**I. Описание работы**

Биткоин – стал синонимом цифровой экономики. Но не утихают споры является ли он прообразом валюты будущего, несёт ли в себе какую-то ценность (value), способен ли защитить от инфляции, является ли чисто спекулятивным инструментом и не является ли просто очередным «пузырём» (<https://www.cnbc.com/2018/01/16/bitcoin-is-a-pyramid-scheme-warns-ex-wells-fargo-ceo-dick-kovacevich.html>). С юридической точки зрения до сих не закрыт вопрос следует ли принимать его, как платёжное средство, или запретить для конвертации в товары внутри страны. Стоят вопросы по регуляции и налогообложению действий майнеров, которые своими действиями по сути создают деньги из воздуха (<https://www.cnbc.com/2018/02/21/everything-you-need-to-know-about-bitcoin-and-your-taxes.html>). Всё это лишь позволяет отметить, что по сравнению с ситуацией ещё пару лет назад, зарождением биткойна в 2009-ом году, внимание к криптовалюте заметно увеличилось. Прежде скептические реакции о возможном будущем биткойна сменились на более серьёзную аналитику. Многие сайты начали выпускать ежедневные обзоры по крипторынкам. Даже в таких серьёзных изданиях, как Bloomberg, в день выходит несколько десятков новостей о криптовалюте. Появились специализированные сайты, которые пишут только о криптовалюте. Всё больше людей начинают следить за вводом новых протоколов, начинают разбираться в принципе работы блокчейна, в том, что такое lighting, зачем нужен segwit, как грядущий fork повлияет на цену криптоактива. Но как все эти знания могут помочь опытному инвестору? В данной работе я попытаюсь доказать или опровергнуть факт, что новостной фон вокруг биткойна содержит возможность для создания роботизированного алгоритма для предсказания движения криптовалюты в следующий момент.

В работе, используя базовые методы NLP (Natural Language Processing), я оцениваю, какой вклад дают публикации на новостных сайтах в будущую цену биткойна. В качестве новостных сайтов я выбрал пять наиболее известных сайтов: Bitcoin.com, Coindesk.com, Cointelegraph.com, Bloomberg, Forbes. Биткойн выбран, как одна из основных криптовалют. Моей задачей являлось предсказать рост или падение цены биткойна после публикации новости. При этом результаты моей работы можно распространить дальше на другие криптовалюты или другие новостные сайты. Также результаты работы можно использовать для создания собственных торговых роботов или просто для торговли на бирже. Кроме этого результаты работы являются проверкой теории эффективности рынка, т.е. что вся публичная информация уже учтена в цене криптовалюты. В случае если вся информация уже учтена в цене акции, построение подобных торговых роботов невозможно. В работе будут проанализированы, как дневные, так и минутные данные. Анализ минутных данных позволит говорить о внутридневных корреляциях и о том, как быстро рынок ребалансируется при публикации новостей.

**II. Биткойн, как криптовалюта**

В данном параграфе я рассмотрю основные аспекты работы биткойна и некоторые общие свойства, присущие другим криптовалютам и блокчейну. Я дам описание, как работает биткойн, как происходят транзакции, как работает крипто-кошелёк, что такое алгоритм консенсуса. Также я попытаюсь дать некоторое представление о том, почему биткойн является валютой, раскрою некоторые из взглядов на то является ли он пузырём или нет. Биткойн является инновацией не только благодаря тому, что является цифровой валютой, но и из-за блокчейна, который лежит в его основе, поэтому при описании биткойна я так же обращаю внимание на некоторые аспекты работы блокчейна и полезных свойств, которые заложены в нём. Следует отметить, что многие свойства биткойна могут быть обобщены в целом на любую криптовалюту, поэтому моя работа по описанию биткойна имеет практический интерес для тех, кто хочет разобраться, как работает этот финансовый инструмент.

