

Федеральное государственное автономное образовательное
учреждение высшего образования



ПОЛИТЕХ
Санкт-Петербургский
политехнический университет
Петра Великого

Институт компьютерных наук и технологий
Высшая школа интеллектуальных систем и суперкомпьютерных технологий

Отчёт по лабораторному практикуму N3
по дисциплине "Цифровые ресурсы в научных исследованиях"
по теме "Распределенные реестры (блокчейн) -
новые концепции и технологии последних лет"

Выполнил студент гр. 3540901/21501

_____ С.А.Мартынов
(подпись)

Преподаватель

_____ Е.Н.Бендарская
(подпись)

« _____ » _____ 2022 г.

Санкт-Петербург
2022

Оглавление

Задание 1. Поиск узкоспециализированных источников.	4
Задание 2. Представить полный список отобранных источников.	6
2.1. Актуальность поставленной проблемы	6
2.2. Формулировка сути проблемы	6
2.3. Объект и предмет исследований и разработок	7
2.4. Соотношение темы с другими смежными темами и ее особенности, значимость	7
2.5. Исторический аспект появления и развития рассматриваемой проблемы . . .	7
2.6. Современное состояние проблемы	8
2.7. Кто занимается данной проблемой	8
2.8. Основные теоретические и методические положения	8
2.9. Подходы (методы и методологии) к решению проблемы с программной или аппаратной реализацией	8
2.10. Области применения результатов разрешения проблемы	9
2.11. Выбор одного из рассмотренных подходов и поиск необходимых библиотек, платформ и фреймворков для моделирования.	9
Задание 3. Дополнить список цифровых источников.	10

Задание 4. Поиск источников с подходящими БД.	11
Заключение	13

Задание 1. Поиск узкоспециализированных источников.

Задание: Поиск узкоспециализированных источников по теме – сайты и работы отдельных научных коллективов:

- НИИ;
- научные лаборатории;
- научные группы;
- отдельные ученые.

Представить по каждому подпункту по 5-7 наиболее релевантных источников.

Санкт-Петербургский государственный университет: Центр технологий распределенных реестров СПбГУ (<https://dltc.spbu.ru/>). Осуществляет исследования, разрабатывают образовательные программы, оказываются экспертные услуги и создают программные решения с использованием технологий распределенных реестров (блокчейн).

Федеральное государственное автономное образовательное учреждение высшего образования «Пермский государственный национальный исследовательский университет» (ПГНИУ): Лаборатория криптоэкономики и блокчейн-систем (<http://www.psu.ru/fakultety/ekonomicheskij-fakultet/laboratorii/laboratoriya-kriptoekonomiki-i-blokchejn-sistem>). Исследования направлены на создание математических моделей и практическое применение блокчейн-систем и смарт-контрактов для решения прикладных задач, таких как межбанковские и биржевые расчеты, финансовые расчеты в международных холдингах и

группах, открытое электронное голосование, подтверждение авторских прав на цифровой контент и электронный нотариат.

Федеральное государственное образовательное бюджетное учреждение высшего образования "Финансовый университет при Правительстве Российской Федерации": Блокчейн-лаборатория (<http://www.fa.ru/org/science/irce/blockchainlab/Pages/Home.aspx>). Финансовый университет выиграл и выполнил тендер Государственной думы РФ на проведение экспертно-аналитического исследования по теме, связанной с технологией блокчейна и криптовалютам. Проводит Научно-исследовательские и опытно-конструкторские работы, связанные с внедрением технологии блокчейн в государственном и корпоративном секторах.

Калифорнийский университет в Беркли (The University of California, Berkeley): Harness the power of blockchain and cryptocurrencies (<https://www.edx.org/professional-certificate/uc-berkeleyx-blockchain-fundamentals>)

MIT (Массачусетский технологический институт): Blockchain Technologies: Business Innovation and Application (<https://executive.mit.edu/course/blockchain-technologies/a056g00000URaa.html>)

Stanford School of Engineering: Cryptocurrencies and Blockchain Technologies (<https://online.stanford.edu/courses/xcs251-cryptocurrencies-and-blockchain-technologies>).

- Корхов Владимир Владиславович (Доктор экономических наук, Профессор кафедры управления рисками и страхования СПбГУ).
- Сергей Владимирович Ивлиев (Кандидат экономических наук, Руководитель Лаборатории криптоэкономики и блокчейн-систем ПГНИУ)
- Виталий Бутерин (Создатель Ethereum)
- Горгадзе Владимир Вячеславович (Ph.D. теоретической физики, Калифорнийский университет в Беркли)
- Баргер Артем Иосифович (Научный сотрудник IBM, работающий в группе Cloud Foundation в Хайфской исследовательской лаборатории)

Задание 2. Представить полный список отобранных источников.

