Искусственные общества 2013-2024

ISSN 2079-8784

URL - http://ras.jes.su

Выпуск 3 Том 16. 2021

Оценка масштабирования системы отслеживания социальных контактов на основе блокчейн

Шмелев Илья Александрович
старший научный сотрудник, ГАУГН
ФИЦ ИУ РАН
Российская Федерация, Москва

Тарханов Иван Александрович

Аннотация

Во время пандемии одной из важнейших задачи является отслеживание социальных контактов с заболевшими. В данной работе классифицируются проекты, посвященные отслеживанию этих контактов. Делается вывод, что гибридная архитектура такого решения на основе эксклюзивного блокчейн будет иметь ряд преимуществ и описана концептуальная модель такой системы. Однако анализ существующих проектов, использующих технологию блокчейн, показал, что основная их проблема — это непроработанный вопрос масштабирования, что в условиях создания глобальной цифровой инфраструктуры общества становится ключевым вопросом. Далее проводится оценка масштабирования концептуальной модели системы на основе открытых источников о московском метро и подтверждаются основные выводы о выбранных архитектурных решениях.

Ключевые слова: социальные контакты, COVID-19, блокчейн, фиксация контакта, масштабирование

Дата публикации: 16.09.2021

Ссылка для цитирования:

Шмелев И. А. , Тарханов И. А. Оценка масштабирования системы отслеживания социальных контактов на основе блокчейн // Искусственные общества – 2021. – Том 16. – Выпуск 3 [Электронный ресурс]. URL: https://artsoc.jes.su/S207751800015809-8-1 (дата обращения: 23.04.2024). DOI: 10.18254/S207751800015809-8

1

Введение

Идея отслеживания социальных контактов давно привлекает внимание исследователей по всему миру и особенно актуальна в эпоху пандемии. В настоящий момент существует ряд различных технологических решений данной проблемы разной степени проработанности [9]. Большинство из этих решений пытаются решить проблемы распространения тех или иных вирусов [6]. Ряд подходов подразумевают высокую степень участия человека (например, при подтверждении факта посещения заведения), другие включают современные технологические и цифровые инструменты, использующие Internet of Things (IOT) [13], Radio Frequency Identification (RFID) [6], Bluetooth Low Energy (BLE) [6], [12], Distributed Ledger Technology (DLT) [13] и др. Все они имеют разную архитектуру, но состоят из нескольких типовых компонентов:

- Инструмент для идентификации местоположения субъекта.
- Хранилище информации о местоположении субъектов.
- Алгоритм определения социальных контактов между субъектами.
- Процедура публикации информации о заражении субъекта.
- Уведомление субъектов о риске инфицирования.
- Комплекс мер для обеспечения безопасности персональных данных.

2 Целью данной статьи является разработка концептуальной модели построенной на эксклюзивном блокчейне (блокчейн, в котором обработка транзакций осуществляется определенным списком идентифицированных участников) [19], а также проработка вопроса масштабирования таких систем и использование технологии блокчейн как дополнительного и независимого хранилища информации о контактах с заболевшими.

3

Анализ архитектуры проектов

Согласно обзору [18] *инструментов идентификации местоположения* субъекта, выделяют два основных аспекта – определение геолокации субъекта и определение близости между субъектами. Определение *геолокации* может выполняться с помощью:

- Модуля глобальной навигационной системы (GPS, Глонасс и т.д.) [6], [12].
- Вышек сотовой связи [18].
- Предварительно заданной информации о местоположении модуля фиксации: таких как BLE, RFID модулей, Wi-Fi или QR-кодов [8], [18].

4 В то же время определение близости между субъектами производится с помощью BLE, RFID [9]. Данная информация может храниться как на удаленном сервере, так и в локальной памяти устройств, в зависимости от подхода к обеспечению безопасности персональных данных субъектов.

5 Публикация информации об инфицировании субъекта в большинстве случаев производится должностным лицом, имеющим на это право после положительного теста на наличие инфекции. После этого непосредственно мобильное устройство пользователя или центральный сервис системы проводит риск-анализ на основе собранных данных и уведомляет контактировавших.

