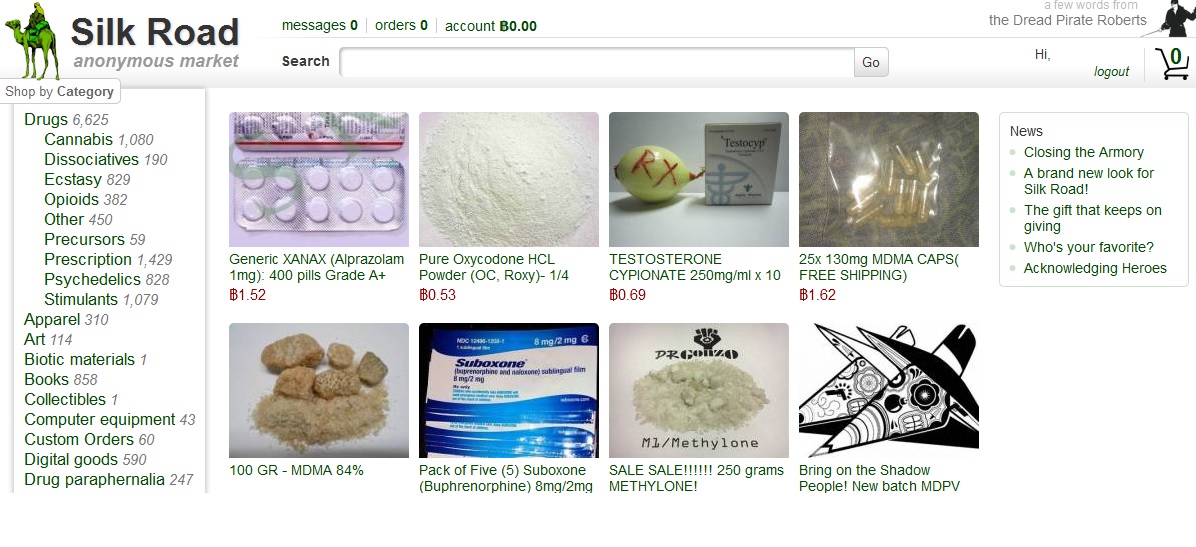


Рисунок 1 Домашняя страница Silk Road 1.0

Изложенные выше принципы лежат в основе, пожалуй, всех анонимных торговых площадок, хотя каждая имеет свои особенности. Самой старой торговой площадкой признается Silk Road, которая возникла в 2011. Для того, чтобы иметь возможность заключать сделки, пользователю необходимо зарегистрироваться, при этом учётную запись (аккаунт) продавца можно было приобрести только за плату, взимаемую администрацией сайта. Имея аккаунт продавца пользователь получает право создавать объявления о продаже товара. Каждое объявление имеет чёткую форму, необходимо указать название товара, его ключевые характеристики (цена, количество, способ упаковки, особенности доставки и т.д.). После создания объявления продавец ждёт, пока потенциальный покупатель обнаружит его на сайте среди других объявлений, ознакомиться с ним и нажмёт на кнопку «Place order». Далее покупатель указывает адрес, на который доставляется посылка и перечисляет деньги продавцу через Silk Road, который является посредником в этом денежном переводе. Данная площадка работала во всех странах мира, кроме России[[4]](#footnote-4), однако не в каждой стране были продавцы товаров. Транзакции могли производиться между резидентами из разных стран, хотя риск обнаружения товара на таможне или в почтовой организации всегда сохранялся. Данный сайт привлек к себе большое внимание правоохранительных органов США, которым удалось взломать его и поймать администрацию в октябре 2013 года. Однако спустя месяц возникла реинкарнация сайта – Silk Road 2.0 – которая также была взломана ФБР почти ровно через год (2014). С тех пор существует вторая реинкарнация сайта – Silk Road 3.0 – которая на данный момент в целях собственной безопасности плавно трансформируется в Silk Road 3.1.

Хотя данная торговая площадка является самой известной в мире, почти одновременно с ней возникло и много других, таких как Agora, Evolution. Позже к ним добавились и другие, такие как Tochka или Valgalla. Некоторые из торговых площадок взламываются хакерами, другие удаётся закрыть ФБР США, поэтому постоянно какие-то торговые площадки закрываются, а на их место приходят другие.

Во всём этом многообразии площадок, российские площадки сильно выделяются. Во-первых, так сложилось, что на территории России не имеет смысла использовать международные площадки, при этом российские площадки охватывают только Россию и частично страны бывшего СССР. Во-вторых, после взлома и закрытия первой версии Silk Road самая известная российская площадка RAMP может считаться самой старой до сих пор действующей анонимной торговой площадкой.[[5]](#footnote-5) И, наконец, в-третьих, в отличие от западных площадок, о российских известно гораздо меньше. Если брать Silk Road и его реинкарнации, то о них уже написано достаточно много научных работ, российские же площадки пока что до сих пор обделены вниманием.[[6]](#footnote-6)

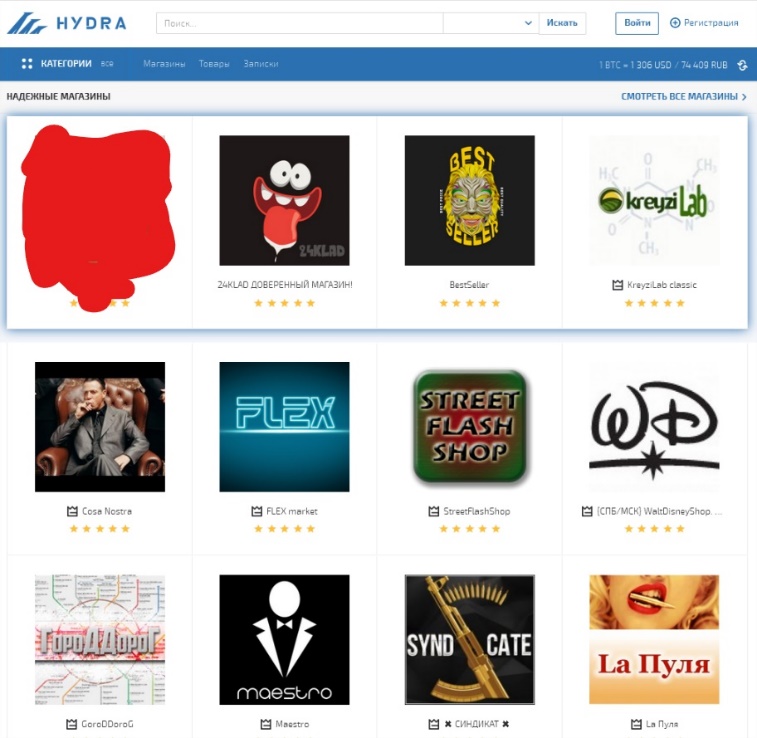
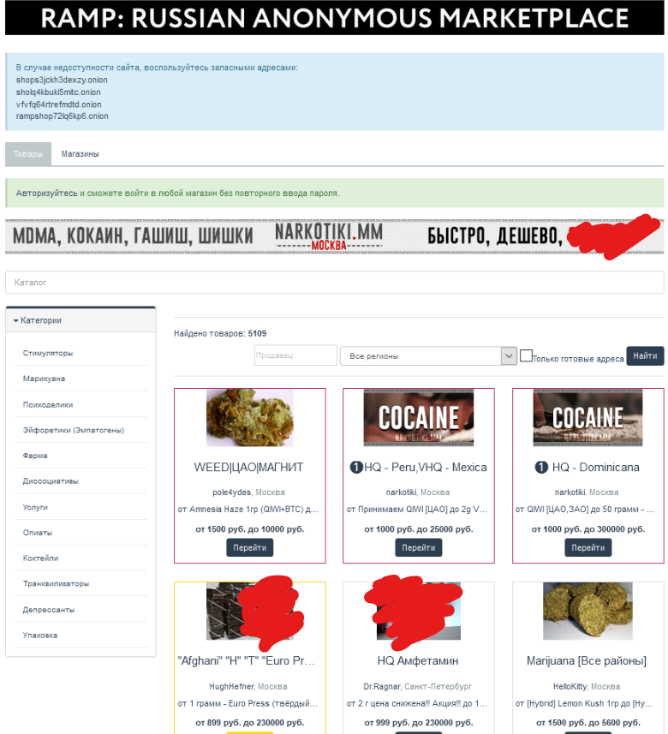


Рисунок 2 Моментальные магазины HYDRA Рисунок 3 Моментальные магазины RAMP

Среди российских торговых площадок самыми известными являются RAMP, HYDRA, Runion, Rutor (не путать с торрент-трекером), Anthill, SwimPool, Pekarmarket, Matanga, SpiceForum.[[7]](#footnote-7) Российские торговые площадки условно можно разделить на две категории: торговля ПАВ и торговля всем остальным. В первой категории RAMP является, как уже было сказано ранее, самой старой (функционирует с сентября 2012)[[8]](#footnote-8) и популярной торговой площадкой. Торговля на RAMP ведется в двух форматах: форумный, подразумевающий абсолютно свободное взаимодействие между участниками, и формат моментальных магазинов. В первом случае продавец арендует у администрации сайта право иметь собственную ветку на форуме, где он размещает информацию о своих товарах и контактах. Во втором случае все магазины продавцов объединены в структурированный список, чем-то похожий на список Silk Road. Однако ключевое отличие в том, что доступ к товару покупатель получает сразу (моментально) после оплаты: как только покупатель согласится на сделку, сайт торговой площадки автоматически предоставит ему место, где зарыт «клад».

Следующей площадкой является HYDRA, возникшая в 2015 году как постепенная консолидация площадок LegalRC, WayAway, RuSilkRoad и некоторых других. HYDRA изначально предлагала торговлю в формате моментальных магазинов, однако многие продавцы также имеют форумные ветки на LegalRC и WayAway. Данная площадка, несмотря на короткую жизнь, является основным конкурентом RAMP и вторым по популярности местом торговли, поскольку ведёт успешную стратегию развития и продвижения. Предположительно, на обе площадки приходится наибольшая доля анонимного рынка ПАВ.

Самым популярным представителем второй категории является сайт Runion. Данный ресурс имеет несколько профилей деятельности, и самый главный из них – это неформальный статус центра общения для представителей субкультуры русскоязычного даркнета. Благодаря основному профилю деятельности, потенциальным покупателям и продавцам легче встретить друг друга. С этого ресурса новость о спросе или предложении какой-либо специфичной услуги потенциально распространяется быстрее всего в остальную часть *даркрунета*. Товаром и услугой может быть всё что угодно, однако скорее всего предмет сделки будет нарушать уголовный кодекс РФ. Либо это уникальный или редкий товар, либо услуга, нарушающая законодательство или требующая редкой и сомнительной квалификации исполнителя. При этом потенциальные участники сделок стараются дозировать доступ к информации о предмете сделки, соблюдая баланс между поиском контрагента и привлечением излишнего внимания остальных незаинтересованных пользователей. Второе преимущество Runion-а как места общения – наличие известных авторитетных пользователей, заработавших доверие и имеющих полезные связи. Такие пользователи часто становятся посредниками, помогая найти контрагента для сделки и впоследствии играя роль гаранта (хотя доверить обе этих обязанности одному и тому же лицу – это некоторый риск столкнуться с недобросовестным поведением).

Остальные площадки также способны представлять собой интерес, однако их предполагаемая доля на рынке анонимных сделок мала, и поэтому, до тех пор, пока не изучены основные ресурсы даркнета, они должны иметь второстепенную важность.

Вместе с тем, при первом знакомстве с русскоязычной частью даркнета становится ясным, что это большой и неизученный сегмент интернета, который имеет свою теневую экономику, функционирующую по своим собственным принципам. Данная работа отражает стремление дать представление об этой экономике, исследование которой возможно статистическим способом на основе больших даных.

## Обоснование целей, поставленных в работе.

Для исследования торговых площадок статистическим способом через анализ загруженных снимков выбирается парсинг, поскольку данный автоматический способ считывания информации с html-страниц является довольно мощным, по сравнению с ручным анализом, когда требуется зрительно искать нужную информацию, записывать её и структурировать. Но для применения парсинга необходимо, чтобы загружаемые в ходе производства снимков html-страницы были одинаковыми по структуре и содержали однородную информацию. В связи с этим наиболее удобным для исследования являются площадки RAMP и HYDRA, которые практически полностью используются только для продажи ПАВ, поскольку однородные транзакции означают более однородную информацию о них.

