|  |
| --- |
| МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ |
| Федеральное государственное автономное образовательное  учреждение высшего образования |
| **НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЯДЕРНЫЙ УНИВЕРСИТЕТ «МИФИ»** |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  **КАФЕДРА ФИНАНСОВОГО МОНИТОРИНГА**  Направление подготовки 10.05.03  «Информационно-аналитические системы безопасности» |



**Пояснительная записка к учебно-исследовательской**

**работе на тему:**

«Выявление подозрительных транзакций в блокчейне»

Студент гр. C15-501…………………………………………/………………………./

Научный руководитель …………….………………………/………………………./

Москва, 2019

СОДЕРЖАНИЕ

ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ 3

ВВЕДЕНИЕ 4

ОСНОВНАЯ ЧАСТЬ 6

1 Изучение принципов функционирования блокчейна 6

2 Формирование гипотезы 7

3 Изучение и использование API блокчейна 8

3.1 Как устроено API? Какие данные можно получать? К каким сайтам обращаться? 8

3.2 Получение данных 8

4 Формирование датасета для кластеризации 9

5 Применение методов снижения размерности. Проведение кластеризации 10

5.1 Теория по кластеризации (k-means, прочие) + tsne (pca) 10

5.2 Результаты кластеризации 10

6 Анализ полученных результатов 11

6.1 Изучение состава кластеров (сколько объектов в каждом, куда попали подозрительные) 11

6.2 Построение графа 11

6.3 Анализ результатов на графе (отследить че куда) 11

ЗАКЛЮЧЕНИЕ 12

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 13

# ОПРЕДЕЛЕНИЯ, ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

# ВВЕДЕНИЕ

Традиционная банковская система развивалась и формировалась на протяжении многих столетий и все это время не имела аналогов. Однако, начиная с 2008 года, когда впервые была описана технология блокчейн [1], наравне с фиатной валютой начала набирать популярность и криптовалюта. Анонсирование блокчейна сделало возможным существование децентрализованных платформ, продемонстрировало, что хранить и передавать ценности надёжно и безопасно можно и без использования банковских структур.

Уникальность блокчейна заключается том, что функционирование и защита системы основаны на криптографических методах. Децентрализованность технологии и высокий уровень безопасности обеспечил необходимое доверие пользователей. Возможность совершать транзакции анонимно также сделало блокчейн привлекательней по сравнению с традиционными банковским системами.

Однако, вышеописанные уникальные характеристики блокчейна помимо очевидных преимуществ несут в себе также и угрозу злоупотребления анонимностью путем совершения транзакций по отмыванию доходов полученных преступным путем, по финансированию наркотрафика, незаконной покупке и продаже оружия, поддельных документов, а также прочей незаконной деятельности. Разумеется, пользователи, осуществляющие подобную деятельность, составляют лишь малую долю среди всех пользователей блокчейна, тем не менее известно достаточное количество случаев, в которых криптовалюта использовалась для финансирования терроризма, краж, наркотрафика, мошенничества и вымогательства. Одним из наиболее известных примеров является закрытие Silk Road [5].

Существование возможности злоупотребления уникальными свойствами криптовалюты подрывает доверие к технологии со стороны финансовых регуляторов, правоохранительных органов и разведывательных служб [2-4]. Для правительства факт использования блокчейна для совершения транзакций, связанных с преступной деятельностью, представляет серьезную опасность. В связи с этим возникает необходимость изучить возможность выявления подозрительных транзакций в блокчейне с целью последующего контроля за кошельками, которые в таких транзакциях участовали.

Целью данной работы является изучение основных принципов функционирования технологии блокчейн, а также применение методов машинного обучения для проверки возможности верхнеуровнего выделения подозрительных транзакций из общей массы с целью их последующего детального рассмотрения.