6 Большинство анализируемых решений одним из основных приоритетов декларируют вопрос обеспечения безопасности персональных данных людей [6], [8]. Как минимум, реализуют шифрование передаваемых персональных данных. Как максимум, производят генерацию уникальных анонимных ключей на устройствах пользователей, которые знают лишь контактирующие стороны, но не могут идентифицировать друг друга [8], [16]. Здесь мы рассматриваем только решения, использующие пользовательские устройства для идентификации контактов между людьми, т. к. они наиболее безопасны с точки зрения возможностей идентификации, контактирующих лиц. В работе [18] описаны различные виды решений, построенные как на централизованной, так и на децентрализованной архитектуре. Классификация таких решений предложена в [6], где авторы разделили существующие технологические решения на 3 основные класса: централизованные, децентрализованные и гибридные.

7 **Централизованные** решения полагаются на центральный сервис, который инкапсулирует в себе процедуры по созданию идентификаторов устройств пользователей, отслеживанию контактов и уведомлению всех сторон о рисках заражения. Примером такой архитектуры являются системы, построенные на базе BlueTrace протокола [8]. Основной их проблемой является тот факт, что необходимо полагаться на 3-ю сторону, которая может быть умышленно или неумышленно скомпрометирована, что может поставить под угрозу безопасность персональных данных или корректность функционирования всей системы в целом. Данный аспект является критически важным для социально-значимых сервисов коим и является рассматриваемая система. Примеры централизованных систем - BlueTrace [8], ROBERT [17]. Мобильные приложения, запущенные на основе централизованных алгоритмов - TraceTogether, CovidSafe (BlueTrace), StopCovid (ROBERT).

8 **Децентрализованные** решения также полагаются на центральный сервис, но только как на посредника для передачи вспомогательных данных необходимых для риск-анализа, проводимого на самих устройствах. Примером может служить алгоритм PACT [16], который подразумевает, что все устройства обмениваются периодически генерируемыми кодами и сохраняют их для последующего анализа рисков инфицирования. Для генерации случайных кодов используются seed-данные, как основа для вычислений псевдослучайных функций. После того, как соответствующий государственный орган подтвердит, что участник был инфицирован, участник получает право загрузить seed-данные на центральный сервис, откуда остальные участники могут их получить и проверить риск заражения вирусом. Примеры децентрализованных систем - PACT (EAST-COAST), PACT (WEST-COAST), TCN, DP-T3, TP-T3 unlinkable, Pronto-C2.

9 **Гибридные** решения, как понятно из названия, это сочетание централизованных и децентрализованных решений, где, например, создание идентификаторов пользовательских устройств и обмен ими происходит с помощью клиентских приложений, а уведомления о заражении и анализ взаимодействия людей на предмет риска инфицирования производится центральным сервисом. Примеры гибридных систем - Desire [10], [17], EpiOne, ConTra Corona. Примеры приложений, использующих децентрализованные алгоритмы: SwissCovid (DP-3T), CovidWatch (TCN), CovidSafe (PACT WEST-COAST). Подробности см. [6], [9].

10 Очевидным плюсом гибридных решений является то, что они позволяют обеспечить анонимность пользователей (за счет того, что те сами управляют процессом раскрытия своих данных) и одновременно агрегировать информацию о контактах, заболевших **на общем для всех участников сервисе**. Такой подход предоставляет возможность организовать реестр больших данных, который в дальнейшем можно использовать для анализа поведения общества во время пандемий, нахождение очагов распространения заболевания, прогнозирование их появления, позволяя принимать точечные меры для борьбы с заболеванием, уменьшая ущерб, наносимый экономике.

11 Благодаря тому, что гибридная архитектура позволяет гибко совместить в себе разные аспекты централизованных и децентрализованных систем, именно её целесообразно взять за основу для разработки концептуальной модели. Несмотря на то, что анонимность пользователей обеспечивается в гибридных архитектурах, на центральный сервис налагается очень высокая ответственность по сохранению целостности информации и защите от вмешательства в функционирование. Также полезные для общества данные находятся под контролем одной стороны и нет никаких гарантий, что эти данные будут открыты для исследователей в исходном виде и не подвергнуты тем или иным манипуляциям.

12

Хранение данных в блокчейн

Для обеспечения неизменности и прозрачности статистических данных в качестве хранилища целесообразно использовать распределенный реестр (блокчейн, Distributed Ledger Technology, далее DLT) в том или ином виде. В работе [13] предложена концепция системы, использующей DLT, для построения саморегулирующейся открытой децентрализованной системы для отслеживания социальных контактов между людьми в целях контроля распространения заболеваний. Однако в ней мало места уделено проблеме масштабирования такой сети и адаптации к реальной нагрузке, что является существенной проблемой.

