ПОИСК ЭФФЕКТИВНОГО АЛГОРИТМА, РЕАЛИЗУЮЩЕГО МЕХАНИЗМ ДОВЕРИЯ В P2P СЕТИ РАБОТАЮЩЕЙ ПО ПРОТОКОЛУ DHT KADEMLIA

SEARCH FOR AN EFFECTIVE ALGORITHM THAT IMPLEMENTS THE TRUST MECHANISM IN A P2P NETWORK RUNNING ON THE KADEMLIA DHT PROTOCOL

УДК-004

Бухтияров Иван Романович, Студент НИЯУ МИФИ, Россия, г. Москва
Михин Андрей Михайлович, Студент НИЯУ МИФИ, Россия, г. Москва
Сайфуллин Данис Русланович, Студент НИЯУ МИФИ, Россия, г. Москва
Тен Станислав Валерьевич, Студент НИЯУ МИФИ, Россия, г. Москва

Buhtiyarov Ivan Romanovich, Student of NRU MEPhI, Russia, Moscow
van4elotti@gmail.com
Mihin Andrej Mihajlovich, Student of NRU MEPhI, Russia, Moscow
Sajfullin Danis Ruslanovich, Student of NRU MEPhI, Russia, Moscow
Ten Stanislav Valer'evich, Student of NRU MEPhI, Russia, Moscow

Аннотация

Репутация и управление доверием продолжают играть значимую роль в совместных вычислениях (collaborative computing) для широкого спектра приложений, от облачных вычислений, децентрализованных сетевых вычислений, мобильных сервисов до социальных сетей и онлайн-сообществ. Общая проблема, с которой сталкиваются эти системы, заключается в том, как эффективно сотрудничать в выполнении задач, одновременно противостоя атакам на протяжении всей совместной работы. Это связано с тем, что участники большинства таких систем могут иметь мало знаний о других

участниках, с которыми ранее не было взаимодействия. Основная задача механизма доверия — распознавать и изолировать вредоносные узлы в сети. В данноӗ работе изложен механизм доверия для P2P сети основанноӥ на протоколе Kademlia DHT. Мы воспользуемся идееӥ механизма доверия EigenTrust [14], широко используемого к неструктурированным P2P-сетям. Мы определим место хранения рейтингов и значений доверия, правила предоставления положительных и отрицательных оценок и надежную систему оценивания, чтобы узлы не давали неверные оценки, а узлы доверительного управления не подделывали их. Прежде чем приступить к анализу существующих алгоритмов необходимо изложить основные алгоритмы и понятия, используемые при его построении.

Annotation

Reputation and trust management continue to play a significant role in collaborative computing for a wide range of applications, from cloud computing, decentralized network computing, mobile services to social networks and online communities. The common problem that these systems face is how to effectively collaborate on tasks while simultaneously resisting attacks throughout the entire collaboration. This is because participants in most of these systems may have little knowledge of other participants with whom they have not previously interacted. The main task of the trust mechanism is to recognize and isolate malicious nodes in the network. This paper describes the trust mechanism for a P2P network based on the Kademlia DHT Protocol. We will use the idea of the EigenTrust trust mechanism [14], which is widely used for unstructured P2P networks. We will determine where ratings and trust values are stored, rules for providing positive and negative ratings, and a reliable rating system so that nodes do not give incorrect ratings and trust management nodes do not fake them. Before proceeding to the analysis of existing algorithms, it is necessary to describe the main algorithms and concepts used in its construction.

Ключевые слова: алгоритм, сеть, механизм, P2P, DHT Kademlia

Keywords: algorithm, network, mechanism, P2P, DHT Kademlia

Одноранговые сети

P2P это распределенная система, в которой одноранговые узлы совместно используют ресурсы (файлы, компьютеры и службы) В системах P2P все узлы имеют одинаковые обязанности и возможности. В системе клиент-сервер ограничением является определение сервера и множества клиентов, в то время как в системе P2P это ограничение отсутствует. Таким образом, каждый узел играет симметричную роль, что означает, что одноранговый узел может быть клиентом в одном случае и сервером в другом случае. Узлы обмениваются сообщениями таким образом, что им не требуется центральный сервер. P2P-системы масштабируемы и могут быть гетерогенными; одноранговые узлы могут присоединяться или выходить из сети по желанию, а когда один из одноранговых узлов выходит из строя, другие одноранговые узлы продолжают работать.

