Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.724

Проведенцев Евгений Сергеевич

Децентрализованные социальные сети

АВТОРЕФЕРАТ
на соискание академической степени
магистра технических наук

по специальности 1-40 80 05 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

Научный руководитель
Сечко В. В.
к.т.н., доцент

Минск 2015

КРАТКОЕ ВВЕДЕНИЕ

В настоящее время Интернет занимает огромное место в жизни человека. В отличие от Интернета «раннего» периода, в настоящий момент он не представлен просто сайтами, на которых размещена та или иная информация, а также ссылки на прочие ресурсы либо каталоги файлов. Это уже и не просто совокупность сервисов наподобие онлайн-переводчиков, электронных словарей или web-интерфейсов для электронных почтовых ящиков.

В современном Интернете web-сервисы начинают интегрироваться с «настольными» приложениями, перенимая типичные для них интерфейс, поведение и возможности. Среда Интернет позволяет также организовать многопользовательскую работу с документами, проектами, обмениваться различной информацией в едином сервисе с людьми, зарегистрированными в нем и разделенными большими расстояниями.

Одним из видов сервисов коллективной работы пользователей являются социальные сети. Социальная сеть – это совокупность социальных элементов (людей, групп, организаций, сообществ), которые обмениваются между собой различными данными. Такими данными могут быть как простые текстовые сообщения, так и изображения, видео- и аудиофайлы.

Переходя к теме простых сообщений, следует отметить, что в большинстве случаев они не несут значимой смысловой нагрузки и сводятся к простому диалогу между пользователями, который мог бы быть произойти, к примеру, при личной встрече. Исходя из этого было принято решение использовать социальную сеть для обмена действительно полезной информацией. Причем информация эта необязательно должна быть полезна определенному кругу людей – знакомых, коллег, учащихся одного учебного заведения: можно давать информацию, которая была бы полезна более широким группам людей, в конечном итоге – их социуму.

В настоящее время практически все популярные социальные сети централизованы. Вместе с преимуществами это несет ряд недостатков, таких как: риск утечки информации с центральных серверов, центральные точки подвержены риску выйти из строя, после чего сеть перестает работать.

Децентрализованная структура сети означает, что она не находится исключительно в одном месте и не контролируется только одной организацией. Обмен данными происходит непосредственно между клиентами. Используя децентрализованную структуру возможно реализовать социальную сеть, предоставляющую аналогичную классическим сетям функциональность.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является разработка протокола передачи данных для децентрализованной социальной сети. Протокол должен обеспечивать возможность реализации всех необходимых функций, присущих социальным сетям.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Изучить принципы работы с одноранговыми сетями.
2. Изучить разновидности криптосистем для шифрования данных и выбрать среди них наиболее подходящие для решения поставленной задачи.
3. Разработать протокол передачи данных, на базе которого может быть построена децентрализованная социальная сеть. Протокол должен обеспечивать возможность реализации всех необходимых функций, присущих социальным сетям.
4. Разработать архитектуру клиентского приложения программного средства.
5. Реализовать клиентское приложение на базе разработанного протокола передачи данных, с помощью которого пользователи получают возможность обмена сообщениями и файлами, управлять своим и просматривать профили других пользователей, организовывать и участвовать в сообществах.

Объектом исследования являются децентрализованные социальные сети.

Предметом исследования методы и алгоритмы передачи и шифрования данных в децентрализованных сетях, применение данных методов и алгоритмов в целях реализации функций социальных сетей.

Основной *гипотезой*, положенной в основу диссертационной работы, является возможность использования децентрализованной архитектуры социальной сети с целью избегания угрозы безопасности персональных данных пользователя социальной сети.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

Работа выполнялась в соответствии с научно-техническим заданием и планом работ кафедры «Программное обеспечение информационных технологий» по теме «Разработать модели, методы, алгоритмы для оценки параметров, повышения надежности и качества функционирования аппаратно-программных средств систем и сетей сложной конфигурации и внедрить в современные обучающие комплексы» (ГБ № 11-2004, № ГР 20111065, научный руководитель НИР – В. В. Бахтизин).

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя В. В. Сечко, заключается в формулировке целей и задач исследования.

Апробация результатов диссертации

Основные положения диссертационной работы докладывались и обсуждались на IX Международной научно-практической конференции «Научный поиск в современном мире» (Махачкала, Россия, 2015).