**Биткоин**

Биткоин – это цифровая платёжная система, состоящая из цепи транзакций. В основе этой цепи лежит алгоритм консенсуса между участниками системы. Под консенсусом в сети мы имеем в виду, что транзакция в сети считается валидной, если за неё проголосовала большая часть пользователей системы. Биткоин является первой peer-to-peer платёжной системой, полностью поддерживаемой пользователями, без участия регулирующих органов и посредников. («Pantera Primer», Ronald A.Glantz, 2014) Что такое peer-to-peer сеть? Также её называют p2p, одноранговой, пиринговой или децентрализованной. P2P – это компьютерная сеть, созданная поверх основной сети (в нашем случае интернета) и основанная на равноправии участников. В такой сети отсутствуют выделенные серверы, а каждый узел (peer) является, как клиентом, так и выполняет функции сервера). (<https://en.wikipedia.org/wiki/Peer-to-peer>)

С точки зрения пользователя всё намного проще, биткойн просто является валютой в интернете. Хотя стоит отменить, что кроме того биткойн является самой перспективной и надёжной книгой учёта транзакций в мире. Надёжность и сохранность достигается за счёт того, что все данные много раз скопированы и хранятся на компьютерах пользователях сети. На момент написания этой статьи в интернете присутствовало 1592 криптовалюты (<https://coinmarketcap.com/all/views/all/>), при этом биткойн по капитализации занимал лидирующее место с капитализацией в 37% от всех существующих криптовалют (<https://coinmarketcap.com/charts/>).

Опишем принцип, по которому работает биткойн. Для работы с биткойном нужен специальный компьютерный софт или мобильное приложение, которые содержат информацию о цифровом кошельке пользователя и позволяют получать и отправлять биткойны через них. Получение и отправление биткойнов происходит посредством добавления новых записей в специальный электронный журнал транзакций, который называется «блокчейн». Этот журнал транзакций находится в публичном доступе, и любой пользователь может при желании скачать себе информацию обо всех транзакциях всех пользователей интернета. Подлинность каждой транзакции обеспечивается электронной подписью отправителя, таким образом только пользователь кошелька управляет тем, кому и сколько биткойнов он отправит. Также любой пользователь может осуществить процессинг транзакций в блокчейне, используя мощности своего компьютера или специальных серверов. Этот процессинг называют «майнингом». В вознаграждением за майнинг система даёт некоторое количество биткойнов.

Итак, каждый пользователь кошелька имеет публичный и приватный ключи. Аналогом публичного ключа является банковский счёт. Вы можете переводить деньги на любой публичной ключ. Но только обладать приватного ключа может перевести/потратить биткойны. Все публичные ключи лежат в открытом доступе, но мы не знаем, кому принадлежит какой ключ. Поэтому говорят, что адреса биткойнов псевдоанонимны. При создании кошелька и перевода на него биткойнов, кошелёк сигнализирует в блокчейн, что на него поступили деньги. Кошелёк генерирует два ключа – публичный и приватный (приватный может не генерироваться, если кошелёк создан на бирже, например, Bitsamp). Если вы захотите заплатить за какую-то услугу, то вы идёте к продавцу и говорите, что с вашего публичного ключа ему прийдут биткойны. Дальше мы идём в кошелёк или на биржу и переводим средства с нашего публичного ключа на публичный ключ продавца. Эта транзакция попадает в p2p сеть, которая проверяет её, смотрит, есть ли у вас деньги, и отправляет для записи в узлы сети. Таким образом, ваша транзакция попадает майнеру, который добавляет её в «блок». Блок состоит из серии транзакций разных пользователей. Блок передаётся в сеть дальше, где он попадает в общий журнал транзакций, т.е. в блокчейн. Теперь ваши деньги переведены продавцу. («Pantera Primer», Ronald A.Glantz, 2014)

**Преимущества биткойна**

* Свобода в совершении транзакций. Вы можете посылать деньги любому участнику сети в любом объёме в любое время. Сеть работает без выходных, без географических границ, без других финансовых ограничений.
* Маленькая комиссия за транзакцию. Теоретически процессинг транзакций может быть бесплатным для вас. Практически для ускорения процессинга вы платите небольшую сумму майнерам.
* Привлекательность для микротранзакций. Так как комиссия за транзакцию очень маленькая, то биткойн может использовать в развивающихся странах продавцами для осуществления микро платежей.
* Прозрачность сети. Вся информация о платежах находится в блокчейне. В любой момент кто угодно может проверить и верифицировать ваш платёж.
* Нейтралитет. Так как биткойн представляет из себя p2p сеть, т.е. децентрализирован, то никто не может манипулировать и контролировать его протокол.