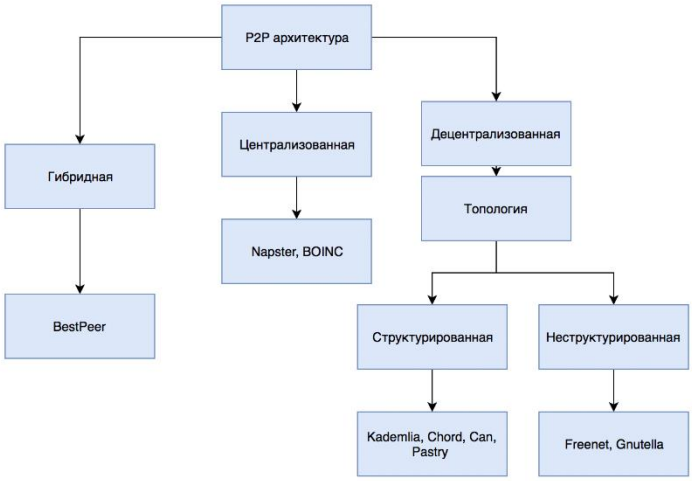

Рис. 1: Топология P2P систем

Маршрутизация

P2P сети обычно реализуют некоторую форму виртуальной сети, наложенную поверх физической сети, где узлы образуют подмножество узлов в физической сети. Данными по-прежнему обмениваются непосредственно над базовой TCP/IP сетью, а на прикладном уровне узлы имеют возможность взаимодействовать друг с другом напрямую, с помощью логических связей. Наложение используется для индексации и обнаружения узлов, что позволяет

системе P2P быть независимой от физической сети. На основании того, как узлы соединены друг с другом внутри сети, и как ресурсы индексированы и расположены, сети классифицируются на неструктурированные и структурированные (или как их гибрид).

Неструктурированные сети

Неструктурированная P2P сеть не формирует определенную структуру сети, а случайным образом соединяет узлы друг с другом. Так как не существует глобальной структуры формирования сети, неструктурированные сети легко организуются и доступны для локальных оптимизаций. Кроме того, поскольку роль всех узлов в сети одинакова, неструктурированные сети являются весьма надежными в условиях, когда большое количество узлов часто подключаются к сети или отключаются от нее.

Однако, из-за отсутствия структуры, возникают некоторые ограничения. В частности, когда узел хочет найти нужный фрагмент данных в сети, поисковый запрос должен быть направлен через сеть, чтобы найти как можно больше узлов, которые обмениваются данными. Такой запрос вызывает очень высокое количество сигнального трафика в сети, требует высокой производительности, и не гарантирует, что поисковые запросы всегда будут решены.

Структурированные сети

В структурированных P2P сетях организуется определенная топология, и протокол гарантирует, что любой узел может эффективно участвовать в поиске в файла/ресурса, даже если ресурс использовался крайне редко.

Наиболее распространенный тип структурированных сетей P2P реализуется распределенными хеш-таблицами (DHT), в котором последовательное хеширование используется для привязки каждого файла к конкретному узлу. Это позволяет узлам искать ресурсы в сети, используя хеш-таблицы, хранящие пару ключ-значение, и любой участвующий узел может эффективно извлекать значение, связанное с заданным ключом.

Тем не менее, для эффективной маршрутизации трафика через сеть, узлы структурированной сети должны обладать списком соседей, которые удовлетворяют определенным критериям. Это делает их менее надежными в сетях с высоким уровнем оттока абонентов (т.е. с большим количеством узлов, часто подключающихся к сети или отключающихся от нее).

Файлообменная сеть

Одна из областей применения технологии одноранговых сетей — обмен файлами. Пользователи файлообменной сети выкладывают какие-либо файлы в папку общего доступа («расшаренную» от англ. share — делиться) на своём компьютере, содержимое которой доступно для скачивания другим пользователям. Какой-нибудь другой пользователь сети посылает запрос на поиск какого-либо файла. Программа ищет у клиентов сети файлы, соответствующие запросу, и показывает результат. После этого пользователь может скачать файлы у найденных источников. В современных файлообменных сетях информация загружается сразу из нескольких источников. Её целостность проверяется по контрольным суммам.