Опубликованность результатов диссертации

По теме диссертации опубликована одна работа в сборниках трудов и материалов международных конференций.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, пяти глав, заключения, списка использованных источников, списка публикаций автора и приложений. В первой главе представлен анализ предметной области, выявлены основные существующие проблемы в рамках тематики исследования, показаны направления их решения. Вторая глава посвящена исследованию предметной области и поиску существующих методов и алгоритмов в сфере децентрализованных социальных сетей. В третьей главе предложены методы и алгоритмы передачи данных для построения протокола передачи данных, предназначенного для использования в децентрализованных социальных сеях. В четвертой главе предложена практическая реализация клиентского приложения децентрализованной социальной сети, выявлены особенности реализации децентрализованных алгоритмов передачи данных. В пятой главе продемонстрирован внешний вид клиентского приложения, а также даны рекомендации по установке и эксплуатации клиентского приложения.

Общий объем работы составляет 52 страницы, из которых основного текста – 40 страниц, 11 рисунков на 10 страницах, 4 таблицы на 5 страницах, список использованных источников из 26 наименований на 2 страницах и 1 приложение на 5 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

Во **введении** определена область и указаны основные направления исследования, показана актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначена практическая ценность работы.

В **первой главе** проведен анализ применяемых архитектурных решений систем, используемых при разработке онлайновых социальных сетей. Выявлены основные недостатки и угрозы безопасности социальных сетей, построенных на централизованной архитектуре. Определены основные пути решения главных недостатков социальных сетей, построенных на централизованной архитектуре.

Одноранговые сети являются наиболее легким и дешевым типом сетей для установки. Большинство одноранговых сетей требует наличия на компьютерах, кроме сетевой карты и сетевого носителя (кабеля), только операционной системы. Как только компьютеры соединены, пользователи немедленно могут начинать предоставление ресурсов и информации в совместное пользование.

Каждый пользователь в одноранговой сети является одновременно сетевым администратором. Это означает, что каждый пользователь в сети управляет доступом к ресурсам, расположенным на его компьютере. Он может дать всем остальным неограниченный доступ к локальным ресурсам, дать ограниченный доступ, а может не дать вообще никакого доступа другим пользователям. Каждый пользователь также решает, дать другим пользователям доступ просто по их запросу или защитить эти ресурсы паролем.

Основной проблемой использования одноранговых сетей при разработке социальных сетей является применение классических алгоритмов передачи данных, предполагающих наличие централизованного сервера хранения и передачи данных, в децентрализованной структуре сети.

Во второй главе произведен анализ существующих методов и алгоритмов передачи данных в сфере децентрализованных социальных сетей.

В первой части второй главы рассматривается алгоритм идентификации записей социальной сети с помощью панхэшей. Каждая запись в программном средстве имеет уникальный идентификатор – композитный (составной) ключ. Этот идентификатор состоит из хеш-кодов используемых полей. Текстовые поля хешируются при помощи алгоритма SHA1, даты кодируются с использованием 3 байт (в днях, начиная от 1 января 1900 года), географическая координата кодируется в «ленточную» 4-байтовую координату от Северного полюса, при ссылке на другие записи используются ключи этих записей. Неполные записи (ключи которых имеют нулевые пропуски) могут соотносится с более полными (у которых ключи имеют меньше пропусков). Такое соотношение будем называть *подобием*. Поиск подобия может быть полезен в запросах и настраивается дополнительно.

Во второй части главы рассматривается ленточная координата. Широта имеет диапазон 180 градусов, а долгота – 360 градусов. Если отводить по 2 бай-

та на каждую координату, то погрешности распределятся ассиметрично: 0,0014 и 0,0027 градусов соответственно.

Чтобы выровнять погрешность, используется «ленточная координата», при которой 4-байтовое число показывает порядное смещение в матрице от левого верхнего угла (Северного полюса), равного 0, до правого нижнего (Южного полюса), равного чуть меньше 255^4 (размер матрицы ограничен 4 байтами).

Матрице в градусах 360*180 = 64 800 соответствует матрица значений 92 681*46 340 = 42 94 837 540.