Однако одного этого требования недостаточно. Например, весь сайт RAMP, который включает в себя форумную часть и моментальные магазины, тоже не удобен для парсинга. Форумная часть предлагает большую свободу во взаимодействии между контрагентами, в связи с чем формат объявлений о продаже товара индивидуален для каждого продавца, поэтому информация в таких объявлениях от разных продавцов имеют разную структуру относительно друг друга.

Поэтому наиболее удобными ресурсами для исследования являются моментальные магазины площадок RAMP и HYDRA, которые подразумевают наличие уже готовых для моментальной покупки закладок с ПАВ. Для того, чтобы покупатель мог купить товар в любой момент времени – независимо от времени, когда продавец готов обслужить его – необходим автоматический формат совершения транзакций, в ходе которого и денежные средства, и информация о месте закладки будет проходит через посредника — администрацию торговой площадки. Более того, продавцы тоже должны иметь возможность оформить объявление о продаже новой закладки в независимости от способности администрации её оформить. И это обуславливает то, что на обоих торговых площадках моментальные магазины имеют четкую структуру с одинаковым форматом объявления о продаже для всех товаров.

Однако, если выбрать для исследования только моментальные магазины, становится важным понимание того, насколько большую часть теневой экономики российского даркнета удаётся охватить в ходе исследования. На самом деле, существуют достаточно серьёзные соображения, по которым данные моментальные магазины действительно охватывают наибольшую часть торговли. Как уже было сказано ранее, в России слабо распространены – если вообще распространены – англоязычные торговые площадки, поэтому отсутствие к ним внимания никак не отразится на охвате российского анонимного рынка. Если говорить о выборе между площадками для ПАВ и площадками для торговли другими запрещенными или спорными товарами и услугами, то очевидно, что площадки с ПАВ занимают наибольшую часть анонимного рынка, поскольку иные площадки не имеют четкой специализации на товарах и известных постоянных продавцов (за некоторыми исключениями), которые могли бы возникнуть при стабильно высокой активности на них. Если говорить об игнорировании форумной части RAMP и других форумных торговых площадках, то это пожалуй единственные конкурентные ресурсы, способные занять существенную долю анонимного рынка. Однако форумные площадки подходят больше для оптовой торговли, когда обоим сторонам важно иметь больше свободы для взаимодействия при обсуждении деталей сделки. Розничные же продажи удобнее производить через моментальные магазины, где и продавец, и покупатель несут меньше транзакционных издержек, поскольку условия сделки и сама закладка товара готовы заранее, что не вынуждает обе стороны вступать друг с другом в контакт для заключения сделки. Ну и разумеется, покупателям удобнее производить выбор в условиях моментальных магазинов, где в большом ассортименте предложений ему легче ориентироваться, поскольку он структурирован. По той же причине остальные торговые площадки также не будут занимать слишком большой доли анонимного рынка, поскольку в силу сетевого эффекта данный рынок постоянно тяготеет к консолидации в одну или две площадки (транзакционные издержки у участников рынка снижаются, если они находятся на одной и той же площадке). Поэтому, в силу выше сказанного, выбор моментальных магазинов RAMP и HYDRA не только удобен, но ещё и должен охватить большинство сделок на анонимном рынке даркрунета.

Отдельной темой является определение основных задач анализа этих торговых площадок. В итоге в качестве основных было выбрано определение структуры типов ПАВ, которые торгуются на обоих площадках, оценка месячного объема транзакций на RAMP и построение динамики активности на этой площадке, поскольку решение таких задач даёт хорошее представление об этих площадках и является хорошей базой для продолжения анализа. Так же немаловажным является изучение взаимосвязи характеристик обоих площадок с остальным миром. Для этого было решено сосредоточиться на RAMP и протестировать наличие прогнозной силы курса биткоина к рублю на динамику активности на площадке RAMP. Во всём выше сказанном, кроме тестирования прогнозной силы внешнего показателя, данная работа близка по содержанию к статье (Christin, 2012).

Наличие причинно-следственной взаимосвязи между курсом биткоина и активностью на RAMP также является интересным вопросом, поскольку пользователи используют эту валюту при покупке товаров. Если связь будет установлена, то это будет важным фактом как для тех, кто исследует данную валюту, так и для исследователей анонимных торговых площадок. Однако её выяснение, что сложнее чем исследование прогнозной силы, требует достаточно усилий, чтобы посвятить ей отдельную статью.

## Сбор данных

Сбор данных в этой работе является ключевым звеном, определяющим дальнейшие шаги в анализе площадок. К добываемым данным предъявляются два основных требования. Во-первых, необходимо извлечение как можно большего количества полезной информации. Во-вторых, требуется обеспечение как можно более высокого качества загружаемых данных, когда число сбоев и недозагруженных из-за них html-страниц было бы минимальным. В дополнение к основным требованиям есть также ряд второстепенных. Во-первых, это экономия места на жёстком диске для загружаемых данных; во-вторых, уменьшение времени активной загрузки каждого снимка; в-третьих, сохранение анонимности исследователя и анонимности участников площадок (по этическим соображениям).

Поэтому, следовало выбрать такой метод загрузки данных, который наиболее полно удовлетворяет всем вышеизложенным требованиям. Хорошим решением оказался предложенный анонимным исследователем Gwern Branwen способ, который находится в приложении к архиву даркнета, опубликованном им в 2015 году (Branwen et al., 2015). Он основан на функции зеркальной загрузки html-страниц, которую имеет программа Wget. Данная программа является своеобразным интернет браузером, не имеющим пользовательского интерфейса. Получая первоначальную url-ссылку на одну из страниц сайта, он загружает её и все те html-страницы, ссылки на которые указаны в содержимом первой страницы. Также привлекалась программа Privoxy, которая является внутренним прокси между компьютером и внешней сетью. Компьютер был настроен таким образом, чтобы все запросы во внешнюю сеть проходили через это прокси. В свою очередь, Privoxy, с помощью регулярных выражений, был настроен на запрет загрузки «лишних», т.е. бесполезных, не несущих полезной информации, html-страниц. Поэтому, благодаря использованию Privoxy достигался меньший объем места на жёстком диске компьютера, а также меньшее время производства каждого снимка по сравнению с обычной рекурсивной загрузкой через Wget. Также использовался Cron для автоматического запуска работы Wget согласно расписанию. Так как Wget и Cron предполагают работу через консольный терминал, то в качестве операционной системы была взята операционная система из семейства Linux.

При выборе стратегии загрузки снимков учитывались следующие факторы. Во-первых, загрузка занимает достаточно много времени и места на жестком диске: на производство снимка RAMP уходит около 12 часов и 3,5 Гб памяти, для HYDRA, в свою очередь, требуется 5 часов. Во-вторых, при написании выпускной курсовой работы даётся не так много времени на сбор данных, что можно компенсировать высокой частотой производства снимков. Поэтому, в качестве золотой середины было принято решение производить снимки обоих площадок два раза в неделю: в вечер четверга и вечер воскресения. С одной стороны, такой выбор имеет недостаток, поскольку между четвергом и воскресением проходит 3 дня, а между воскресением и четвергом 4 дня, поэтому промежутки между днями производства снимков несимметричны. С другой стороны, такой подход может показать, как меняется активность на площадках в будние и выходные дни.

Сайты площадок начали регулярно кэшироваться с декабря 2016 года, и на протяжении всего периода производства снимков периодически возникали внештатные ситуации. При ограниченности в ресурсах, которые были доступны для исследовательской работы, иногда случалось, что кэширование приходилось пропускать. В большинстве случаев проблему удавалось решить путем начала загрузки снимка как можно раньше, следующим днём ночью или утром, и это не должно было сильно влиять на содержимое данного снимка. Иногда кэширование приходилось отложить полностью, поэтому некоторые наблюдения оказались пропущенными.

Далее, по объективным причинам, в середине февраля загрузку RAMP пришлось приостановить до второй половины апреля, а плановую загрузку HYDRA прекратить вовсе в середине. После прекращения плановой загрузки пришлось также сделать три внеплановых снимка для того, чтобы иметь возможность построить регрессию наблюдаемого у магазинов HYDRA количества совершённых транзакций на количество оставленных отзывов. Таким образом, имеется ряд наблюдений состояния RAMP с начала декабря по начало мая с перерывом в 2 месяца, а также ряд наблюдений HYDRA с начала декабря по начало марта. Для RAMP в целом имеется 22 пропущенных снимка, из которых 18 были в рамках запланированного перерыва. Для HYDRA всего было пропущено 3 снимка. Более конкретные характеристики произведенных снимков можно наблюдать в таблицах 1 и 2.

Таблица 1 Характеристики загруженных состояний (снимков) HYDRA



В итоге можно утверждать, что качество почти всех произведённых снимков является достаточно хорошим. Для измерения качества использовалось предложение из (Munksgaard, Demant and Branwen, 2016): для оценки полноты загрузки сайтов площадок используется соотношение числа загруженных товаров к числу обнаруженных. В ходе рекурсивной загрузки html-страниц программа Wget обнаруживает ссылки на другие html-страницы, некоторые из которых принадлежат html-страницам товаров площадки. Поэтому, если в ходе рекурсивной загрузки наблюдаются систематические сбои, то Wget пропустит загрузку большого числа этих ссылок, из-за чего многие обнаруженные товары окажутся незагруженными. Конечно, как указывают авторы статьи, низкая доля незагруженных товаров не гарантирует, что снимок содержит в себе 100% всех доступных html-страниц. К примеру, если страница, с которой начинается рекурсивная загрузка, содержит в себе ссылку только на одну следующую страницу, в ходе загрузки которой произойдёт сбой, то такая метрика покажет 50% долю числа загруженных страниц в числе всех обнаруженных страниц. Но это не значит, что была загружена половина все существующих на сайте страниц. Однако, судя по тому, что количество загруженных товаров для обоих площадок от снимка к снимку находится на стабильном уровне (см. таблицы1 и 2), можно сделать вывод о том, что снимки достаточно хорошо приближают реальное состояние сайтов. Более того, сайт площадки RAMP указывает общее количество товаров, которое доступно для покупки на нём, поэтому для таблицы 2 указано отношение числа загруженных товаров и числа обнаруженных товаров к этому количеству. И если доверять этой информации, то почти все снимки RAMP имеют весьма высокую полноту на уровне не менее 70% по числу загруженных товаров, что говорит о хорошем качестве загруженной базы данных.[[9]](#footnote-9)

Таблица 2 Характеристики загруженных состояний (снимков) RAMP



В отношении HYDRA нужно сделать замечание по резкому падению загруженных файлов после 16 февраля: такое изменение вызвано только оптимизацией процесса производства снимков, приведшей к отказу от загрузки ненужных файлов (речь идет о файлах .jpg и .img формата).

## Парсинг

Следующей важной частью является парсинг произведенных снимков. Парсинг – это работа направленная на извлечение и структурирование информации, содержащейся в html-страницах. Каждая html-страница представляет из себя сочетание обычного текста (контента страницы) и компьютерного кода html, который преобразуется браузером в уникальное оформление страницы. Парсинг основан на возможности обработки html-файлов языком регулярных выражений. Последний позволяет искать в файле фрагменты текста, удовлетворяющие условиям, которые задаются тем, кто осуществляет парсинг.

Таким образом, существует возможность, зная в какой части html-файла находится название товаров, цены продажи и другие характеристики, составить таблицу, в которую с помощью языка регулярных выражений будут отнесены данные по каждому товару. Получившаяся таблица окажется достаточно хорошо подготовленной для анализа, который необходим в целях решения поставленных в данной работе задач.

Осуществление парсинга представляет из себя написание компьютерного кода. Его содержимое отражает последовательное и структурированное открытие каждого html файла, поиск посредством языка регулярных выражений всей желаемой информации, и сохранение её в таблицы. Так как одинаковые действия нужно проделать для каждого html-файла (вернее, для каждого типа html-файлов), то такой код должен включать в себя большое количество циклов и массивов. В качестве языка программирования был выбран Python, который хорошо подходит для работы с текстовым форматом, имеет развитый подключаемый пакет с командами для применения языка регулярных выражений, а также удобен при работе с циклами и массивами.

#### Парсинг моментальных магазинов RAMP

Так как на RAMP каждый магазин оформлен как отдельный сайт с субдоменным адресом, то загруженный снимок RAMP-а представляет собой контейнер папок, одна из которых отображает html-страницы со списком всем выставленных на площадке на продажу товаров и списком всех функционирующих магазинов, а все остальные – посвящены конкретному магазину и каждому товару, который данный магазин предлагает. Страничка каждого товара не имеет постоянного url-адреса, что с одной стороны сильно усложняет парсинг и время производства каждого снимка, но с другой стороны существенно повышает надёжность и качество каждого произведенного снимка, поскольку каждый товар имеет несколько попыток для загрузки в случае сбоя обработки запроса.