Kademlia DHT

Kademlia — это реализация распределённой хеш-таблицы для одноранговых компьютерных сетей, разработанная Петром Маймунковым и Давидом Мазьером (David Mazi`eres). Протокол Kademlia определяет строгую структуру сети, регулирующей связь между узлами, и способ обмена информацией в ней. Узлы сети, работающей по протоколу Kademlia, общаются между собой по протоколу транспортного уровня UDP. Узлы Kademlia хранят данные посредством распределённых хеш-таблиц (DHT). В итоге над существующей LAN/WAN создаётся новая виртуальная или оверлейная сеть, в которой каждый узел обозначается специальным номером («Node ID»). Этот номер также выполняет и другие функции.

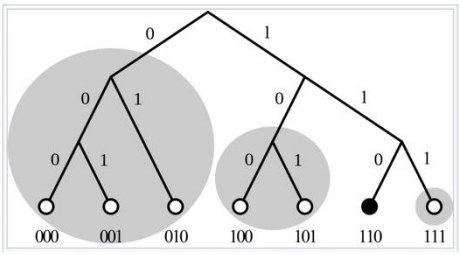

Рис. 2: Разделение сети относительно 110

В начале поиска узел создает временную таблицу поиска, содержащую k ближайших узлов к целевому идентификатору. Эта таблица всегда сортируется таким образом, что ближайший узел находится вверху списка, за ним следует следующий ближайший и т. д. Затем узел запрашивает α случайно выбранных узлов из этой таблицы для поиска более близких узлов. Исходные значения Kademlia для k и α равны 20 и 3 соответственно. Узлы, содержащиеся в ответах на поиск, вставляются в соответствующие места в таблице поиска. Исходный алгоритм Kademlia останавливается, как только найден первый узел, в котором хранится нужный элемент контента.

Реализация совместного использования файлов Kad в eMule использует 10 в качестве значения для k. Kad завершает поиск, когда все верхние k узлов таблицы поиска были опрошены для поиска более близких узлов и не вернули ни одного. Затем эти k узлов-кандидатов используются для выполнения Put или Get. Таким образом, в зависимости от выполняемой операции несколько копий элемента контента могут быть сохранены или извлечены.

Маршрутизация выполняется итеративно, то есть ответы отправляются на узел источника, и этот узел принимает решение о следующем переходе. Рекурсивная маршрутизация будет означать, что исходный узел запускает запрос, промежуточные узлы выполняют и пересылают его, а исходный узел только получает результат, не имея возможности влиять на выбор следующего перехода. Итеративная маршрутизация необходима для того, чтобы узел мог полностью контролировать процесс маршрутизации.

Атака насыщения вредоносными узлами в P2P сети

Одноранговые сети, как правило, полагаются на существование нескольких независимых и удаленных объектов, чтобы уменьшить угрозу от

фиктивных одноранговых узлов. Ключевая идея атаки насыщения вредоносными узлами или Sybil [9] состоит в том, чтобы добавить вредоносные узлы, которые все контролируются одним физическим объектом. Это позволяет при условии, что зловредные вершины расположены в определенных местах, получить контроль над частью сети P2P или даже над всей сетью, отслеживать трафик или злоупотреблять протоколом DHT другими способами. Мы предполагаем, что наш сценарий по своей сути подвержен влиянию вредоносных узлов, поэтому весьма вероятны угрозы, исходящие от злонамеренного поведения этих узлов. Угрозы, которые они могут генерировать, можно условно разделить на следующие две категории:

• Атака на хранение и извлечение данных

• Атака на таблицу маршрутизации

Вторая группа атак может принимать различные формы, такие как неправильная маршрутизация поиска или неправильные обновления таблиц маршрутизации. Первый тип атак может быть представлен следующим поведением: отказ от сохранения данных, имитация, что ресурсы не принадлежат узлу (хотя на самом деле принадлежат), или путем предоставления поддельных данных. В нашей работе мы сосредоточимся только на атаках на хранение и извлечение данных, оставляя другие атаки за рамками.

В настоящее время никто не нашел общего и эффективного распределенного способа противостоять атаке Sybil, несмотря на множество общих подходов, представленных в литературе. Нашей целью было разработать эффективную схему, применимую к конкретной структурированной P2P-сети Kademlia.