Задание: Представить полный список отобранных источников (использовать результаты ЛР1 И ЛР2 и п.1 данной ЛР3), распределив их по пунктам

Списки находятся в файлах library.eng.bib и library.rus.bib.

При классификации материалов, я максимально старался избегать повторов. Но это оказалось довольно затруднительной задачей, т.к. в одной работе могут встречаться и исторические аспекты и современное состояние.

2.1. Актуальность поставленной проблемы

- Анализ блокчейн-технологии: основы архитектуры, примеры использования, перспективы развития, проблемы и недостатки
- Технологии, реализующие принцип распределенного реестра, и возможности их применения в информационных системах

2.2. Формулировка сути проблемы

- The essence of cryptocurrencies: Descriptive and comparative analysis
- Технологические принципы функционирования блокчейн
- Анализ блокчейн-технологии: основы архитектуры, примеры использования, перспективы развития, проблемы и недостатки

2.3. Объект и предмет исследований и разработок

- Технологические принципы функционирования блокчейн
- Анализ блокчейн-технологии: основы архитектуры, примеры использования, перспективы развития, проблемы и недостатки

2.4. Соотношение темы с другими смежными темами и ее особенности, значимость

- К вопросу о выборе платформы распределенного реестра при проектировании информационных систем финансового сектора экономики
- Концепция распределенного реестра на основе графов в системах распределенной энергетики
- Применение системы распределенного реестра для повышения оперативного учета и контроля технического состояния элементов распределенной информационной системы
- Разработка элементов распределенной библиотеки на базе распределенного реестра для систем мониторинга и диагностики
- Разработка модели распределенного хранилища данных на базе технологии распределенного реестра для систем мониторинга

2.5. Исторический аспект появления и развития рассматриваемой проблемы

- Технологические принципы функционирования блокчейн
- Анализ блокчейн-технологии: основы архитектуры, примеры использования, перспективы развития, проблемы и недостатки

2.6. Современное состояние проблемы

- Анализ практической реализации технологии распределённого реестра

2.7. Кто занимается данной проблемой

Изучением, развитием и практическим применением блокчейна занимается огромное количество как коммерческих компаний (это было видно при просмотре патентов, выданных в Китае), так и научных организаций (в задании 1 текущей работы несколько примеров из ведущих ВУЗов), и даже независимые сообщества разработчиков.

2.8. Основные теоретические и методические положения

- Методика обеспечения своевременности и полноты обмена информационными ресурсами в корпоративных сетях с распределённым реестром
- Анализ открытого блокчейна в рамках комплексной классификации технологии распределённого реестра
- Анализ практической реализации технологии распределённого реестра
- Proof-of-randomness protocol for blockchain consensus: the white paper version 1.0

2.9. Подходы (методы и методологии) к решению проблемы с программной или аппаратной реализацией

- Decision Making using the Blockchain Proof of Authority Consensus
- Proof of Learning (PoLe): Empowering neural network training with consensus building on blockchains
- Consensus mechanism for software-defined blockchain in internet of things
- Weighted Byzantine Fault Tolerance consensus algorithm for enhancing consortium blockchain efficiency and security

2.10. Области применения результатов разрешения проблемы

- Практическое применение технологий распределенного реестра
- Blockchain Platform for Industrial Internet of Things
- Blockchain-based Database to Ensure Data Integrity in Cloud Computing Environments

2.11. Выбор одного из рассмотренных подходов и поиск необходимых библиотек, платформ и фреймворков для моделирования.

В качестве интересного практического примера можно взять блокчейн Ethereum: этой осенью они осуществили переход от одного алгоритма консенсуса к другому.

Задание 3. Дополнить список цифровых ИСТОЧНИКОВ.

Задание: Дополнить список цифровых источников ЛР1, ЛР2 и ЛР3 своими 2-3 дополнительными ресурсами. При необходимости провести в них поиск источников.

Хорошим примером перехода от одной модели консенсуса к другой (от proof-of-work к proof-of-stake) является сеть Ethereum. Дополнительным цифровым источником будет являться исходный код ноды блокчейна, размещённый в хранилище кода GitHub:
<https://github.com/ethereum/go-ethereum>.