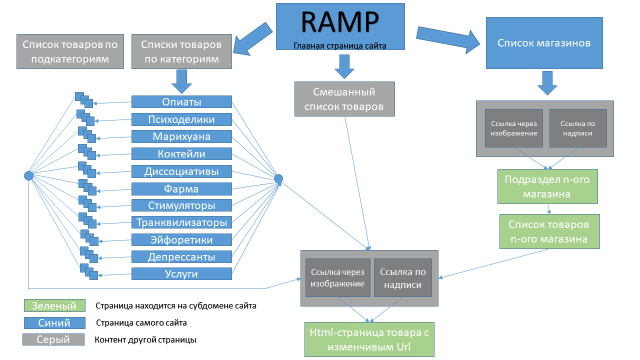


Рисунок 4 Схема сайта моментальных магазинов RAMP

Как уже было сказано раньше, основным предметом купли и продажи на RAMP являются ПАВ, которые сайт предлагает либо единым списком, либо разделяя их на категории. Товар можно выбрать в зависимости от того, какой магазин им торгует, в зависимости от типа ПАВ или от города, где находится закладка.

Важным является вопрос, как много информации можно извлечь из RAMP с помощью парсинга. На страничке каждого товара можно наблюдать его название; название магазина, который им торгует; отзывы покупателей; а также выбор из числа нескольких закладок, которые отличаются между собой весом или размером упаковки; районом города, где произведен заклад; и, наконец, ценой. Информацию о типе или подтипе ПАВ нельзя извлечь из страницы товара, поэтому она доступна только для тех товаров, которые удалось загрузить по ссылке, которая была указана в списке товаров по типу или подтипу.

В итоге, парсинг RAMP оказался довольно трудоемким процессом. В результате для каждого произведенного снимка было создано две таблицы: сокращенная и расширенная. Расширенная включает в себя всю выше изложенную информацию, в то время как сокращенная не содержит сведений о конкретных закладках для каждого товара (нет районов города, веса закладки и её цены).

Конкретные закладки и их характеристики (см. на рисунке 6) в данной работе называются опциями товара. Каждый товар на обоих площадках обладает такими опциями, и покупатели при покупке товара должны выбрать конкретную опцию из списка доступных.

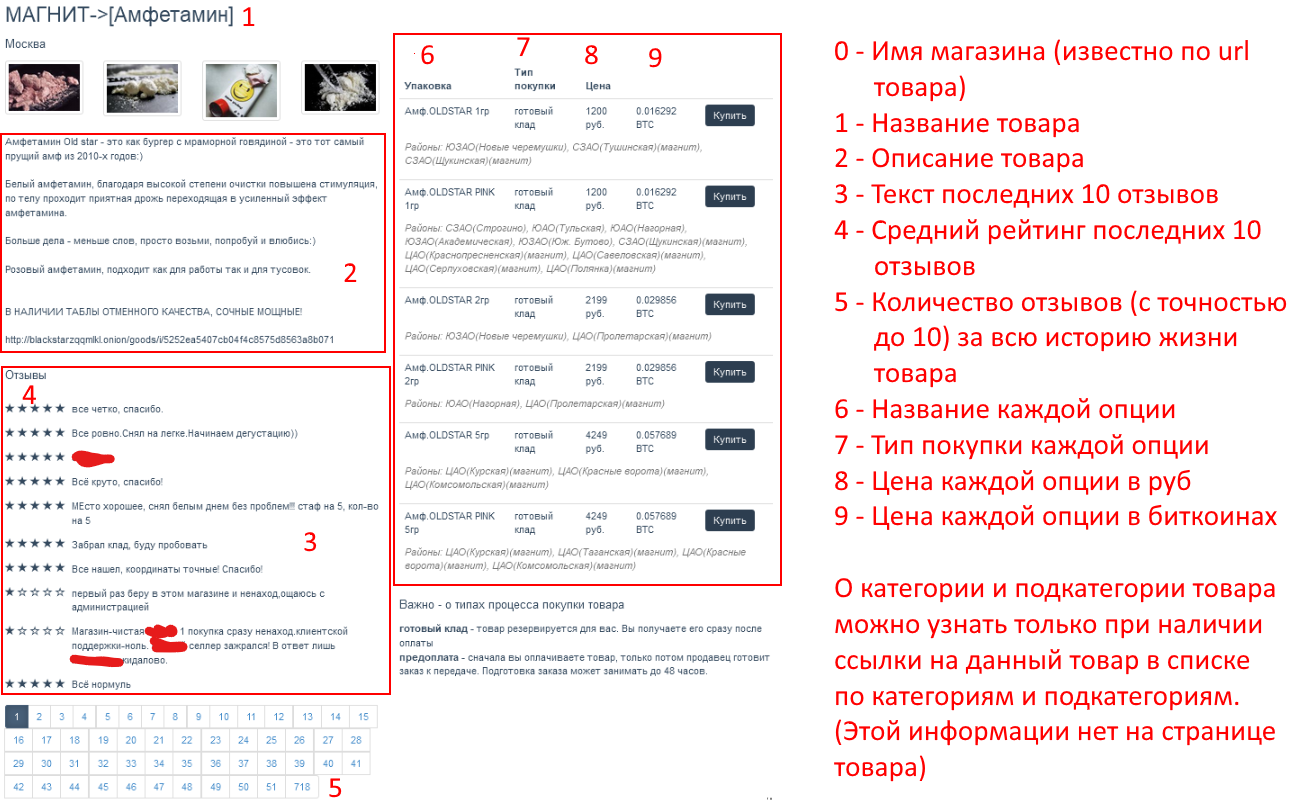


Рисунок 6 Пример страницы товара на RAMP с указанием информации, доступной для парсинга

Одной из особенностей RAMP является то, что хотя для всех магазинов и указан диапазон, в котором лежит число совершенных им сделок, реальное число совершенных сделок узнать невозможно. Поэтому единственный способ хотя бы как-нибудь измерять количество проданного товара — по числу оставленных отзывов. Оно имеет особое значение в контексте данной работы. Число комментариев известно с точностью до 10, и оно показывает все отзывы, которые были оставлены за всё время существования товара на сайте. Благодаря этой информации возможно дать приблизительную оценку объема сделок, которые совершаются на площадке за некоторый период, о чём пойдет речь ниже.

#### Парсинг моментальных магазинов HYDRA

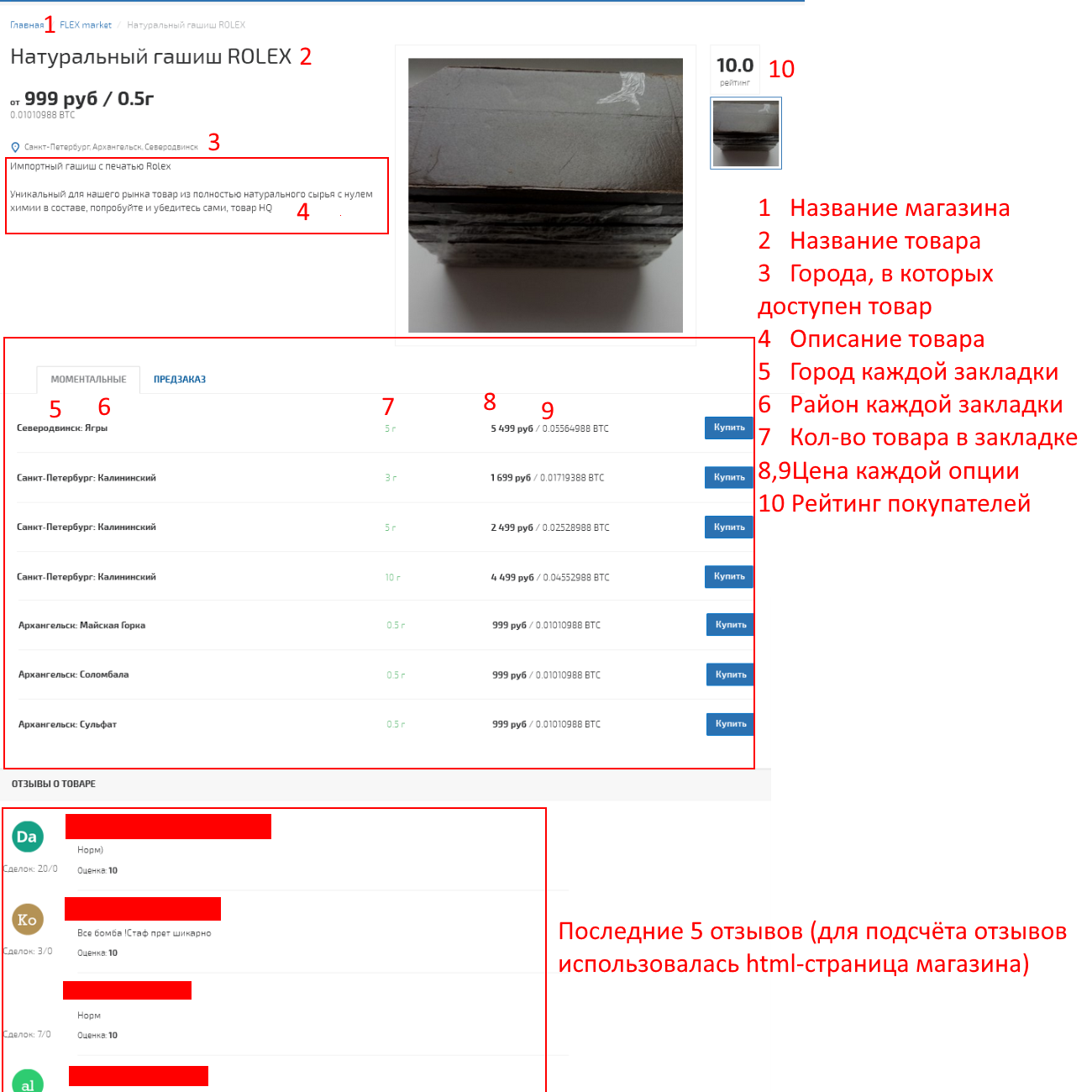


Рисунок 7 Пример страницы товара на HYDRA с указанием информации, доступной для парсинга

Парсинг HYDRA аналогичен RAMP по трудоемкости и принципу написания программного кода. В результате также было создано две таблицы. Информация, доступная для парсинга, указана на рисунке 7. Первая таблица показывает всю известную информацию о товаре, кроме той, что варьируется в зависимости от конкретной опции. Вторая таблица включает информацию из первой таблицы, а также информацию по опциям. Важной особенностью HYDRA является то, что один и тот же товар может иметь опции с закладками в совершенно разных городах, поэтому данная площадка принципиально не может показать почти никакой информации о характеристиках анонимного рынка в регионах РФ. С другой стороны, в отличие от RAMP, в HYDRA в опциях для разных товаров указан единый стандартизированный тип фасовки и количество товара в едином формате. Оба этих факта делают HYDRA очень удобной для анализа ценовых характеристик товаров. Также на данной площадке для некоторых магазинов, в отличие от RAMP, доступно для наблюдения точное число совершённых сделок. Дело в том, что обе площадки скрывают количество сделок для большинства магазинов и указывают только примерный диапазон, в котором оно лежит, но HYDRA делает исключение для новых развивающихся магазинов, у которых число сделок ещё не достигло 500. В заключение, если на RAMP для каждого товара указываются все оставленные когда-либо отзывы, то HYDRA отображает только последние 100 оставленных отзывов для каждого магазина (не для товара). Такая практика применяется с целью дать возможность магазинам «стирать темные пятна» в своей репутации.

## Статистический анализ данных по HYDRA

После произведенного парсинга стал возможным статистический анализ полученных таблиц для обоих площадок. В этом разделе сначала будет дан обзор внутренней структуре HYDRA, а затем — структуре RAMP.