Логично предположить, что такие системы предназначены для работы с большим количеством данных, а проблема масштабирования DLT является актуальной [21].

13 В роли хранилища данных для рассматриваемой системы может выступать эксклюзивный блокчейн [19], контролируемый консорциумом лиц ответственными за платформу. У данного решения есть несколько преимуществ:

14
1. Эксклюзивный блокчейн проще поддерживать, вертикально масштабировать и в том числе существует возможность обеспечить SLA, что чрезвычайно важно для такого рода сервисов.
2. Неизменность данных записанных в блокчейн в сочетании с правильной балансировкой интересов сторон, поддерживающих сеть, позволят сделать прозрачную для всех пользователей и интересантов структуру, где можно строго идентифицировать лиц кто вносит информацию в сеть.
3. Контролируемая приватность данных – в зависимости от потребностей регуляторов, такой блокчейн можно сделать публичным для общества, а можно сделать приватным и предоставлять информацию через определенный интерфейсы взаимодействия с сервисом.
4. Существуют прозрачные и верифицируемые правила для всех участников сети по которым работает система.

15 Данное решение реализуемо на базе блокчейн фреймворка, который поддерживает смарт-контракты, DPoS (Delegated Proof-of-Stake) или PoA (Proof-of-Authority) алгоритмы консенсуса и настройку прав для эксклюзивного блокчейна. Например, ими могут быть Substrat, Exonum, Cosmos SDK и др. Далее сформирован Proof-of-Concept системы для оценки её возможностей при росте нагрузок.

16

Структура программного комплекса

Концептуальная схема гибридного программного решения изображена на рисунке 1.

17

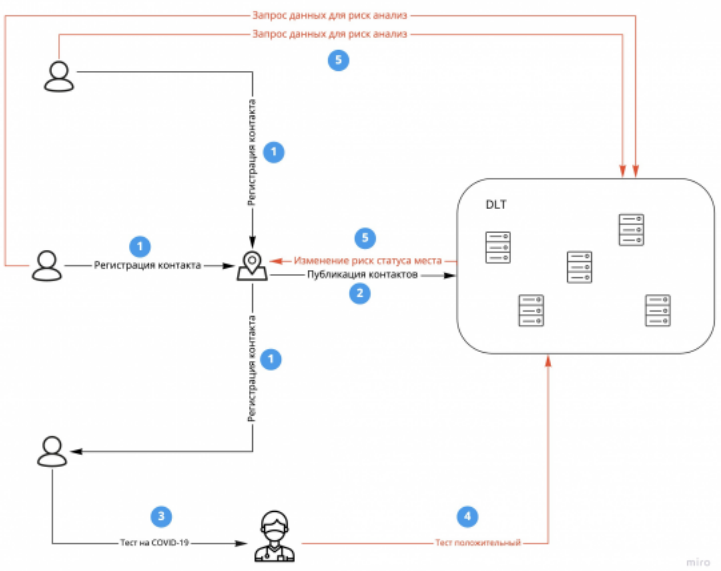

Рис. 1. концептуальная схема программного решения

18 Предложенная здесь концептуальная модель во многом базируется на [13], где рассматривается сеть с IOT устройствами, использующими BLE технологию. Эти устройства выступают в роли свидетелей физического присутствия пользователя в конкретной локации. Процесс состоит из 5 основных этапов:

- регистрация социального контакта в определенной локации, в которой установлено IOT устройство;
- публикация информации этим устройством в DLT;
- тестирование человека на COVID-19 и выявление факта инфицирования;
- публикация информации об инфицировании в сеть;
- уведомление людей, посещавших локацию и изменение риск-статуса данной локации.

19 В предложенной концептуальной схеме стоит обратить внимание на следующий ряд ключевых компонентов:

1. алгоритм фиксации контакта;
2. хранилище данных;
3. уведомления об инфицировании и анализ риска заражения;
4. процесс публикации информации об инфицированном в сеть.

20 Компоненты из п.1 и п.3 выбраны на основе [6], [18] и адаптированы под использование с выбранным хранилищем данных. Процесс из п.4 не является чем-то специфичным, но опирается на решения, принятые в предыдущих пунктах. Псевдокод смарт-контрактов, представляющих реализацию данной системы, представлен в приложениях А, Б, В.