Определение доверия и репутации

Определение 1 Kumar et al. определяют [7] доверие как веру узла в способности, честность и надежность другого узла на основании его собственного непосредственного опыта.

Определение 2 Согласно Kumar et al. [7], репутацию определим как веру узла в возможности, честность и надежность другого узла в сети на основе отзывов и результатов транзакций других узлов, а также своего непосредственного опыта.

Для вычисления доверия должны быть приняты во внимания ее свойства:

• Неуверенность в среде: ценность доверия высока, когда большое количество участников должны сотрудничать, чтобы получить результаты из непредсказуемой среды, такой как P2P-сети.

• Контекстозависимость: доверие ограничено в определенном контексте

• Доверие субъективно: два узла могут иметь совершенно разные представления о третьем, чью надежность оценивают.

• Доверие не транзитивно.

• Доверие является однонаправленным: в оценке доверия между агентом и субъектом может отсутствовать взаимность.

Основное различие между понятиями доверия и репутации заключается в том, что доверие представляет собой личное или субъективное убеждение, зависящее от индивидуального восприятия или убеждения, тогда как репутация это глобальное восприятие или убеждение.

Отзывы и рейтинги, полученные от других узлов сети, влияют на изменение репутации [8]. Системы доверия и репутации помогают отделять легитимные и зловредные узлы в системах P2P собирая, распределяя и суммируя отзывы о действиях и поведении узлов в прошлом. Мы рассматриваем все взаимодействия между двумя партнерами как транзакции; например: запросы (GET / PUT), обмен файлами и.т.д.

Механизмы для вычисления репутации, могут быть разделены на две группы: Механизмы доверия, основанные на индивидуальной репутации, и механизмы доверия, основанные как на индивидуальной репутации, так и на социальных отношениях. В следующем разделах будут представлены некоторые системы доверия, основанные на индивидуальной репутации и социальных отношениях.

Итак, прежде всего, мы представим в следующем разделе анализ существующих известных мер для предотвращения атаки насыщения вредоносными узлами. А после этого эффективную контрмеру для Kademlia, основанную на доверии и репутация.

Существующие подходы

Механизмы управления репутацией используются для создания доверия в P2Pсистемах путем мониторинга поведения узлов в системе и предоставления им возможности оценивать транзакции. В этом разделе мы рассмотрим современное состояние систем управления репутацией. Будет также рассмотрен представитель систем безопасности, которые не используют механизмы доверия — S/Kademlia[10].

S/Kademlia

Баумгарт и Миес представляют S/Kademlia в [10]. S/Kademlia стремится улучшить безопасность Kademlia путем противодействия нескольким видам атак: авторы предлагают использовать безопасный процесс генерации идентификатора для предотвращения атак Sybil и Eclipse. Кроме того, другие узлы не должны иметь возможность украсть идентификатор узла. Для этого S/Kademlia использует пары асимметричных ключей: используя цифровые подписи, каждое сообщение аутентифицируется, подтверждая владение соответствующим закрытым ключом. Идентификатор узла состоит из значения хеш-функции открытого ключа. Чтобы противостоять атакам Sybil и Eclipse, авторы предлагают использовать либо центральный центр сертификации (CA) для подписи открытого ключа узла, чтобы подтвердить его подлинность, либо использовать криптопазлы(crypto puzzles). Для последнего варианта они вводят два типа криптопазлов: статический это отрицает свободное генерирование идентификатора узла и динамическое, что увеличивает сложность создания большого количества узлов Сибил.

1. Однако использование центрального органа (CA) против атаки Сибил требует взаимодействия с ним только один раз во время начальной загрузки и

один раз каждый раз, когда изменяется идентификатор узла. Если бы он также управлял

оценками и значениями доверия, ему потребовалось бы значительно больше ресурсов, потому что любая транзакция любого узла в сети вызывала бы запросы доверия во время транзакций и оценки после транзакции.

2. Использование крипто-головоломок является спорным: если сеть P2P работает в однородной среде в отношении вычислительной мощности, крипто-головоломки могут быть контрмера против атаки Sybil. Если вычислительная мощность узлов отличается на несколько порядков, однако, то такое решение не будет подходить.