За всё время сбора данных число магазинов на HYDRA стабильно росло (см. таблицу 1), начиная свой рост с 267 магазинов и заканчивая 429 на 11 марта. Такое внушительное изменение в числе активных магазинов за почти 2,5 месяца говорит о том, что данная площадка находится на стадии активного роста. Отдельно стоит упомянуть резкий рост числа магазинов (а также числа доступных для покупки товаров, см. таблицу 1), который произошёл на снимке от 12 февраля: открылось около 23 новых магазинов, а число товаров возросло на 325. Такие изменения можно связать с тем, что как раз в тот период происходило мирное поглощение HYDRA площадки RuSilkRoad: администрации обоих площадок объявили об объединении, а магазинам RuSilkRoad предложили переехать на льготных условиях в HYDRA[[10]](#footnote-10), прежде чем данная площадка прекратит существование. Через какое-то время сайт RuSilkRoad стал зеркалом, перенаправляющим запросы на HYDRA. В отношении числа товаров также верно, что оно стабильно растёт: с 1230 в начале периода наблюдений до 2332 в конце.

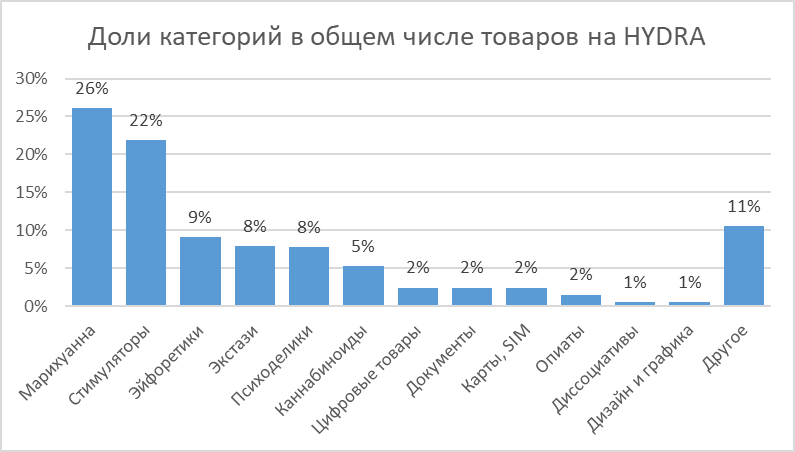


Рисунок 8 Доли категорий в общем числе товаров на площадке HYDRA

Также весьма интересна структура типов товаров, которая торгуется на этой площадке, которую можно наблюдать на рисунке 8. Как и предполагалось, большую часть покупаемых товаров на данной площадке составляют наркотики (более 80%). Наибольшую долю имеют марихуанна и стимуляторы, занимающие по четверти рынка. Сильные наркотики гораздо реже распространены. Например, опиаты составляют всего 2%. Также в приложении 4 приведена более подробная разбивка товаров на подкатегории. Из неё, например, видно, что кокаин также имеет долю 4%. Неожиданностью стало появление товаров, не относящихся к ПАВ. Так, например, с недавнего времени стало возможным приобретать подделанные документы, хакерские услуги по взлому личных данных и т.п. Неожиданно появились услуги по дизайну, которые нацелены на вновь открываемые магазины. Данные категории товаров возникли уже после того, как началось наблюдение за площадкой, поэтому сейчас ещё рано сделать вывод о том, что площадка стала стабильно использоваться для покупки товаров, не относящихся к ПАВ (хотя и видно, что она находится на пути к этому). Отдельное место занимают товары, относящиеся к категории «Другое». В большинстве случаев товары из этой категории – это способ финансово поблагодарить сотрудников магазинов за их работу («дать на чай») или дать объявление об открытых вакансиях по заготовке кладов.

Таблица 3 Ценовые характеристики товаров HYDRA



Благодаря характеристикам площадки весьма удобно собрать информацию о ценах на все категории товаров (см. таблицу 3). На ней отражены средние, медианные и максимальные цены, посчитанные по опциям товаров каждой категории. Так, например, согласно таблице, средняя цена за 1 грамм конопли составляет 1200 руб, а за наркотики из категории стимуляторов – 2300 руб. Видна связь между ценой товаров и вредом, который они наносят организму: имеют наиболее низкие средние цены имеют принадлежат самым вредным типам ПАВ (см. барбитураты, синтетические каннабиноиды). Хотя, опиаты, например, являются самым дорогими из ПАВ. Высокая цена на товары категории «Другое» можно объяснить тем, что в них присутствуют объявления о работе. «Найм на работу» на анонимных рынках весьма специфический: на самом деле сотрудники магазинов теоретически имеют возможность полностью сберегать свою анонимность даже от начальства. Однако для того, чтобы стать «кладменом», необходимо внести залог за первую партию товара, которую он будет прятать в кладах.

Следующей интересной темой является продолжительность активного существования товаров на HYDRA. Для того, чтобы её определить, для каждого товара оценивалось первая дата, в которую он впервые возник в базе наблюдений за текущим состоянием площадки, а также последняя дата, когда его всё ещё можно было в ней обнаружить. Сделано предположение, что до обнаружения в базе наблюдений товар не существовал на площадке, а после исчезновения из неё – был убран с продажи и на самой площадке. Даты возникновения были разделены на несколько временных промежутков, по семестрам каждого месяца. После этого был составлен рисунок 9, на котором наглядно видно, что площадка успела обновить ассортимент товаров на 50% к середине февраля (т. е. за 2,5 месяца). Всего же, за весь период наблюдений, с учетом резкого роста числа товаров после поглощения RuSilkRoad, ассортимент обновился примерно на 65%.

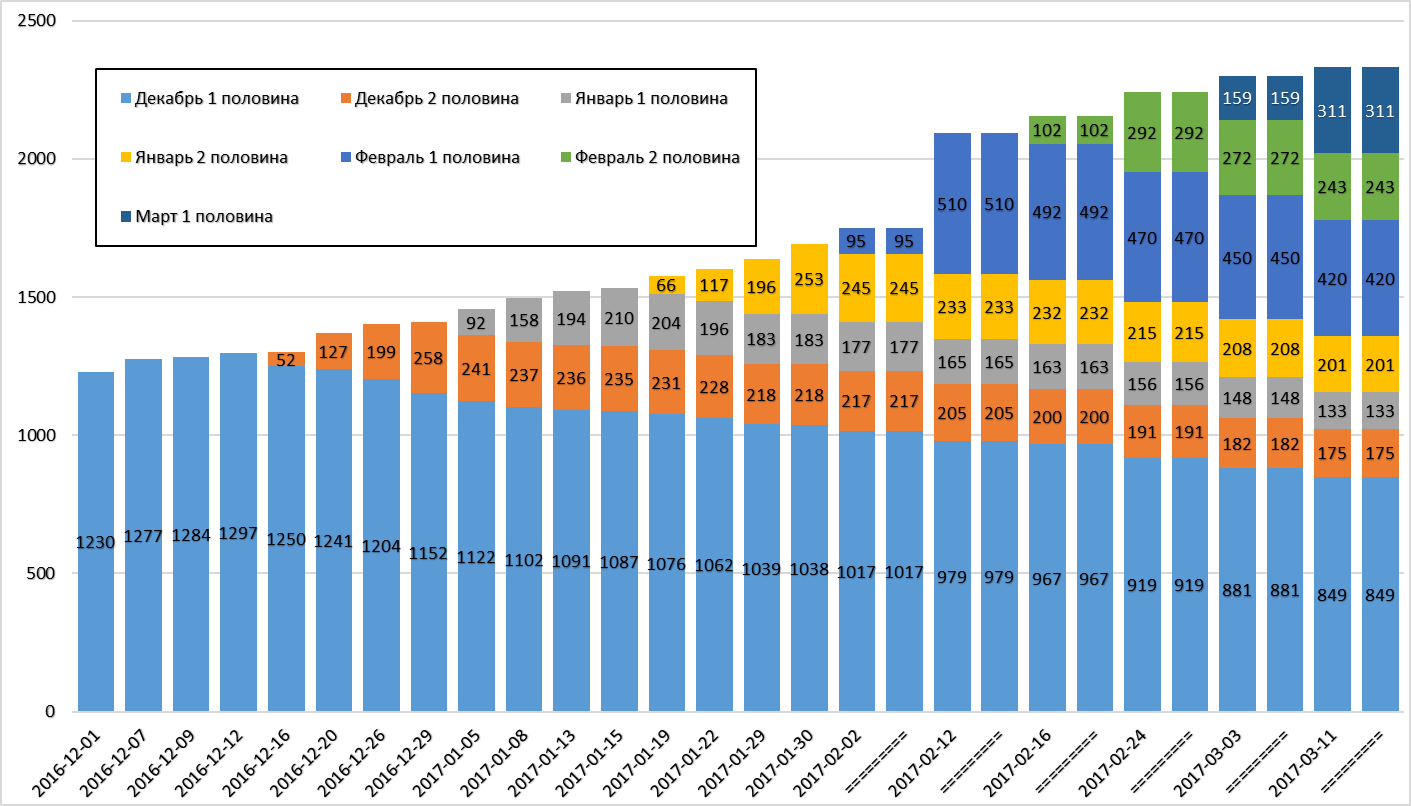


Рисунок 9 Появление и выбытие товаров на площадке HYDRA

В заключение, в приложении 6 можно найти сведения о размере магазинов площадки. Для этого использовалась информация на сайте HYDRA о количестве сделок, которые совершил магазин за всё время своего существования. Сведения приведены для разных снимков[[11]](#footnote-11) с примерно недельным промежутком в наблюдениях. Из таблицы видно, что магазины с предположительно низким оборотом (менее 500 сделок за всю историю) составляют около 70% процентов от общего числа.

#### Регрессия числа сделок на число оставленных комментариев.

Одной из целей данной работы является оценка объема транзакций, совершаемых на анонимных площадках. И хотя данная оценка будет производиться на примере площадки RAMP, это было бы невозможным, если бы не был найден способ конвертации оставленных отзывов пользователей о качестве покупаемых ими товаров в число реально совершённых ими сделок. К счастью, поскольку на HYDRA по некоторым магазинам (число сделок менее 500) указывается точное число совершенных сделок, данная конвертация возможна.

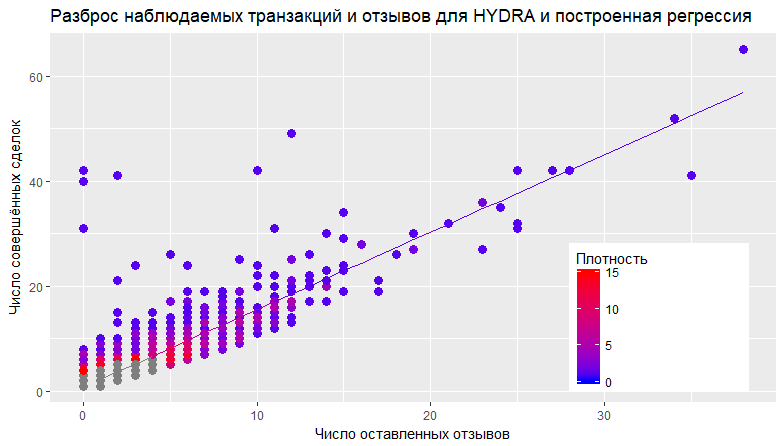
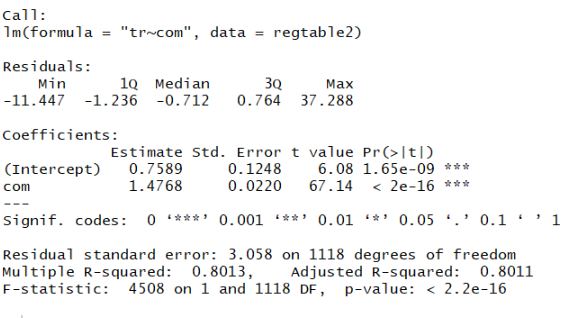


Рисунок 10 Диаграмма разброса и регрессия сделок на комментарии. Серый цвет: плотность более 15

Причины, по которым возникают серьезные основания в возможности такой конвертации просты: в силу специфичности анонимных площадок (наличие механизма репутационных потерь) покупатели постоянно обращают внимание на оставленные до них отзывы о товаре – они выступают информационными сигналами, говорящими о качестве товара (произведенного, к тому же, в кустарных условиях) и надёжности поведения продавца. Если в магазине давно не появляется новых положительных отзывов, то многие покупатели начнут думать, что магазин закрылся или стал обманывать клиентов. Поэтому все магазины стремятся к тому, чтобы их покупатели оставили отзыв об опыте употребления товара.