21

Оценка нагрузки на сеть

Оценить порядок цифр и примерные объемы данных, которые будут циркулировать в системе, можно на основе открытой статистики о московском метро. Для этого определим набор вводных данных и рассчитаем объем пиковых нагрузок на разные уровни системы - IOT локации, DLT в целом.

Вводные данные:

1. Максимальный зафиксированный пассажиропоток (*S*) в московском метро за сутки по известным открытым данным [5], историческое пиковое значение от 2014 года: 9 715 635 пассажиров –Smax
2. Количество станций метро [5]: 278 (01.04.2021)
3. Допустим, что DLT будет использовать алгоритм ecdsa для создания криптографической подписи с эллиптической кривой (secp256k1), как один из наиболее известных в реализации DLT [20], [15]
4. Регистрация контактов с IOT устройством будет создавать основную нагрузку на сеть и в качестве параметров там выступает: organizationPublicKey (64 байта), userPublicKey (64 байта), timeframe (4 байта), userSignature (64 байта). Данные параметры — это аргументы метода смарт-контракта из приложения Б: registerContact.
5. Человек за время поездки в среднем проезжает 7 станций –k. Данная приблизительная оценка получена из следующих соображений на основе информации из [2], [5]. Самая длинная ветка метро: арбатско-покровская 45,1 км [5]. Средняя продолжительность поездки пассажира 14,5 км [5]. Количество станций на арбатско-покровской линии: 22 [2].

Исходя из вводных данных можно определить базовые параметры для расчетов:

1. Размер транзакции: минимальный размер: 64 * 3 + 4 = 196 байта. Реальный размер зависит от имплементации протокола DLT и по этой причине округлим в большую сторону до 512 байт или 0,5 кб – d_T
2. Средний пассажиропоток на станцию метро: 34 948 пасс. или округлив 35 тыс. пасс – s_{avg}
3. Если исходить из оценок, опубликованных в материалах по московскому метрополитену [3] и петербургскому метрополитену [4], то можно обнаружить, что существует ярко выраженные часы пик в метро: утром 08:00 - 09:00 (18.6% дневного трафика), вечером 18:00 - 19:00 (19.2% дневного трафика). Для упрощения расчетов положим, что на каждый из этих часов приходится по 20% дневного трафика – p .

Оценка пиковой нагрузки на IOT устройство

Чтобы оценить пиковую нагрузку на одной станции метро, возьмем для ориентира наиболее загруженную станцию и обозначим ее пассажиропоток, как:

$s_{max} = s_{avg} * C$, где C – коэффициент отклонение s_{max} от среднесуточного пассажиропотока.

Следовательно, можно оценить количество запросов в час при пиковой нагрузке следующим образом:

$lh_{max} = s_{max} * p = 35\,000 * 4 * 0{,}2 = 28\,000$

Количество запросов при пиковой нагрузке в секунду:

$ls_{max} = lh_{max} / 3600 = 7.7 \sim 8$

Оценка пиковой нагрузки на DLT

Оценим пассажиропоток в час пик:

$Lh_{max} = S_{max} * p = 9\,715\,635 * 0{,}2 = 1\,943\,127$

Количество запросов в секунду:

$Ls_{max} = Lh_{max} / 3600 = 539.45 \sim 540$

Учитывая, что запросы будут отправляться при посещении пассажиром каждой станции, тогда: $Ws_{max} = Ls_{max} * k = 540 * 7 = 3780$

Оценка роста размера DLT

Оценим объем данных публикуемых в DLT за сутки:

$D_d = S_{max} * d_T * k = 9\,715\,635 * 0{,}5 * 7 = 34\,004\,722{,}5$ кб $\sim 32{,}4$ гб

Публикуемый объем данных можно сократить за счет введения механики сэмплирования данных. Учитывая, что входящий поток данных на IOT устройство является стохастическим, тогда фильтрация потока путем случайного выбора элементов для обеспечения необходимого процента k , не должно повлиять на репрезентативность конечной выборки при правильном выборе числа k . Данное число может конфигурироваться на разных уровнях:

1. единое для всех;
2. определение на уровне конкретной географической точки;
3. определение на уровне конкретной географической точки, таймфрейма и дня недели для более точной обработки данных за день.

Наиболее точным выглядит решение под п.3, потому что позволяет адаптивно менять число k в зависимости от таймфрейма и не позволять данным, поступающим в час пик негативно влиять на репрезентативность выборки.