PowerTrust

Разработанная система репутации PowerTrust [11] динамически выбирает небольшое количество узлов с наибольшей репутацией, используя механизм распределенного ранжирования. Эти узлы называются «Power Nodes» и играют ключевую роль в вычислении глобальных и локальных значений репутации других узлов. PowerTrust известен своей способностью повысить глобальную точность репутации и скорость агрегации значений доверия. С другой стороны, требуется большое время, чтобы узлы были выбраны и идентифицированы.

TrustMe

В отличие от систем, использующих социальные сети, TrustMe [12] стремится сохранить анонимность узлов. TrustMe стремится скрыть узлы управления доверием узла от самого себя, чтобы узел не мог пытаться манипулировать своей ценностью доверия. Доверительные узла выбираются случайным образом во время начальной загрузки и не определяются алгоритмом, а значения доверия запрашиваются с помощью широковещательных рассылок.

Другие механизмы доверия

EigenTrust [14] и FuzzyTrust [13] являются двумя популярными системами доверия для неструктурированных P2P сетей. Их алгоритм работы похож на рейтинговую систему eBay, поскольку каждая транзакция может быть оценена

как положительная или отрицательная. Оба алгоритма определяют методы получения информации о доверии из этих значений. EigenTrust использует распределенный способ для вычисления значения глобального доверия. А в FuzzyTrust каждый узел имеет свой собственный набор нечетких правил чтобы получить значение доверия. Ценность доверия основана на количестве положительных и отрицательных отзывов и имеет значение в диапазоне [-1; 1].

Литература

1. [1] Guido Urdaneta, Guillaume Pierre и Maarten van Steen, a Survey of DHT Security Techniques, ACM Computing Surveys, vol. 43, № 2, стр. 2, июнь. 2011. Июль 2014 года
2. [2] Петр Maymounkov и Дэвид Mazieres, "адреса: в одноранговой информационной системы на основе Яндекс Метрики"в IPTPS: пересмотренные документы с Первого международного семинара по равный-одноранговых систем, Кембридж, Массачусетс, США, 2002, стр. 53-65.
3. [3] Исправив, С. Д., Шлоссер, М. Т., Гарсия-Молина, Х., 2003. Алгоритм EigenTrust для управления репутацией в P2P-сетях. В: Тез.Докл. 12-й Международной конференции по всему миру.
4. [4] Xiong, L., Liu, L., 2004. PeerTrust: поддержка доверия на основе репутации для одноранговых электронных сообществ. IEEE Trans. Знания. Сведения Англ. 16 (7), 843–857.
5. [5] Чжоу, Р., Хван, К., 2007. PowerTrust: надежная и масштабируемая система репутации для доверенных одноранговых вычислений. IEEE Trans. Параллель и распределение. Система. 18 (4), 460–473.
6. [6] Диего Гамбетта. Можем Ли Мы Доверять Доверию?, глава доверие: установление и разрыв кооперативных отношений, страницы 213-237. Факультет социологии Оксфордского университета, 1980 год.

7. Сандип Кумар, Чандер Дивейкер и Амит Чаудхари. Система репутации в одноранговой сети: проектирование и классификация. Журнал глобальных исследований в области компьютерных наук, 2 (8), 2011.
8. Омар Абдель Вахаб, Джамаль Бентахар, Хади отрок и Аззам Мурад. Обзор моделей доверия и репутации для веб-служб: одиночных, составных и сообществ. Системы Поддержки Принятия Решений, 74: 121-134, 2015.
9. [10] J. R. Douceur, The Sybil Attack, in: Proceedings of the IPTPS First Inter - national Workshop, Cambridge, MA, USA, March 2002.
10. [11] Ingmar Baumgart and Sebastian Mies, "S/Kademlia: A Practicable Approach to Secure Key - Based Routing", in Proceedings of the 13th IEEE International Conference on Parallel and Distributed Systems (ICPADS), Hsinchu, Taiwan, 2007, PP.1-8.
11. [12] Чжоу, Р., Хван, К., 2007. PowerTrust: надежная и масштабируемая система репутации для доверенных одноранговых вычислений. IEEE Trans. Параллель и распределение. Система. 18 (4), 460–473.
12. [13] Aameek Синг и Лю Лин, "поверь мне: анонимный управления доверительных отношений в децентрализованной P2P-системам,"в материалах 3-й Международной конференции IEEE по одноранговой вычислительной (P2P), Линко пинг, Швеция, 2003, с. 142-149.
13. [14] Nathan Griffiths, Kuo-Ming Chao, and Muhammad Younas, "Fuzzy Trust for Peer-to-Peer Systems", in Proceedings of the 26th IEEE International Conference Workshops on Distributed Computing Systems (ICDCSW), Lisbon, Portugal, 2006.
14. [15] Сепандар Д. Камвар, Марио т. Шлоссер и Гектор Гарсия-Молина, "алгоритм Eigentrust для управления репутацией в P2P-сетях", в трудах 12-й Международной конференции ACM по всемирной паутине (WWW), Будапешт, Венгрия, 2003, стр. 640-651.