Таблица 4 Результаты регрессии числа сделок на оставленные отзывы на площадке HYDRA



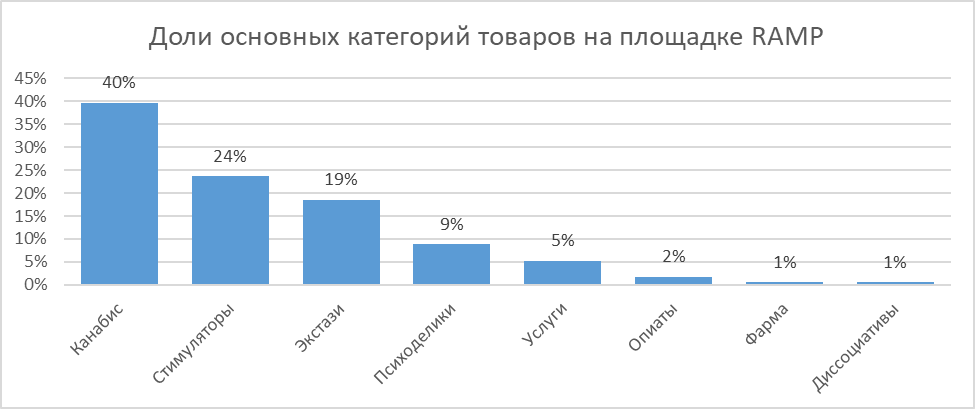
В качестве оценки соотношения между совершенными сделками и оставленными отзывами использовалась простая парная регрессия с константой. Оценивалось изменение за период между снимками и числа совершенных за всё время сделок наблюдавшихся для магазина , где , множество магазинов, для которых отображается точное число совершенных сделок. Хотя данные, использовавшиеся для данной регрессии имеют панельный вид, допускается, что нет никаких оснований наличия разницы между разными магазинами по свойству конвертации отзывов.

Результаты построенной модели неудивительны: диаграмма разброса наблюдений изображена на рисунке 10, а статистика модели отражена на таблице 4. Оба коэффициента отличны от 0 на 5% значимости, также модель имеет высокий на уровне 80% объяснённой вариации зависимой переменной. Хотя не найдено оснований подозревать гетероскедастичность, применение робастных ошибок не меняет полученных выводов (см. приложение 3). По результатам построенной регрессии для дальнейшей работы принята модель, в которой число сделок в примерно полтора раза превышает число комментариев:

## Статистический анализ данных по RAMP

Предварительный анализ структуры RAMP показал, что обе площадки весьма похожи между собой, поэтому в этом разделе фокус будет смещен в пользу ещё не затронутых характеристик и отличий между двумя площадками.

Рисунок 11 Разбиение товаров площадки RAMP по категориям



На рисунке 11 изображена относительная структура RAMP по категориям товаров (более подробна информация изложена в приложении 5). Как и у HYDRA, самыми распространёнными товарами являются конопля и стимуляторы, однако доля конопли значительно больше: 40% против 26% у HYDRA. Стимуляторы занимают примерно такую же долю; в то же время эйфоретики занимают примерно такую же долю: 19% у RAMP против 17% у HYDRA, на которой экстази выделено в отдельную группу. Если говорить о том, каковы объемы площадок относительно друг друга, то на этот счёт сделано не так много изысканий. Однако можно взглянуть на число обнаруженных товаров на одну и ту же дату произведения снимков: если взять снимок от 16 февраля, то на RAMP обнаружено 3437 товаров, на HYDRA – 2108. Таким образом, при допущении что качество обоих снимков одинаково, можно сделать вывод, что RAMP крупнее HYDRA, а пропорции их размеров находится на уровне около 60% к 40%. В отличие от последней, RAMP до сих пор является площадкой, где продаётся безусловно только ПАВ. Недавно один из магазинов стал продавать знания о том, как выращивать коноплю или создавать индивидуальный VPN, но это – в порядке исключения.

#### Оценка денежного объёма совершаемых транзакций на RAMP

Следующей важной частью работы является оценка объема анонимной торговли ПАВ на примере площадки RAMP. Идея основана на том, что для каждого товара известно количество оставленных отзывов в разные моменты времени и цены опций. Благодаря построенной выше регрессии транзакций на комментарии можно оценить число покупок каждого товара за период времени, которое остаётся только умножить на некую цену, чтобы получить объем продаж для каждого отдельного товара площадки. Далее, если сложить все объемы продаж по товарам, то получится весь объем продаж площадки.

Данная идея имеет три слабых места. Во-первых, неизвестно реальное число покупок, и для его оценки используются оставленные отзывы. Во-вторых, пользователи покупают опции какого-либо одного товара, которые имеют разные цены, но отзывы отображаются единым потоком по данному товару. Иными словами, число оставленных отзывов неизвестно для опций, а известно только суммарное число отзывов по всем опциям товара. Поэтому необходимо решить, каким способом усреднить цену по опциям. В-третьих, каждый снимок RAMP имеет свои сбои при загрузке, поэтому не все существующие товары в него попадают, поэтому многие товары имеют пропущенные наблюдения в числе оставленных комментариев. В таблице 5 приведены условные товары, которые переживают разные процессы. Товар может постоянно попадать в выборку, в таком случае он переживает условно обычный процесс. Некоторые магазины иногда меняют названия своих товаров, из-за чего в ходе парсинга создается впечатление, что один товар исчезает, но вместо него появляется точно такой же. Из-за сложностей парсинга, решение данной проблемы займёт неоправданно много труда, поэтому необходимо таким образом оценить изменение отзывов, чтобы попытаться как можно больше нивелировать этот эффект.

Таблица 5 Наблюдаемое число отзывов для товаров, подверженным разным процессам



Для этого предлагается следующая стратегия: если товар не наблюдаем, делается допущение, что он не существует. Для каждого наблюдаемого товара будет оцениваться число отзывов , оставляемое в среднем за промежуток времени между датами снимков и . При этом те периоды, в которые товар не существовал, также будут учитываться при подсчете среднего значения. Поскольку для каждого наблюдения известно оставленное за всю историю жизни товара число отзывов, то нет необходимости при подсчете использовать промежуточные наблюдения отзывов – достаточно просто посмотреть на разницу между числом отзывом в начале времени наблюдения товара и в конце. Таким образом, для всей площадки суммарное число оставляемых в среднем за 1 период отзывов для товара равно:

где – даты первого появления товара в выборке и последнего нахождения в ней, а – даты начала и конца наблюдения выборки из всех загруженных товаров (даты первого и последнего снимка). Поначалу такой выбор может вызывать недоумения, поэтому в таблице 10 приведены альтернативные стратегии оценки числа комментариев на площадке. Более подробно сравнение этих метрик приведено в приложении 1. После подсчёта для каждого товара , оно умножается на некую цену этого товара и получается оценка дневного денежного оборота площадки RAMP, которая для удобства умножается на 30, чтобы получить месячную оценку:

Далее необходимо выбрать цену , по которой считается денежная стоимость одной совершенной сделке по покупке товара . После просмотра некоторого числа html-страниц товаров выдвинуто предположение, что продавец выбирает разбивку на опции по следующим соображениям. Указывается опция с очень небольшим весом и ценой для новых покупателей, которые желают проверить качество товара или испытывают проблемы с ликвидностью и поэтому на данный момент не готовы сразу одной покупкой удовлетворить весь свой спрос на товар. Далее подбирается пара-тройка таких опций, которые оптимальны для постоянных покупателей: цена и количество товара заметно выше, размер клада учитывает тот факт, что каждый раз при его поиске и подборе покупатель рискует собой, а также попросту тратить своё время и силы. Далее есть опции для оптовой покупки, при этом цена опций настолько высока что становится ясным, что каждая сделка не может быть заключена на такую сумму. И более того, часто встречаются опции с нереалистично высокими ценами (например, 1 000 000 руб.), часто это делается для привлечения внимания покупателей по каким-либо причинам.

В связи с этим, применение среднего арифметического как способа агрегировать информацию по ценам имеет свои недостатки. Поэтому наряду с ним объем сделок считался с помощью и других способов тоже. Среди них расчёт медианы и подсчёт среднего арифметического при искусственном ограничении установлении потолка цен, выше которого цены не принимаются в расчёт.

Таблица 6 Оценка месячного объема сделок на RAMP в зависимости от способа агрегирования



В итоге, в таблице 6 представлены оценки денежного объема сделок, совершаемых на RAMP за месяц, в зависимости от того, как выбрать . В число таких цен входят минимальные и максимальные цены по опциям каждого товара , чтобы иметь представление о низшем и высшем пределе интервала, в котором можно дать оценку представленным в этой работе способом. Согласно таблице, месячный объем сделок на RAMP, рассчитанный по простому среднему, составляет 37 млрд. руб. Однако если при расчете не принимать внимание цены выше хотя 100 000, то оценка резко падает до 2 млрд. руб. В целом, наиболее адекватный реализации это использование медианной цены, при котором выходит, что за месяц через площадку проходит 1.66 млрд руб. Стоит принимать во внимание, что при медианной цене не учитываются сделки, совершаемые оптом, поэтому следует уточнить, что речь идет об оценке месячного объема продаж в розницу. Можно привести сравнение с заявлением главы ФСКН В. Иванова в 2015 году[[12]](#footnote-12), который оценивает месячный объем рынка наркотиков в России в около 125 млрд. руб. Если оценка, сделанная в данной работе, а также оценка В. Иванова верна, то стоит предположить, что рынок анонимной интернет торговли в розницу очень мал по сравнению со всем рынком наркотиков, однако не исключено, что в настоящее время значимая доля оптовых сделок совершается через анонимные площадки.

Таблица 7 Доли 20-ти крупнейших магазинов площадки RAMP



После произведение данной оценки на её основе был произведён дальнейший анализ площадки. В таблице 7 приведена разбивка денежного объёма сделок по магазинам. Как видно благодаря ней, доли крупнейших магазинов не превышают и 10%, что говорит о низкой степени отраслевой концентрации. Также был подсчитан индекс Герфиндаля-Хиршмана, значение которого на уровне 140 говорит о том же самом. Следовательно, можно выдвинуть гипотезу, что отрасль розничной продажи наркотиков имеет высокую степень рыночной конкуренции.

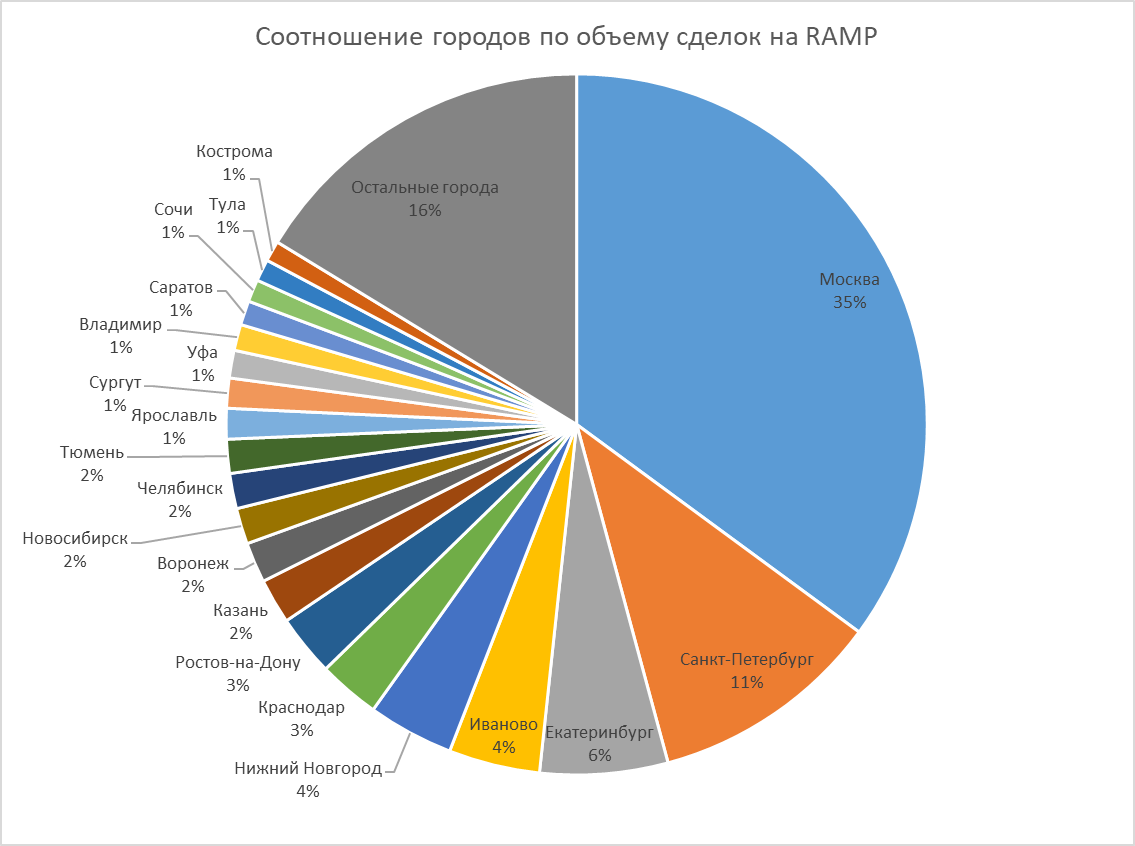


Рисунок 12 Доли городов РФ в общем объеме сделок на RAMP

На рисунке 12 представлена региональная структура RAMP. Москва и Санкт-Петербург занимают суммарно почти половину всей розничной торговли ПАВ. Далее за ними следуют города – центры субъектов РФ. В выборку также попали города других стран: упоминается Минск, Киев и несколько других крупных населённых пунктов Украины, однако объём сделок в зарубежных городах ничтожно мал (см. приложение 9).