(«Pantera Primer», Ronald A.Glantz, 2014)

**Биткойн, как денежная единица**

Ценность биткойна определяется возможностью использовать его в виде денежной единицы. Его принимают в виде платёжного средства, на него можно купить что-то, его можно обменять на другие денежные единицы. Кроме того, биткойн обладает и другими свойствами денег: продолжительный срок службы, портативность, возможность обменять на что-то, редкость, делимость на более мелкие единицы, несёт в себе некоторую цену (store of value), узнаваемость. Отличие от других денежных единиц состоит в том, что биткойн не несёт в себе никакие физические свойства (как золото или серебро) или его ценность не базируется на веру в органы власти (как обычные фиатные деньги). Биткойн целиком основан на математических свойствах. Так как биткойн обладает всеми свойствами денег, то его цена в главную очередь определяется принятием его другими контрагентами. Со временем всё больше контрагентов будут принимать биткойн в качестве оплаты за услуги и товары.

Как и у другой денежной единицы цена биткойна определяется спросом и предложением. Предложение биткойна ограничено математически – существует ограничение на общее количество биткойнов, которое будет сгенерировано системой. Спрос на биткойны определяется пользователями и участниками сети. В последнее время цена биткойна часто имела сильную волатильность из-за резкого наплыва и последующего исчезновения новых участников рынка, которые двигали цену вверх или вниз. В будущем, когда капитализация биткойна достаточно вырастет он будет менее подвержен подобным шокам.

(«Pantera Primer», Ronald A.Glantz, 2014)

**Является ли биткойн пузырём**

To be done

**Легализация биткойна**

To be done

**Алгоритм Консенсуса**

<https://habrahabr.ru/company/bitfury/blog/327468/>

**DAO**

<https://medium.com/universablockchain/decentralized-autonomous-organization-what-is-a-dao-company-eb99e472f23e>

**III. Источники данных**

Для моделирования эффекта выхода публикации на курс биткойна я буду использовать несколько источников новостей. Всего будет рассмотрено пять разных сайтов: Bitcoin.com, Coindesk.com, Cointelegraph.com, Bloomberg, Forbes. При отборе источников я руководствовался в первую очередь узнаваемостью и цитируемостью источника. Так первые три сайта появляются первыми в выдаче в Яндексе и Google по запросам «биткойн» и «биткойн новости». Bloomberg выбран, как один из основных источников новостей о финансовых рынках. Forbes выбран, как альтернатива Bloomberg. По содержанию статьи на Forbes являются более полными и предоставляют более широкую аналитику, чем публикации на сайтах bitcoin или coindesk, но по характеру контента Forbes больше похож на них, потому что при публикации контента придерживается модели «contribution model» (статьи написаны самими читателями или независимыми журналистами).

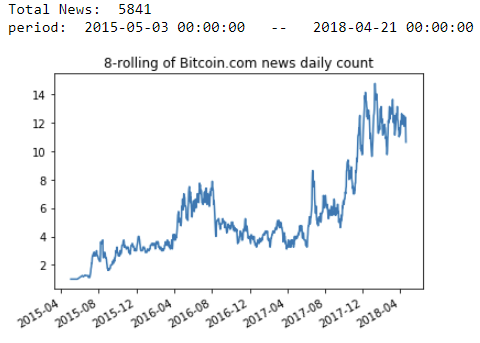
В качестве предварительной гипотезы я могу предположить, что контент на сайтах Bitcoin.com, Coindesk.com, Cointelegraph не будет оказывать значимого влияния на курс биткойна, т.е. по содержанию контента мы не сумеем предсказать движения криптовалюты. Предположение сделано на основе того, что данные сайты не предоставляют аналитику в реальном времени и по больше части публикуют новости post factum. Таким образом, новости будут сильно коррелировать с динамикой внутри дня, но не оказывать влияния на движение валюты на следующий день. Аналогичную ситуацию мы будем наблюдать с сайтом Forbes.com; отличие состоит в том, что высокая посещаемость сайта может создавать критическую массу и всё-таки оказывать влияние на движение биткойна. Bloomberg предоставляют аналитику финансовых рынков в реальном времени и является основным источником информации для большинства трейдеров, поэтому для него я ожидаю увидеть значимую корреляцию с движением криптовалюты. Кроме того данные о новостях с сайта Bloomberg доступны с точностью до минуты публикации, что даёт возможность исследовать вклад публикации в цену биткойна в следующую за публикацией минуту.