Базовой имплементацией такого решения может быть таблица таймфреймов за неделю, где каждому таймфрейму соответствует количество людей, прошедших через данную точку на прошедшей неделе в этот день $V_d(n)$ в таймфрейм n . Затем в зависимости от задач оптимизации объема данных выбрать максимальное количество транзакций допустимые для публикации в сутки ld_{max} для каждой точки. Исходя из этого можно рассчитать необходимый коэффициент:

44 $k(n) = min(1, \frac{ld_{max}}{V_d(n) * N})$ где N – количество таймфреймов в сутках.

45 Это наглядная модель, которая демонстрирует подход к решению проблемы, данный вопрос требует дальнейшего исследования и развития с учетом факторов, которые могут быть выявлены при эксплуатации.

46

Масштабирование

В контексте обсуждения проблем масштабирования систем на базе DLT выделяют 3 слоя [21], на которых можно решать проблему масштабирования решений на базе блокчейна:

- слой 0 – методики связанные со скоростью распространения данных по DLT сети;
- слой 1 – методики, связанные с конфигурациями: размера блока, алгоритмов консенсуса, эксперименты со структурой DLT; (шардирование, направленный ацикличный граф);
- слой 2 – методики, связанные с интеграцией внешних источников: платежные каналы, вычисления вне DLT, интеграция других DLT.

47 Так как слой 0 полностью зависит от конкретной реализации DLT, то здесь рассмотрим масштабирование за счет методик слоя 1 и слоя 2. В связи с тем, что подобная сеть требует регулирования государством, то мы рассматриваем эксклюзивную модель DLT. Исходя из этой установки, сеть будут поддерживать гос. организации и аккредитованные интересанты. Соответственно это определяет элементы *слоя 1*, которые будут конфигурироваться:

- размер блока;
- скорость создания блоков;
- алгоритм консенсуса, который определяет нижнюю границу скорости создания и финализации блоков.

48 Касательно *слоя 2*, в связи с тем, что устройства публикующие данные в DLT имеют четкую привязанность к геолокации и определенным юрисдикциям (города, регионы, страны, то разумным выглядит применения подхода, подразумевающего интеграцию различных DLT в один главный. Наша модель подразумевает создание своего DLT для каждого объекта мониторинга, которыми могут быть населенный пункт или совокупность населенных пунктов, объединенных по территориальной принадлежности и юрисдикции. Соответственно узлы, поддерживающие DLT сеть будут являться представителям конкретных юрисдикций и населенных пунктов, что позволит распределить нагрузку на сеть и одновременно повысить устойчивость сети за счет такого рода децентрализации. Затем применяется подход анкоринга [7] – периодически “слепки” данных из этих этого DLT будут публиковаться в главный DLT, которым может быть известный независимый публичный блокчейн (такой как Bitcoin, Ethereum и др.) или же свой DLT поддерживаемый представителями государства и общества. Подобный подход позволяют осуществить платформы, описанные в работе [11] такие как Polkadot, Cosmos, Multichain.

49 **Масштабирование на слое 2**

50 Как было продемонстрировано в разделе 4, скорость роста DLT является довольно значительной и нетрудно рассчитать, когда он достигнет размеров около 1 Тб данных и более. Потому в нашей работе предлагается композитная архитектура DLT, а именно каждая группа управляющих субъектов будет иметь 2 DLT:

1. Краткосрочный DLT с информацией о посещениях локации, данные в которой будут урезаться со временем ввиду потери актуальности. Имеющий 30-дневный срок хранения данных.
2. Долгосрочный DLT, в который будет отправляться информация от сотрудников мед. организации для подтверждения факта посещения инфицированным определенной локации.
3. Раз в период времениTr(который определяется используемой технологией DLT), будет производиться анкоринг долгосрочного DLT с главным DLT.

51 Данное решение подразумевает взаимодействие между двумя DLT, т. е. узлы одного DLT должны быть способны обратиться к другому DLT для получения информации из хранящегося в нем смарт-контракта. Так в приведенном в приложении псевдокоде подразумевается, что *SicknessRegistrar* находится в краткосрочной цепочке блоков, а *SicknessController* и *AccessController* в долгосрочной цепочке блоков.

52 Для оптимизации используемой смарт-контрактами памяти раз в N таймфреймов будет запускаться метод *prune* в *SicknessRegistrar,* который будет стирать более неактуальные данные.