## Построение временного ряда активности на RAMP

Имеющейся информации, полученной в ходе парсинга теоретически должно хватить для того, чтобы было возможным построение индекса активности на RAMP за равные промежутки времени. Идея похожа на тот же принцип, по которому оценивался объем сделок на RAMP, но теперь подсчёт изменения количества комментариев должен происходить от снимка к снимку, а не между последним и первым наблюдением товара в выборке. При этом для простоты решено отказаться от конвертации числа отзывов в число сделок. Поэтому полученный индекс, хотя и не дает точное представление о реальном объеме сделок, тем не менее способен показать, как это количество сделок изменяется во времени. Если все выше сказанное обобщить в математической формуле, то получится:

где – индекс активности на RAMP в момент времени , – условная цена на товар , – количество отзывов для товара на момент времени .

Однако при его построении возникают две принципиальные проблемы. Во-первых, неизвестна точная цена каждой сделки. Как уже было отмечено ранее, на странице каждого товара предлагается выбрать покупки из числа нескольких опций, каждая из которых отличается по цене, количеству товара и месту закладки. В то же время, количество комментариев не показано отдельно для каждой опции, поэтому возникает вопрос, на какую цену умножать изменение комментариев. В качестве возможного варианта в этой работе предлагается использование медианной цены по опциям.

Такой вариант может быть обоснован (см. метод оценки денежного объема сделок на RAMP выше), и в то же время он прост в реализации. При его реализации делается предпосылка о том, что продавцы сами так подбирают опции, чтобы медианная цена отражала объём средней сделки.

Следующей проблемой является то, что любой товар далеко не всегда встречается в каждом снимке, иными словами, для каждого товара присутствует довольно много пропущенных наблюдений числа комментариев (см таблицу 5). Более того, число комментариев в любой момент времени известно с точностью до 10 (в силу того, что получение информации о точном числе комментариев потребовало бы слишком большого времени загрузки каждого снимка). Это объясняется тем, что на RAMP есть возможность посмотреть каждый когда-то кем-то оставленный комментарий, но не все они присутствуют на одной странице. Если это число комментариев на момент времени , то число комментариев, которые наблюдаются в произведенных снимках:

То есть, если, например, по некоторому товару было оставлено 46 отзывов, то для их отображения требуется 5 страниц (по 10 отзывов на первые 4 страницы, и 6 отзывов на 5-ую страницу), поэтому наблюдаемое количество комментариев в снимке равно 50. В итоге вторая проблема состоит в том, что по каждому товару имеется много пропущенных наблюдений количества комментариев и каждое наблюдаемое количество товаров отличается от реального количества, но не более чем на 9 отзывов. И если первая проблема (с поиском условной цены на каждой товар) имеет простое решение, то для того, чтобы решить вторую проблему требуется применение инструментов эконометрики временных рядов.

## Восполнение пропущенных наблюдений и сглаживание через фильтр Калмана

Хотя для каждого товара и имеются пропущенные наблюдения числа комментариев, его можно попытаться восстановить через интерполяцию уже имеющихся данных. Существуют простые способы это сделать, например, использование последних наблюдаемых значений величины или соединение двух краев наблюдаемых промежутков прямой линией. Однако, такой способ не учитывает всю накопленную информацию о поведении временного ряда для того, чтобы сделать прогноз пропущенного наблюдения более точным. Более того, проблема несоответствия наблюдаемого числа комментариев реальному остаётся открытой, а её решение является важной задачей, поскольку если для построения индекса активности на площадке использовать наблюдаемое значение, то изменение этого индекса во времени может получится слишком резким: для большинства товаров разница в числе наблюдаемых комментариев будет равна либо 0, либо 10. Поэтому имеет смысл использовать такой способ, который позволил бы одновременно восполнить пропущенные наблюдения и сгладить слишком резкие колебания наблюдаемой величины. И таким способом является применение фильтра Калмана.

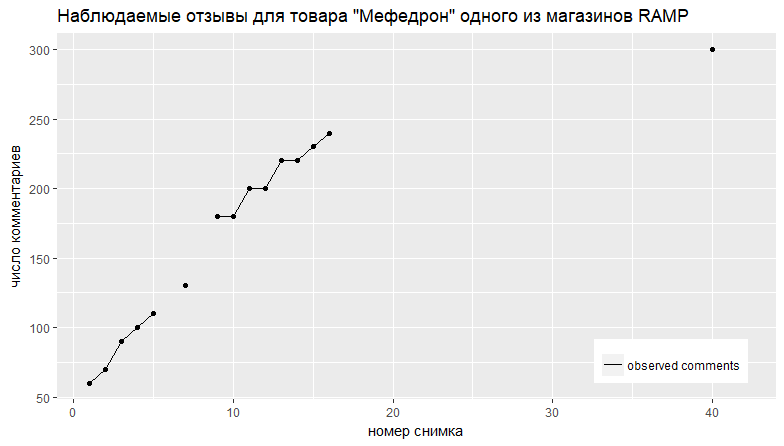


Рисунок 13 Пример наблюдаемого числа комментариев для одного из товаров.

В этой работе используется реализация фильтра Калмана в том виде, в каком он описан в (Durbin and Koopman, 2012) и (Helske, 2016)[[13]](#footnote-13). Фильтр Калмана использует Модель Пространства Состояний (State Space Model). Пусть имеются моменты времени (состояния) . Модель подразумевает наличие векторной переменной состояния , которая изменяется при переходе к новому состоянию :

При этом переменная состояния ненаблюдаема (латентна), доступно её неточное измерение (или несколько таких измерений за каждый период времени или даже только частичное измерение лишь некоторых компонентов вектора ), выражаемое вектором :

Пусть имеет форму , имеет форму , и имеет форму . Пусть в первом состоянии переменная имеет и изначально заданное распределение . Maтрицы , выступают системными параметрами модели, отражающими то, как разные компоненты векторов зависят между собой. Они вместе с ковариационными матрицами могут изменяться во времени. Как правило, системные матрицы задаются условиями конкретной модели, в то время как ковариационные матрицы неизвестны и оцениваются в ходе работы алгоритма.

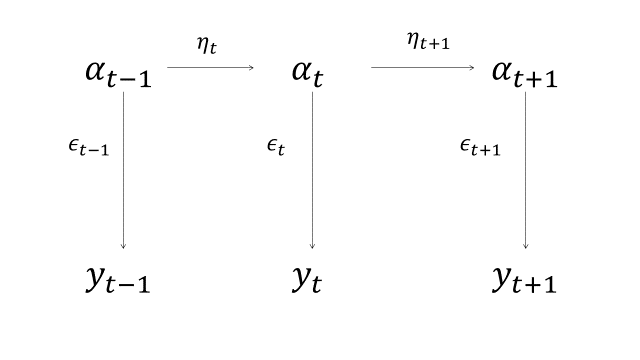


Рисунок 14 Графическая иллюстрация модели пространства состояний

Сам фильтр Калмана производит рекурсивную оценку условного распределения латентной переменной по накопленной истории предыдущих наблюдений. Оценка происходит в два этапа. Сначала оценивается распределение для нового состояния (значение переменной измерения для которого ещё неизвестно: при уже посчитанном условном распределении . Затем производится расчёт при только что подсчитанном Так как имеет нормальное распределение, то необходимо знать только условное математическое ожидание и условную дисперсию латентной переменной. Далее полученной оценки должно хватить для прогноза латентной переменной в будущем состоянии, для которого ещё недоступно значение наблюдаемой переменной . Собственно, такая задача и называется фильтрацией.

Однако, если необходимо сделать прогноз латентной переменной в прошлых состояниях, когда уже накоплена история наблюдения измеряемой переменной даже для будущих состояний, то было бы нецелесообразно использовать для этого оценку распределения , полученную с помощью фильтра – – поскольку это не вся доступная информация. Правильнее было бы производить оценку . И такая задача называется сглаживанием (smoothing). Сглаживание опирается на фильтр Калмана: сначала через фильтр происходит подсчёт , затем рекурсивным способом в обратную сторону происходит подсчёт . После того, как произведена оценка условного распределения для каждого состояния латентной переменной, применяется метод максимального правдоподобия для оценки ковариационных матриц и системных матриц . Это происходит благодаря известным измерениям : если и при этом вместе с имеют нормальное распределение, то и имеет нормальное распределение. После применения метода максимального правдоподобия становится возможным оценивание значений на основе математических ожиданий из оцененных распределений .

Для того, чтобы применить на практике сглаживание методом фильтра Калмана, необходимо сперва построить модель пространства состояний для конкретного случая с наблюдением числа комментариев. Очевидно, что временной ряд, состоящий из наблюдений числа комментариев , является нестационарным процессом. Поэтому пусть для сглаживания будет использоваться Модель местного линейного тренда (Local Linear Trend Model), в которой латентное число комментариев будет иметь предрасположенность к росту :

Тогда в векторной форме модель пространства состояний может быть описана следующим образом:

где:

Тогда единственные параметры, которые остается оценить методом максимального правдоподобия, это скаляр и матрицу , а также распределение первого состояния латентной переменной .

Для того, чтобы на практике построить такую модель для каждого товара и произвести сглаживание, использовался пакет программного обеспечения KFAS, реализованный в R. Точный код, описывающий построение модели можно увидеть в приложении 8. В результате удаётся успешно выполнить обе поставленных задачи – восстановление пропущенных наблюдений и сглаживание ряда. Пример работы фильтра Калмана виден на рисунке 15. На нём видно, что сглаженный ряд проходит насквозь наблюдаемый ряд, зачастую оказываясь то выше наблюдаемого числа комментариев, то ниже. Это может показаться проблемой, поскольку известно, что наблюдаемое число комментариев завышает реальное, следовательно, сглаженный ряд должен нивелировать этот эффект. Однако на самом деле этот аспект никак не скажется на результатах, поскольку для построения временного ряда индекса активности на RAMP используется только изменение в числе комментариев за период между наблюдениями, а не само число комментариев.

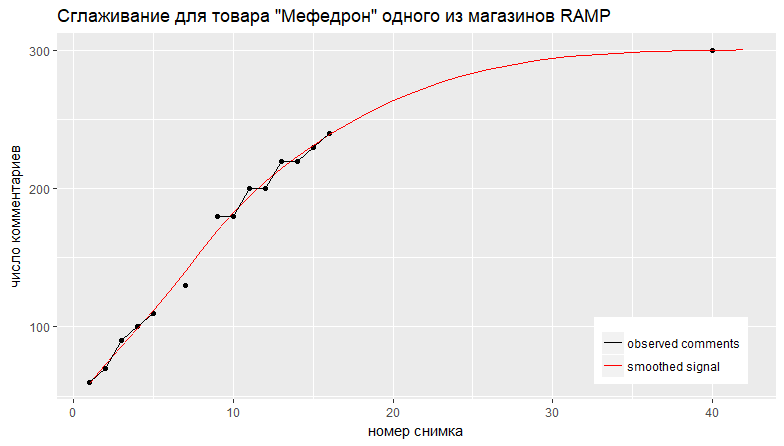


Рисунок 15 Применение фильтра Калмана для одного из товаров площадки RAMP

## Как выглядит динамика активности на RAMP

После того, как для каждого товара произведено сглаживание алгоритмом Калмана, имеется достаточно данных, чтобы построить индекс активности на RAMP. Для его построения использовались значения медианной цены и изменений сглаженного значения комментариев:

Полученный временной ряд показан на рисунке16. Первое, на что следует обратить внимание, это на то, как полное отсутствие снимков с середины февраля по середину апреля отражается на этом индексе. Поскольку промежуток пропущенных наблюдений оказался слишком большим, практически никакой информации о том, как вели себя покупатели и продавцы на площадке за тот период, извлечь не удается. Однако для пропущенных наблюдений в конце декабря удалось восстановить значение индекса, поскольку всего лишь два пропущенных наблюдения не так сильно влияют на работу фильтра Калмана.