# ОСНОВНАЯ ЧАСТЬ

## 1 Изучение принципов функционирования блокчейна

Блокчейн – база данных, совместно используемая всеми пользователями, совершающими транзакции с криптовалютой. Эта база данных представлена в виде цепочки блоков, хранящих информацию о всех совершенных операциях. Блоки в цепи нельзя заменить или сфальсифицировать, так как добавление блока верифицируется, а в дальнейшем копия сформированной цепи хранится у каждого участника. Полная копия цепочки блоков содержит все когда-либо совершенные и подтвержденные транзакции. Используя эту информацию, можно точно узнать сумму и время совершения транзакций, адреса отправителей и получателей, а также прочие параметры транзакции независимо от того, как давно она была совершена. Таким образом, несмотря на анонимность адресов, блокчейн абсолютно прозрачен, а история транзакций в публичном блокчейне доступна для изучения и анализа любому заинтересованному лицу.

Каждый блок содержит хэш предыдущего блока. Это имеет эффект создания цепочки блоков от Блока genesis до текущего блока. Каждый блок гарантированно придет после предыдущего блока в хронологическом порядке, потому что хэш предыдущего блока иначе не будет известен. Каждый блок также вычислительно нецелесообразно изменять, как только он был в цепочке на некоторое время, потому что каждый блок после него также должен быть восстановлен. Эти свойства делают транзакции bitcoins необратимыми. Блоком цепи является главным нововведением биткоин.

Честные генераторы строятся только на блоке (путем ссылки на него в блоках, которые они создают), если это последний блок в самой длинной допустимой цепочке. "Длина" рассчитывается как общая совокупная сложность этой цепочки, а не количество блоков, хотя это различие важно только в контексте нескольких потенциальных атак. Цепочка действительна, если все блоки и транзакции в ней действительны, и только если она начинается с блока genesis.

Для любого блока в цепи есть только один путь к блоку genesis. Однако, если исходить из блока генезиса, могут быть вилки. Одноблочные вилки создаются время от времени, когда два блока создаются всего за несколько секунд. Когда это происходит, генерирующие узлы строятся на любом из блоков, которые они получили первыми. Какой бы блок ни был включен в следующий блок, он становится частью основной цепи, потому что эта цепь длиннее. Более серьезные вилки произошли после исправления ошибок, которые требовали обратных несовместимых изменений.

Блоки в более коротких цепях (или недействительных цепях) не используются ни для чего. Когда биткойн-клиент переключается на другую, более длинную цепочку, все действительные транзакции блоков внутри более короткой цепи повторно добавляются в пул транзакций в очереди и будут включены в другой блок. Вознаграждение за блоки на более короткой цепи не будет присутствовать в самой длинной цепи, поэтому они будут практически потеряны, поэтому существует время созревания 100 блоков для поколений.

Эти блоки на более коротких цепях часто называют "сиротскими" блоками. Это связано с тем, что транзакции генерации не имеют родительского блока в самой длинной цепочке, поэтому эти транзакции генерации отображаются как сироты в вызове RPC listtransactions. Несколько пулов неверно истолковали эти сообщения и начали называть свои блоки "сиротами". На самом деле у этих блоков есть родительский блок и даже могут быть дети.

Поскольку блок может ссылаться только на один предыдущий блок, слияние двух разветвленных цепочек невозможно.

Алгоритм блочной цепи можно использовать в нефинансовых целях: см. альтернативная цепочка.

Цепочка блоков транслируется на все узлы сети с использованием протокола flood: см. раздел Загрузка цепочки блоков.

## 2 Формирование гипотезы

## 3 Изучение и использование API блокчейна

### 3.1 Как устроено API? Какие данные можно получать? К каким сайтам обращаться?

### 3.2 Получение данных

## 4 Формирование датасета для кластеризации

## 5 Применение методов снижения размерности. Проведение кластеризации

### 5.1 Теория по кластеризации (k-means, прочие) + tsne (pca)

### 5.2 Результаты кластеризации

## 6 Анализ полученных результатов

### 6.1 Изучение состава кластеров (сколько объектов в каждом, куда попали подозрительные)

### 6.2 Построение графа

### 6.3 Анализ результатов на графе (отследить че куда)

# ЗАКЛЮЧЕНИЕ

# СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

[1] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” 2008.

[5] N. Christin, “Traveling the silk road: A measurement analysis of a large anonymous online marketplace,” in Proceedings of the 22Nd International Conference on World Wide Web, WWW ’13, (New York, NY, USA), pp. 213–224, ACM, 2013.