53 **Масштабирование на слое 1**

54 Исходя из описанной структуры на слое 2, задача масштабирования слоя 1 наиболее актуальна для краткосрочного DLT. Для DLT долгосрочного хранения рост объема данных не предполагается быстрым, поэтому задача его масштабирования на слое 1 не первостепенна.

55 Долгосрочный DLT предполагается масштабировать за счёт сокращения хранимого объема данных в рамках узлов участников и введения в эксплуатацию архивных узлов [14] (которые будут хранить полный набор данных из DLT), как только объем хранимых данных на узле превысит допустимое значение. Данное значение конфигурируется операторами DLT и зависит от регуляторных требований и допустимых возможностей оборудования.

56 Для краткосрочного DLT ключевыми параметрами для оптимизации являются:

1. Размер транзакции – базовая единица данных, которая будет публиковаться.
2. Скорость выполнения транзакции – потому что все транзакции в блоке должны быть исполнены в ожидаемый период времени.
3. Размер блока определяет максимальное количество транзакций, которые может взять производитель блоков из пула транзакций;

4. Размер пула транзакций – определяет временной буфер для публикации транзакций, которые оказались сверх ожидаемой пиковой нагрузки. Так, например, если определить скорость обработки транзакции в секунду менее, чем нагрузка в час пик, то в пуле транзакций будут накапливаться дополнительные транзакции, которые могут быть обработаны позже. В нашем случае система не требует быстрого отклика, однако данные в пуле транзакций теряют свойства, которые обеспечивает DLT и подвержены риску утраты данных.

Сформируем оценку количества транзакций, которые могут быть помещены в блок в худшем случае:

$$T_c = min((B_{size} - BE_{size}) / T_{size}, BT_{time} / T_{time}),$$

где B_{size} – размер блока, BE_{size} – размер блока без транзакций, T_{size} – размер транзакции, BT_{time} – время, отводимое на выполнение всех транзакций в блоке T_{time} – максимальное время отводимое под выполнение 1 транзакции.

Затем получим оценку среднего количества транзакций в секунду:

$$T_{sec} = B_{time} / T_c ,$$

где B_{time} – время создания блока в секундах, T_c – количество транзакций в блоке.

Следовательно, можно построить оценку необходимого размера пула транзакций:

$$P_{size} = Pmin_{size} + max(0, (Ws_{max} - T_{sec}) * t_{peak} * T_{size}),$$

где $Pmin_{size}$ – минимальный размер пула, который удовлетворяет условию $Pmin_{size} \geq B_{size} - BE_{size}$, Ws_{max} – пиковое количество транзакций в секунду, T_{sec} – оценочное количество транзакций в секунду, t_{peak} – продолжительность допустимого пикового периода, T_{size} – размер транзакции.

Выводы и результаты

Проведенные анализ позволил систематизировать информацию о подходах в реализации систем фиксации контактов между людьми и особенностях реализации ее компонентов. Показано, что гибридная архитектура в сочетании с использованием подхода анкоринга и разделением DLT на две части для долгосрочного и краткосрочного хранения способно решить проблему масштабирования DLT при имплементации даже на такой глобальный инфраструктурный объект как московское метро [1]. Вычисленные оценки показывают, что даже для пиковых нагрузок достаточно использовать относительно недорогие узлы DLT для краткосрочного хранения.

1. tpeak– время, которое DLT может работать при нагрузках, превышающих его пропускную способность. Этот параметр необходим для определения размера пула транзакций и какой объем данных может оставаться вне DLT и в течение какого промежутка времени.
2. ldmax– показатель, определяющий допустимое количество данных, которое может генерировать точка фиксации социальных контактов в сутки. Определяет ограничение на объем данных для публикации при пиковых нагрузках.
3. t– время жизни краткосрочного DLT, определяет требования по объему памяти для всех узлов.

Остальные параметры для расчетов зависят от конкретных имплементаций сети и требуют анализа с учетом данных, полученных в результате эксперимента.