Ниже даны подробные описания каждого источника данных с краткими пояснениями по структуре исходных данных.

**Bitcoin.com**

Компания, предоставляющая сервисы по работе с Bitcoin Core и Bitcoin Cash. Через сайт можно покупать и продавать криптовалюты, а также работать с криптокошельком. На самом сайте содержатся подразделы, где можно посмотреть новости из крипто рынка, построить графики для анализа цены крипто валюты, поиграть в онлайн казино на крипто валюту. Также через сайт существует возможность поучаствовать в майнинге криптовалюты посредством участия в облачной бирже мощностей (<https://en.wikipedia.org/wiki/Bitcoin.com>)

Главный офис компании находится в Токио. CEO Роджер Вер. Домен функционирует с 2014-ого года. Сайт занимается популяризацией биткойна. На апрель 2018-ого года сайт занимал первое место в Яндексе и первое место в Google по запросу «bitcoin news».



**Coindesk.com**

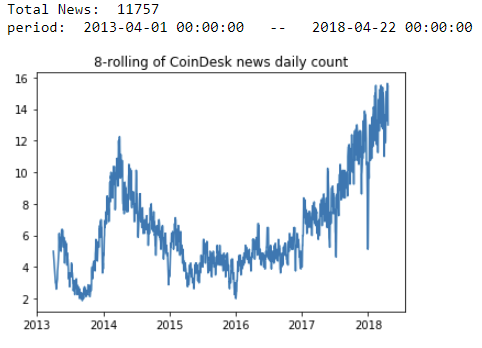
Сайт был запущен в мае 2013-ого года. Портал специализируется на новостях о криптовалютах. Сайт основал Шакил Кан (Shakil Khan), в последствии сайт приобрела компания Digital Currency Group. (<https://en.wikipedia.org/wiki/CoinDesk>). Посещаемость сайта в день составляет 10 миллионов уникальных человек, которые в целом делают 50 миллионов посещений. Coindesk присутствует в социальных сетях: твиттере, фейсбуке, линкедине, - где на него подписаны более 600 000 подписчиков. (<https://www.coindesk.com/about/>)

С сентября 2013-ого года сайт публикует Bitcoin Price Index (CoinDesk BPI). Индекс является средней ценой биткойна нескольких крипто бирж. Индекс часто цитировался в ведущих новостных изданиях, таких как The Wall Street Journal, Financial Times, Bloomberg, CNBC. (<https://www.coindesk.com/about/>)

Кроме того, CoinDesk проводит ежегодную конференцию Consensus summit. Это самая большая встреча в мире, посвящённая технологии блокчейн. Конференция проходит в городе Нью-Йорк. В мае 2017-ого года на саммит было продано свыше 2700 билетов. CoinDesk также проводит саммит Consensus: Invest для институциональных инвесторов, хедж фондов, банков и интересующихся инвесторов, где обсуждается, как инвестировать, где хранить, как оценивать стоимость крипто активов.

Другой известной частью деятельности CoinDesk является отчёт “The State of Blockchain” – самый подробный анализ рынка и перспектив технологии блокчейн

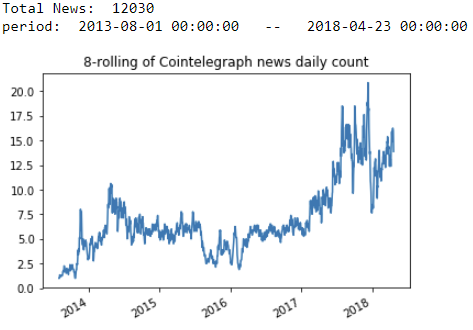
На апрель 2018-ого года сайт входил в топ 10 по запросу «bitcoin news» в Яндексе и занимал второе место в Google по тому же запросу.