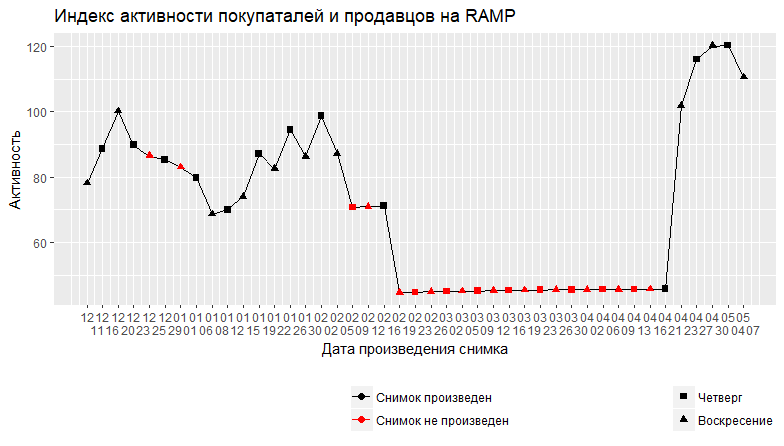


Рисунок 16 Индекс активности покупателей и продавцов на RAMP

Следующей интересной особенностью динамики индекса является его снижение на время новогодних выходных дней. По всей видимости, это может быть объяснено тем, что в этом время многие магазины выходили на каникулы и позволяли своим рядовым сотрудникам прекратить ежедневные закладки товара, чтобы иметь возможность выйти в отпуск. При этом остаётся непонятным, почему после окончания праздников активность на RAMP не восстановилась до былого уровня. Ещё одной интересной характеристикой данного индекса является то, как изменяется его значение в зависимости от дня недели, в который был произведен снимок. Как уже отмечено ранее, снимки производились два раза в неделю: в вечер четверга и вечер воскресения. Однако из полученной динамики нельзя сделать вывод о том, когда пользователи предпочитают производить покупки, поскольку если в декабре активность оказывалась выше на уикендах, то после праздничных выходных пользователи предпочитали делать покупки в основном в будние дни. В целом, необходимо дополнительное изыскание и более длительный ряд наблюдений, для того, чтобы иметь возможность ответить на этот вопрос.

## Влияние курса биткоина к рублю на активность на RAMP

Для совершения сделок покупателям необходимо обменять свои средства на биткоины и зачислить их на свой кошелёк в личном кабинете сайта. Однако валюта, в которой продаётся подавляющее большинство товаров на RAMP, является рублём. То есть, в ходе транзакции между контрагентами обмен происходит в биткоинах, но цены на ПАВ, хотя и указаны в обоих валютах, являются стабильными и не изменяются только в рублёвой валюте. Вместе с тем, известно, что биткоин волатилен и имеет спекулятивную природу.[[14]](#footnote-14) Поэтому возникает вопрос, способно ли изменение курса биткоина к рублю оказывать влияние на активность на RAMP или нет.

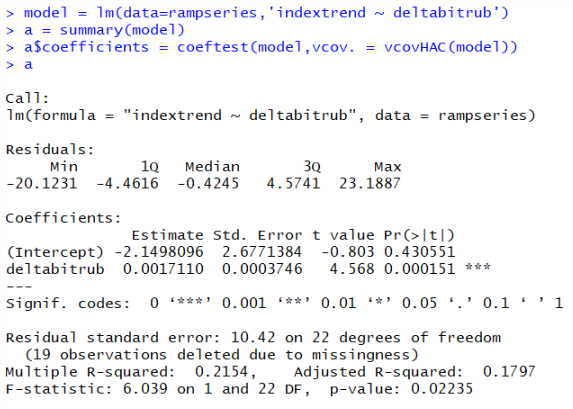
В ходе работы выдвинута гипотеза, что динамика индекса активности RAMP испытывает на себе влияние текущие изменения курса биткоина в следующей форме:

где ошибка является белым шумом, влияющим на активность на RAMP. Данная идея кажется обоснованной. Изменение курса биткоина должно непосредственно влиять на покупательную способность денежных средств, которые имеются у пользователей в личных кабинетах на сайте RAMP, где все цены указаны в рублях. К сожалению, для оценки такой гипотезы требуется предпосылка о наличии экзогенности курса биткоина, проверка которой требует дополнительных теоретических изысканий. Поэтому вместо гипотезы об истинном влиянии курса биткоина на динамику активности акцент будет смещен в пользу простой проверки наличия прогнозной силы текущего изменения курса биткоина на активность в рамках площадки RAMP. То есть, в оцениваемой модели коэффициент будет содержать в себе не только непосредственное влияние изменения курса биткоина, но и влияние ненаблюдаемых факторов, вызывающих одновременную причинность как для зависимой переменной, так и для регрессора. Также следует заметить, что активность на RAMP вызвана не только факторами, изменения которых отражаются на изменении , что может вызвать наличие автокорреляции ошибок в модели. При дальнейшей работе решено было отказать от значения индекса активности в период с 19 февраля по 21 апреля, когда наблюдение за RAMP было прервано (см. рисунок 16), поскольку восстановленное значение индекса плохо справляется с отображением реального значения индекса активности в тот период. В качестве ряда обменного курса использовались цены на момент закрытия торгов в дату, совпадающую с плановой датой производства снимка (см. приложение 8)

Прежде чем оценить коэффициенты модели методом МНК, необходимо убедиться в том, что регрессор и зависимая переменная являются стационарными рядами. Для этого использовались тесты ADF и KPSS для обоих рядов. При их производстве было замечено, что, в связи наблюдением за динамикой индекса активности, а также в связи с тем, что в каждом вновь произведённом снимке RAMP количество обнаруженных товаров постоянно росло, индекс активности, скорее всего имеет стационарную природу, увязанную с временным трендом. Поэтому тесты проводились в спецификации с трендом. Результаты ADF и KPSS для индекса активности RAMP видны в приложении 10. Оба теста при 5% уровне значимости свидетельствуют о том, что детрендированный ряд индекса активности является стационарным. Затем, были проведенные такие же тесты для также говорящие о стационарности для этого ряда, если использовать 5% уровень значимости, см. приложение 11.

Далее, были оценены коэффициенты модели в спецификации с детрендированной динамикой индекса активности. Результаты приведены в таблице 8. Несмотря на использование HAC оценок ошибок, а также маленький ряд наблюдений, тесты на ненулевое значение коэффициентов показывают на наличие связи между активностью на RAMP и курсом биткоина на 1% уровне значимости. Таким образом, можно сделать вывод о том, что курс биткоина как минимум имеет прогнозную силу на поведение индекса активности RAMP.

Таблица 8 Оцененные коэффициенты и тесты на их значимость с использованием HAC оценок стандартных ошибок



Следует также обратить внимание на значение , которое осталось достаточно высоким после детрендирования зависимой переменной. Также видно, что константа не обязательно играет значение в оцененной модели.

После установления наличия прогнозной силы также построено предсказание поведения динамики активности в ретроспективе, которое приведено на рисунке 17.

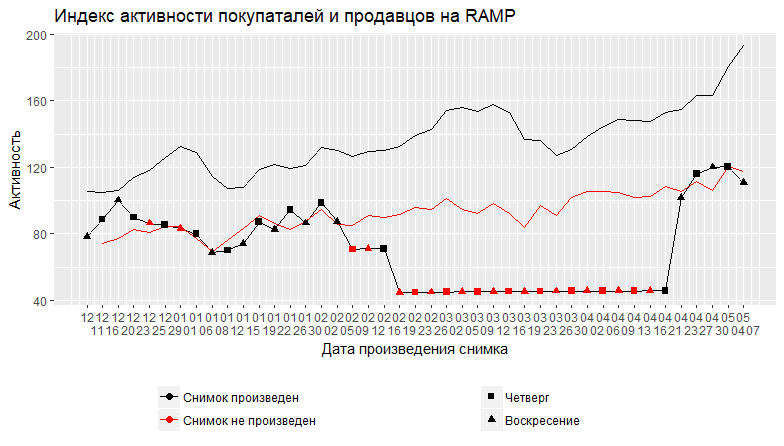


Рисунок 17 Прогноз поведения индекса активности по изменению курса биткоина в ретроспективу. Верхний ряд – отмасштабированный курс биткоина, нижний ряд красного цвета – прогноз в ретроспективу, нижний ряд черного цвета – посчитанный индекс активности.

## Заключение

В ходе проделанной работы была собрана база данных, отражающая состояние моментальных магазинов RAMP и HYDRA, произведён парсинг html-страниц из этой базы, и статистический анализ структуры обоих площадок (их моментальных магазинов). Выяснилось, что основным типом товаров, который продается на площадках – это легкие наркотики, а самая распространённая категория - конопля. Далее была произведена оценка месячного объёма розничных сделок на RAMP, давшая значение 1.6 млрд. руб. (при использовании медианной цены). Оказалось, что RAMP имеет низкий индекс отраслевой концентрации, говорящий о высокой степени конкуренции на площадке. Также оказалось, что Москва, Санкт-Петербург и Екатеринбург занимают самые большие доли оборота данной площадки. Затем с помощью фильтра Калмана был построен индекс активности на RAMP, который, как оказалось, может предсказываться курсом биткоина по отношению к рублю, поскольку было доказано, что он имеет прогнозную силу на данный индекс.

Следует заметить, что все выше сказанное производилось для российских площадок, скорее всего, впервые.

Дальнейшем направлением для исследования можно назвать продолжение производства снимков для RAMP и HYDRA, совершенствование методологии формирования индекса активности на торговой площадке (отказ от использования медианной цены в пользу других вариантов, применение сезонности), а также оценка объема сделок и его динамики для площадки HYDRA. Более того, перспективным представляется изменение подхода к анализу RAMP с абсолютных величин на относительные (анализ цен и долей категорий товаров в разных регионах РФ). В данной работе почти не был затронут анализ факторов формирования цен на товары в зависимости от показателей внешнего мира и структуры разных типов ПАВ, хотя он может оказаться плодотворным.

## Приложения

#### Приложение 1. Более подробное сравнение метрик для оценки среднего числа оставленных комментариев

Если считать данную оценку как:

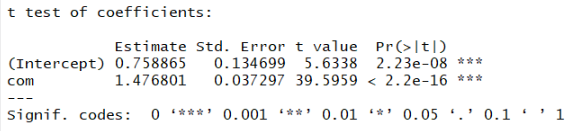
то можно заметить, что данная метрика сильно завышает число оставленных в среднем отзывов, когда встречаются товары, которые переживают любой необычный процесс (см. столбец «деление на ср. кол-во в подпериод»). Эта метрика показывает среднее количество отзывов, которые появляются у любого товара за период его активного существования, и никак не зависит от того, какова была продолжительность этого периода. Поэтому агрегирование таких метрик для всех товаров не подойдет для подсчёта среднего числа отзывов на всей площадке с одним единым для всех товаров периодом. Использование суммы изменений (см. столбец «сумма дельт») также не подходит, поскольку всякий раз, когда товар имеет пропущенное наблюдение, такая метрика занижает реальное число сделок.