Библиография:

1. А.В. Соловьев, И.А. Тарханов "Электронные документы и задача обеспечения сохранности при обмене данными в цифровой экономике", // Труды Института Системного Анализа РАН, Москва, 2018, Том 68, № 1

2. Арбатско-Покровская линия // URL: https://ru.wikipedia.org/wiki/%D0%90%D1%80%D0%B1%D0%B0%D1%82%D1%81%D0%BA%D0%BE-%D0%9F%D0%BE%D0%BA%D1%80%D0%BE%D0%B2%D1%81%D0%BA%D0%B0%D1%8F_%D0%BB%D0%B8%D0%BD%D0 (дата обращения: 03.06.2021). – материал взят с сайта “Wikipedia” https://ru.wikipedia.org/

3. Подземный поиск: запросы из московского метро // URL: https://yandex.ru/company/researches/2017/metro

4. Статистика. Пассажиропоток в метро 2016 г. // URL: https://www.metro-spb.ru/statisticheskie-dannye/2016/

5. Метрополитен в цифрах // URL: https://mosmetro.ru/press/digits/ (дата обращения: 25.05.2021).

6. Ahmed N. et al. A survey of COVID-19 contact tracing apps //IEEE access. – 2020. – Т. 8. – С. 134577-134601.

7. Anchoring Service // URL: https://exonum.com/doc/version/latest/advanced/bitcoin-anchoring/ (дата обращения: 03.06.2021).

8. Bay J. et al. BlueTrace: A privacy-preserving protocol for community-driven contact tracing across borders // Government Technology Agency-Singapore, Tech. Rep. – 2020.

9. Braithwaite I. et al. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19 //The Lancet Digital Health. – 2020.

10. Castelluccia C. et al. DESIRE: A Third Way for a European Exposure Notification System Leveraging the best of centralized and decentralized systems //arXiv preprint arXiv:2008.01621. – 2020.

11. Kan L. et al. A multiple blockchains architecture on inter-blockchain communication //2018 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C). – IEEE, 2018. – С. 139-145.

12. Kleinman R. A., Merkel C. Digital contact tracing for COVID-19 //CMAJ. – 2020. – Т. 192. – №. 24. – С. E653-E656.

13. Lv W. et al. Decentralized blockchain for privacy-preserving large-scale contact tracing //arXiv preprint arXiv:2007.00894. – 2020

14. Nodes and Clients // URL: https://ethereum.org/en/developers/docs/nodes-and-clients/

15. Polkadot Keys // URL: https://wiki.polkadot.network/docs/learn-keys/

16. Rivest R. L. et al. The PACT protocol specification //Private Automated Contact Tracing Team, MIT, Cambridge, MA, USA, Tech. Rep. 0.1. – 2020.

17. Roca V. From ROBERT to DESIRE exposure notification: situation and lessons learned //Workshop on Security and Privacy in Contact Tracing. – 2020.

18. Shubina V. et al. Survey of decentralized solutions with mobile devices for user location tracking, proximity detection, and contact tracing in the COVID-19 era //Data. – 2020. – Т. 5. – №. 4. – С. 87.

19. Sukhwani H. et al. Performance modeling of hyperledger fabric (permissioned blockchain network) //2018 IEEE 17th International Symposium on Network Computing and Applications (NCA). – IEEE, 2018.

20. Wood, G. (n.d.). Ethereum: A secure decentralised generalised transaction ledger.

21. Zhou Q. et al. Solutions to scalability of blockchain: A survey //IEEE Access. – 2020, Т. 8.

Assessing the scaling of a blockchain-based social contact tracking system

Ilya Shmelev
senior researcher, GAUGN
Federal Research Center "Computer Science and Control"
Russian Federation, Moscow
Ivan Tarkhanov

Abstract

During a pandemic, one of the most important tasks is to track social contacts with those who are sick. This article categorizes projects that track these contacts. Projects are classified by architecture and the common components of such systems are highlighted. It is concluded that the hybrid architecture of such a solution based on an exclusive blockchain will have several advantages, and a conceptual model of such a system is described. However, an analysis of existing blockchain projects showed that their main problem is the unresolved issue of scaling such kinds of systems, which is becoming a key issue in the context of creating a global digital infrastructure of society. Further, the scaling of the system's conceptual model is assessed based on open-source information about the Moscow metro, and the main conclusions about the selected architectural solutions are confirmed.

Keywords: social contacts, COVID-19, blockchain, contact-tracing, scaling

Publication date: 16.09.2021

Citation link:

Shmelev I., Tarkhanov I. Assessing the scaling of a blockchain-based social contact tracking system // Artificial Societies – 2021. – V. 16. – Issue 3 [Electronic resource]. URL: https://artsoc.jes.su/S207751800015809-8-1 (circulation date: 23.04.2024). DOI: 10.18254/S207751800015809-8