Кроме того, имеется отдельный репозиторий, который содержит предложения пользователей по улучшению продукта. Эти приложения формально не являются научными документами, однако обладают некоторыми научными признаками:

- Они оформлены по устоявшемуся шаблону
- Они проходят публичное обсуждение
- Они подкреплены ссылками на другие значимые источники

Ссылка на репозиторий: <https://github.com/ethereum/EIPs>

Задание 4. Поиск источников с подходящими БД.

Задание: Предварительный поиск источников с подходящими БД с входными примерами для тестирования и проверки работоспособности выбранного подхода (п.2.11) по теме аналитического отчета, включая проверку наличия необходимых данных в существующих и подходящих по смыслу для исследования подхода:

1. Для задач кластеризации

https://www.uni-marburg.de/fb12/arbeitsgruppen/datenbionik/data?language_sync=1 (с подробным описанием)

https://www.uni-marburg.de/fb12/arbeitsgruppen/datenbionik/Daten?language_sync=1 (удобные архивы)

<http://cs.uef.fi/sipu/datasets/>

<http://ana.cachopo.org/datasets-for-single-label-text-categorization> (категоризация текстов)

2. Для задач классификации

<https://sci2s.ugr.es/keel/category.php?cat=clas&order=ins#sub2>

3. Базы тестовых и реальных данных разного происхождения (медицинские, геологические и т.п.)

<https://archive.ics.uci.edu/ml/index.php>

4. Для разных задач (нужно выбрать тип входных данных) - классификация (тексты, изображения), сегментация (изображения), аннотирование (документов)...

<https://dataturks.com/projects/trending>

5. Сборники Top Sources For Machine Learning Datasets и 10 Standard Datasets for Practicing Applied Machine Learning

<https://towardsdatascience.com/top-sources-for-machine-learning-datasets-bb6d0dc>
<https://machinelearningmastery.com/standard-machine-learning-datasets/>

6. <https://www.openml.org/>

7. <https://data.world/datasets/classification>

Дополнить 2-3 дополнительными ресурсами с базами тестовых примеров (например, из соревнований по машинному обучению, хакатонов, и т.д.)

Если мы говорим про Ethereum, то достаточно запустить ноду блокчейна, и в процессе синхронизации будут получены все данные, которые можно изучать. В то же время, объём этих данных просто огромен, по этой причине специалисты подготовили облегченные датасеты, которые удобно использовать в экспериментах.

Блокчейн Ethereum (март 2019) Полные исторические данные о блокчейне Ethereum в реальном времени (BigQuery).

Датасет содержит 8 760 000 строк записей и подготовлен для экспериментов на платформе Kaggle.

<https://www.kaggle.com/datasets/bigquery/ethereum-blockchain>

Публичный датасет для изучения смарт-контрактов (август 2018)

Датасет был подготовлен в рамках демонстрации аналитических возможностей платформы гугл по анализу данных.

https://bigquery.cloud.google.com/dataset/bigquery-public-data:ethereum_blockchain

Ethereum (2015)

Маленький (килобайты) по объёму датасет с примерами торговли в течение дня.

<https://datahub.io/cryptocurrency/ethereum#data>

Также дополнительным источником датасетов могут служить платформы:

- <https://data.world/datasets/ethereum>
- <https://datahub.io/>

Заключение

Было обнаружено достаточно большое количество узкоспециализированных источников – лабораторий и кафедр при ведущих мировых ВУЗах. Таким образом, можно выделить следующие группы, занимающиеся изучением Блокчейн:

- Научное сообщество
- Коммерческие компании
- Независимые исследователи на GitHub

Кроме хранения исходного кода, GitHub выступает как коммуникационная платформа, где независимые исследователи представляют свои предложения для обсуждения сообществом.

При классификации собранных материалов наибольшую сложность вызвало определение подходящей категории для какой-либо публикации. Зачастую, одна и та же работа могла удовлетворять требованиям различных категорий, т.к. покрывала поставленные там вопросы.

В качестве одного из подходов к реализации блокчейна и используемых алгоритмов консенсуса мы упомянули Ethereum. Этой осенью проект перешёл с proof-of-work на proof-of-stake подход.

Для анализа данных в Ethereum можно установить ноду, и получить все операции которые когда-либо происходили в этой сети с момента её запуска. Однако сообщество заранее подготовило минималистичные датасеты, на которых можно проверить свои модели до запуска на больших данных.