Приложение 2. Код R, применявшийся для построения сглаженного сигнала

Приложение 3. Применение робастных ошибок (метод HC3) для оценки значимости коэффициентов в регрессии отзывов на комментарии для HYDRA



Приложения 4 и 5 Разбиение обоих площадок по подкатегориям товаров (все обнаруженные товары за всё время ведения наблюдений)



Приложение 6 Разбиение магазинов площадки HYDRA в зависимости от числа совершенных сделок



Приложение 7 Разбиение всех обнаруженных за всё время наблюдений товаров площадки RAMP по категориям



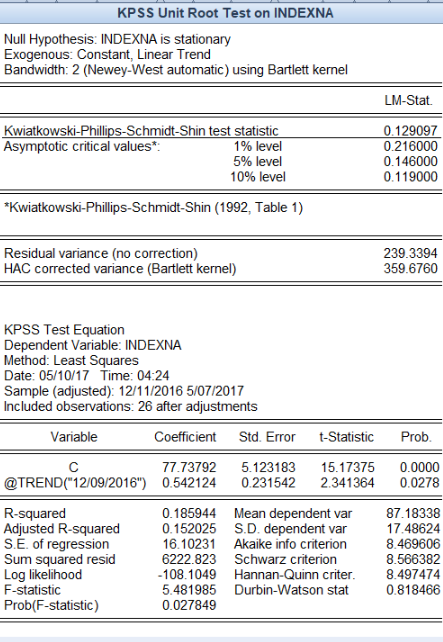
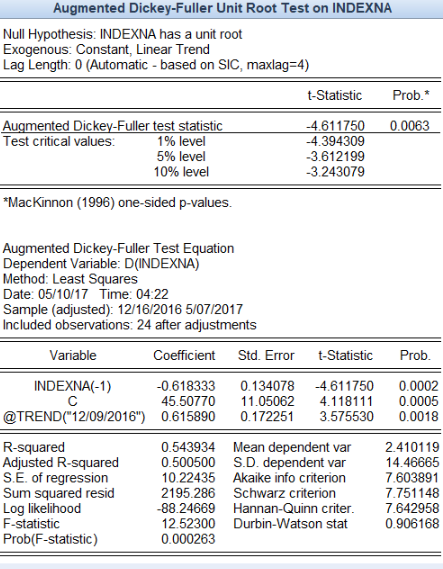
Приложение 8 Индекс активности покупателей и продавцов на площадке RAMP, а также курс биткоина, использованный для регрессии



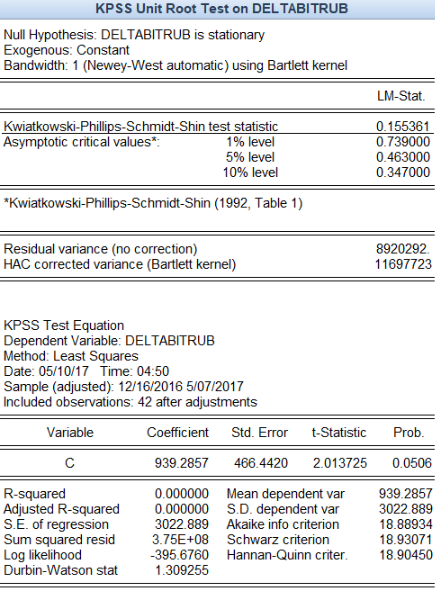
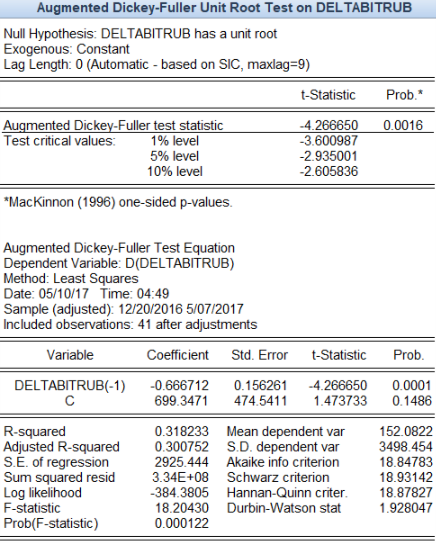
Приложение 9 Месячный объем сделок на RAMP в разбиении по городам



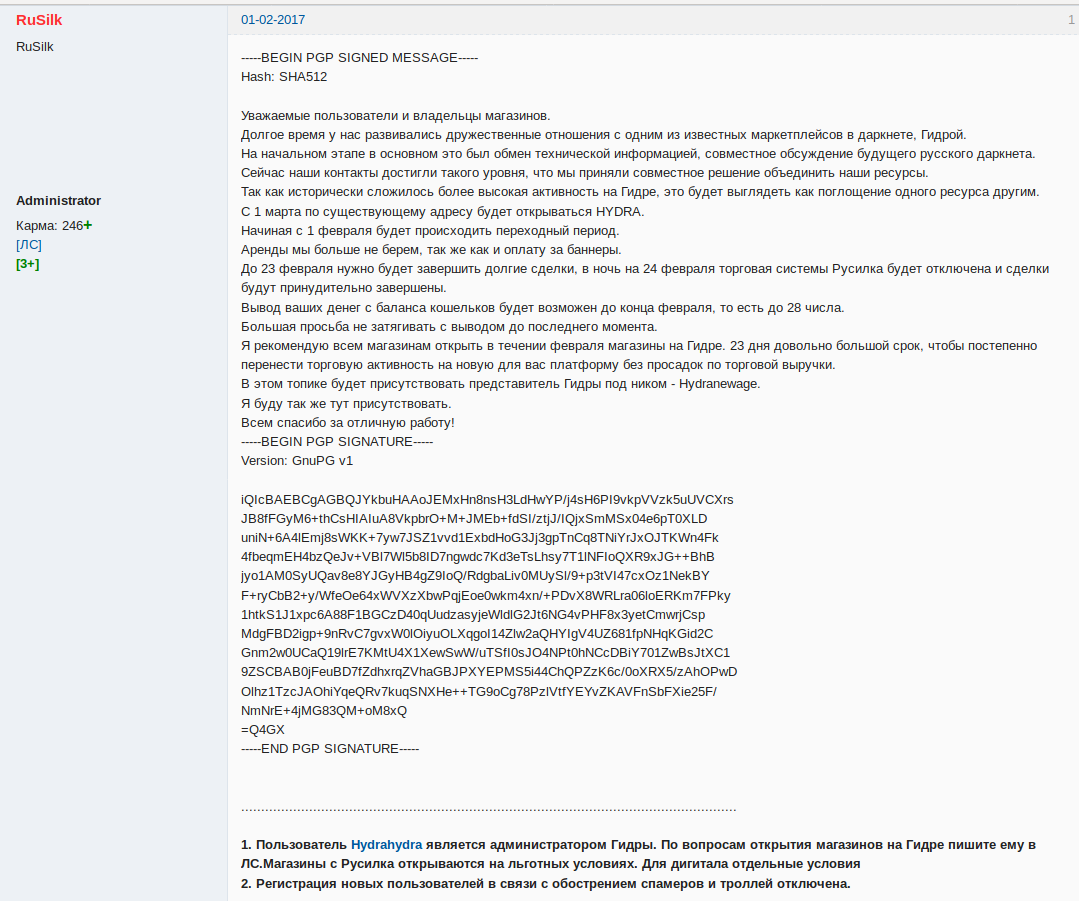
Приложение 10 ADF и KPSS тесты для индекса активности RAMP в спецификации с трендом



Приложение 11 ADF и KPSS тесты для изменения курса биткоина



Приложение 12 Объявление о поглощении RuSilkRoad площадкой HYDRA



## Список литературы:

1. Van Hout, M. C. V., & Bingham, T. (2013). ‘Surfing the Silk Road’: A study of users’ experiences. International Journal of Drug Policy, 24(6): 524-529.

2. Van Hout, M. C., & Bingham, T. (2014). Responsible vendors, intelligent consumers: Silk Road, the online revolution in drug trading. International Journal of Drug Policy, 25(2): 183-189.

3. Barratt, M. J., Ferris, J. A., & Winstock, A. R. (2014). Use of Silk Road, the online drug marketplace, in the United Kingdom, Australia and the United States. Addiction, 109(5), 774-783.

4. Christin, N. (2012, May). Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. In Proceedings of the 22nd international conference on World Wide Web (pp. 213-224). International World Wide Web Conferences Steering Committee.

5. Dolliver, D. S. (2015). Evaluating drug trafficking on the Tor Network: Silk Road 2: The sequel. International Journal of Drug Policy, 26(11), 1113–1123.

6. Munksgaard, R., Demant, J., & Branwen, G. (2016). A replication and methodological critique of the study “Evaluating drug trafficking on the Tor Network”. International Journal of Drug Policy, 35, 92-96

7. Helske., J. (2016, December). KFAS: Exponential Family State Space Models in R. ArXiv e-prints.

8. Durbin J., Koopman S.J. (2012). Time Series Analysis by State Space Methods. 2nd edition. Oxford University Press, New York.

9. Baur, D. G., Hong, K., and Lee, A. D., (2016, June). Virtual Currencies: Media of Exchange or Speculative Asset? SWIFT Institute Working Paper No. 2014-007.

**Электронные ресурсы:**

10. Branwen, G.; Christin N.; Décary-Hétu, D.; Munksgaard, R.; StExo; El Presidente; Anonymous; Daryl Lau; Sohhlz; Kratunov, D.; Cakic, V.; Van Buskirk; & Whom. (2015, July). Dark Net Market archives, 2011-2015, Web. [access date] www.gwern.net/Black-market%20archives

**Статьи из электронных СМИ:**

11. Туровский, Д. (15.08.2016). Зашифрованное подполье. Оружие, наркотики и личности на продажу в русском «глубоком интернете». Meduza.  
https://meduza.io/feature/2016/08/15/zashifrovannoe-podpolie

12. Тодоров, В., Евтушенко, А. (23.08.2016). Наркотическая многоходовочка. Как старейший ресурс в русском Tor уничтожил конкурентов и обманул журналистов. Lenta.ru.

https://lenta.ru/articles/2016/08/23/runionvsramp/

13. Мерзликин, П. (23.08.2016). «Человеческая жизнь в России стоит очень дешево»: организатор торговли в Dark Web — о заказных убийствах, изнасилованиях и принципах. Paper.  
<http://paperpaper.ru/kill/>

14. (23.04.2015) ФСКН оценила наркорынок России в полтора триллиона рублей в год. Lenta.Ru

https://lenta.ru/news/2015/04/23/barko/

15. Greenberg, A. (11.11.2014). How a Russian Dark Web Drug Market Outlived the Silk Road (And Silk Road 2). Wired.

https://www.wired.com/2014/11/oldest-drug-market-is-russian/

16. Gil, P. (17.03.2017). The Darknet: Black Market and Sanctuary. Lifewire.

https://www.lifewire.com/darknet-black-market-2483129

1. How a Russian Dark Web Drug Market Outlived the Silk Road (And Silk Road 2)

   Date of Publication: 11.14.14. Time of Publication: 6:30 am [↑](#footnote-ref-1)
2. [↑](#footnote-ref-2)
3. Иногда участники знакомят условиями сделки только после возникновения конфликтной ситуации. Такая практика возможно благодаря шифрованию условий сделки: договор покупателя и продавца хранится на компьютере гаранта в зашифрованном виде. В случае конфликтной ситуации истец передает гаранту ключ для дешифрования, и последний отныне владеет информацией об условиях сделки. [↑](#footnote-ref-3)
4. Это **единственная** страна, доставка товаров внутри или из/в которую запрещены администрацией сайта, до сих пор, даже на Silk Road 3.1. [↑](#footnote-ref-4)
5. Согласно мнению журнала Wired:  
   Greenberg, A. (11.11.2014). How a Russian Dark Web Drug Market Outlived the Silk Road (And Silk Road 2). [↑](#footnote-ref-5)
6. Во всяком случае, научные работы по ним трудно найти. [↑](#footnote-ref-6)
7. А также закрывшиеся в 2015 году площадки R2D2 и Amberoad [↑](#footnote-ref-7)
8. Greenberg, A. (11.11.2014). How a Russian Dark Web Drug Market Outlived the Silk Road (And Silk Road 2). Wired. [↑](#footnote-ref-8)
9. Для RAMP встречаются снимки, в которых измеряемая метрика полноты (доля загруженных товаров в числе обнаруженных) находится на очень высоком уровне, в то время как качество, определяемое по доле загруженных товаров в числе заявленных, находится на низком уровне: это хорошо иллюстрирует то, что высокое значение данной метрики не дает гарантий. Однако для HYDRA, для которой недоступна информация о заявленном числе товаров, можно довольствоваться хотя бы тем, что число загруженных товаров всегда находилось на стабильном уровне. [↑](#footnote-ref-9)
10. В приложении 19 можно более подробно ознакомиться с этим событием. [↑](#footnote-ref-10)
11. К сожалению, из-за неоптимальных настроек загрузки снимки до 15 феврали не предусматривали возможности собирать данную информацию. [↑](#footnote-ref-11)
12. (23.04.2015) ФСКН оценила наркорынок России в полтора триллиона рублей в год. Lenta.Ru [↑](#footnote-ref-12)
13. [↑](#footnote-ref-13)
14. Данное суждение сделано на основе выводов из статьи (Baur, Hong and Lee, 2016), в которой исследовалась природа данной криптовалюты. [↑](#footnote-ref-14)