Literature

1. [1] Guido Urdaneta, Guillaume Pierre, and Maarten van Steen, A Survey of DHT Security Techniques, ACM Computing Surveys, vol. 43, no. 2, p. 2, June. 2011. Jul 2014
2. [2] Petar Maymounkov and David Mazieres, "Kademlia: A Peer-to-Peer Information System Based on the XOR Metric,"in IPTPS: Revised Papers from the First International Workshop on Peer-to-Peer Systems, Cambridge, MA, USA, 2002, pp. 53-65.
3. [3] Kamvar, S.D., Schlosser, M.T., Garcia-Molina, H., 2003. The EigenTrust Algorithm for Reputation Management in P2P Networks. In: Proc. of the 12th International Conference on World Wide.
4. [4] Xiong, L., Liu, L., 2004. PeerTrust: supporting reputation-based trust for peer-to-peer electronic communities. IEEE Trans. Knowl. Data Eng. 16 (7), 843–857.
5. [5] Zhou, R., Hwang, K., 2007. PowerTrust: a robust and scalable reputation system for trusted peer-to- peer computing. IEEE Trans. Parallel and Distrib. Syst. 18 (4), 460–473.
6. [6] Diego Gambetta. Can We Trust Trust?, chapter Trust: Making and Breaking Co- operative Relations, pages 213–237. Department of Sociology, University of Oxford, 1980.
7. [7] Sandeep Kumar, Chander Diwaker, and Amit Chaudhary. Reputation system in Peer- to-Peer network: Design and classification. Journal of Global Research in Computer Science, 2(8), 2011.
8. [9] Omar Abdel Wahab, Jamal Bentahar, Hadi Otrok, and Azzam Mourad. A survey on trust and reputation models for Web services: Single, composite, and communities. Decision Support Systems, 74:121–134, 2015.
9. [10] J.R. Douceur, The Sybil Attack, in: Proceedings of the IPTPS First International Workshop, Cambridge, MA, USA, March 2002.
10. [11] Ingmar Baumgart and Sebastian Mies, "S/Kademlia: A Practicable Approach Towards Secure Key- Based Routing,"in Proceedings of the 13th

IEEE International Conference on Parallel and Distributed Systems (ICPADS), Hsinchu, Taiwan, 2007, pp. 1-8.

11. [12] Zhou, R., Hwang, K., 2007. PowerTrust: a robust and scalable reputation system for trusted peer- to-peer computing. IEEE Trans. Parallel and Distrib. Syst. 18 (4), 460–473.
12. [13] Aameek Singh and Ling Liu, "TrustMe: Anonymous Management of Trust Relationships in Decentralized P2P Systems,"in Proceedings of the 3rd IEEE International Conference on Peer-to- Peer Computing (P2P), Linko ̈ping, Sweden, 2003, pp. 142-149.
13. [14] Nathan Griffiths, Kuo-Ming Chao, and Muhammad Younas, "Fuzzy Trust for Peer-to-Peer Systems,"in Proceedings of the 26th IEEE International Conference Workshops on Distributed Computing Systems (ICDCSW), Lisbon, Portugal, 2006.
14. [15] Sepandar D. Kamvar, Mario T. Schlosser, and Hector Garcia-Molina, "The Eigentrust Algorithm for Reputation Management in P2P Networks,"in Proceedings of the 12th ACM International Conference on World Wide Web (WWW), Budapest, Hungary, 2003, pp. 640-651.