**Cointelegraph.com**

Сайт существует с 2013-ого года. Портал является независимый источником новостей о криптовалюте, блокчейне и децентрализованных приложениях. На сайте есть новости, анализ, экспертные колонки и информация, касающаяся Финтеха, Блокчейна и Биткойна. (<https://www.crunchbase.com/organization/cointelegraph>)

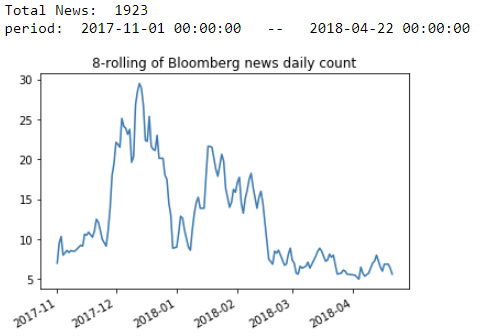
На апрель 2018-ого года сайт входил в топ 10 в Google и в топ 5 в Яндексе по запросу «bitcoin news».



**Bloomberg.com**

Компания была основана в 1981-м году Михаэлом Блумбергом (Michael Bloomberg). Компания должна была стать технологическим поставщиком финансовой аналитики и информации. Сейчас Bloomberg – один из главных глобальных поставщиков круглосуточной информации по финансовым данным и другой финансовой информации, включая текущие и исторические цены на рыночные активы, финансовые отчётности, финансовые новости, аналитику, а также общие новости и новости спорта. Блумберг представляет новости на своей собственной разработанной платформе – терминале – а также на своём ТВ канале (Bloomberg Television), на радио (WBBR), в журналах (Bloomberg Businessweek, Bloomberg Markets, Bloomberg Pursuits) и на сайте (Bloomberg.com). (<https://www.investopedia.com/terms/b/bloomberg.asp>, <https://en.wikipedia.org/wiki/Bloomberg_L.P>.)

В целях анализа использовались только новости с Блумберг, которые можно найти на сайте по запросу «биткойн» (<https://www.bloomberg.com/search?query=bitcoin>). К сожалению, Блумберг не даёт возможности вытаскивать все существующие старые новости в силу технических ограничений портала. Поиск по старым новостям устроен таким образом, что индексируются единичные самые популярные новости и нет возможности выгрузить все существовавшие публикации. В силу этого ограничения для анализа были доступны только новости, начиная с ноября 2017-ого года. Но для каждой новости была известна точная минута публикации, поэтому эти данные можно использовать для анализа связи с минутными данными по цене биткойна.

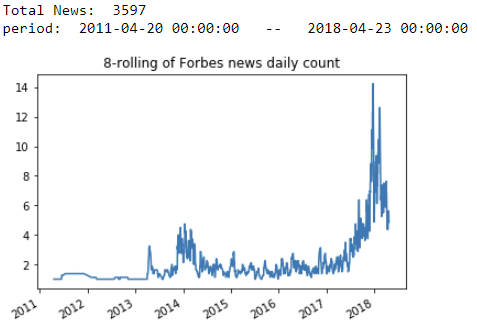


**Forbes.com**

Forbes.com является частью Forbes digital, который в свою очередь является частью Forbes Media LLC. Сайт был основан Дэвидом Чарбаком в 1996-ом году. Так же как и журнал сайт занимается публикацией новостей и статей о долларовых биллионерах, их жизненным увлечениям и об их движимом и недвижимом имуществе. Но на сайте также присутствует много новостей из мира бизнеса и политики. (<https://en.wikipedia.org/wiki/Forbes>)

Forbes.com использует «модель контрибуции» при отборе статей на сайт. То есть контент на сайт создают и публикуют «контрибьюторы». Всем контрибьюторам сайт платит за публикации на основе посещаемости их статей. Всего новости на сайте публикует свыше 2500 человек при этом некоторые из авторов заработали за свою деятельность свыше $100 000 (<https://www.journalism.co.uk/news/the-forbes-contributor-model-technology-feedback-and-incentives/s2/a554255/>).

В целях данной работы использовались новости, которые можно найти на сайте Forbes.com при использовании поискового запроса «bitcoin». Полный список новостей можно посмотреть по ссылке <https://www.forbes.com/search/?q=bitcoin>



Имеются почасовые данные биткойна (BTC) и подневные данные биткойна (BTC) и эфира (ETH)