В третьей части главы рассматривается общая схема передачи данных между пользователями децентрализованной социальной сети. В частности, приводится необходимость ввода хеш-загадки. Хеш-загадка обязывает инициатора обновления затратить аппаратные ресурсы, чтобы перейти к следующему шагу. Если требуется решить хеш-загадку, то последним байтом в фразе задается количество первых бит фразы (по умолчанию 14 бит), которые считаются хеш-загадкой. Инициатор обновления должен подобрать к фразе любую добавку так, чтобы первая часть хеша SHA1 от блока «фраза+добавка» совпала с хеш-загадкой. Так как аналитического решения у этой задачи нет, а составить радужные таблицы для всех фраз длинной 256 байт не представляется возможным, то инициатору обновления приходится искать отгадку методом перебора.

В **третьей главе** описана общая схема функционирования децентрализованной социальной сети: описана структура сети, включающая такие понятия как окружение, ядро окружения, точка входа, доверенный сервис аутентификации. Описаны алгоритмы выполнения базовых операций в децентрализованной социальной сети, построенной на подобной архитектуре. Дан анализ реализации распределенной хеш-таблицы Kademlia, описаны особенности применения алгоритмов данной DHT в описываемой архитектуре. Описана модель данных, применяемая в предлагаемой архитектуре децентрализованной социальной сети.

Описаны способы управления создания и поиска записей в децентрализованной социальной сети. Каждая запись помечается хеш-кодом создателя. Если пользователь ввел запись, она автоматически помечается как «поддерживаемая». До тех пор, пока пользователь держит включенной пометку «поддерживаю» сеть будет хранить эту запись в базе. При поиске узел сети перебирает известные узлы, при этом пользователь ищет слушающие узлы. Если приложение пользователя нашло слушателя, оно подключатся к нему, и начинается обмен данными. Обычно, узел сети находится и в режиме слушания, и в режиме поиска одновременно. Хотя соединение всегда инициирует клиент, который сделал запрос на поиск, сразу после подключения обмен данными идёт в обе стороны. При этом каждый узел выставляет свои запросы, а другой узел должен ответить на эти запросы.

Описаны способы задания уровня доступа (или уровня «доверия») к записям пользователей децентрализованной социальной сети. При указании доверия необходимо указать значение уровня доверия от минус 1 до 1. Доверие равное нулю означает, что пользователь просмотрел эту запись и считает её нейтральной. Когда изменяется доверие к записи, старая подпись остаётся, и рядом за-

водится новая подпись. Таким образом подписей может быть много. История подписей служит показателем эволюции пользовательского мнения. Но при расчёте рейтингов используется только самая последняя по времени создания подпись. Также описан способ обработки комментариев к записям (мнений).

В **четвертой главе** описано создание прототипа клиентского приложения децентрализованной социальной сети. Обоснован выбор технологий разработки приложения. Дано описание назначения основных классов ПС. Описаны особенности протокола передачи данных на транспортном уровне. Протокол разрабатываемого программного средства является байтовым (бинарным). Обмен происходит небольшими порциями — сегментами. Сегмент имеет следующую структуру: индекс, код команды, уточняющий код, данные. Реплики в пределах сеанса последовательно нумеруются. Реплика-ответ кроме своего номера также содержит дополнительно номер реплики-вопроса, на который она отвечает.

В **пятой главе** представлен внешний вид клиентского приложения децентрализованной социальной сети. Даны рекомендации по установке и эксплуатации программного средства.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

В результате исследования был разработан протокол передачи данных, а также разработано программное средство на базе данного протокола для организации социальных взаимоотношений, работающее в одноранговой сети. Отсутствие центрального сервера понижает риск выхода сети из строя и позволяет пользователям безопасно обмениваться данными.

Основные результаты исследования работы:

– проанализированы аналоги разрабатываемого программного средства, на основании чего сформулированы требования к разрабатываемому программному средству;

– разработан протокол передачи данных, достаточный для поддержки всех необходимых функций, присущих социальным сетям;

– спроектировано и разработано клиентское приложение программного средства, выполняющие также функции сервера.

Проведенное тестирование показало, что выполняются все функции, описанные в спецификации, удовлетворяют все требования, связанные с работоспособностью, удобством использования и безопасностью. Система выдерживает необходимую нагрузку и не потребляет много компьютерных ресурсов.

Рекомендации по практическому использованию результатов

1. Полученные результаты формируют теоретическую и практическую базу для разработки архитектур децентрализованных социальных сетей. Они могут быть использованы для модернизации и дальнейшего развития существующих децентрализованных социальных сетей.

2. Разработанные методы и алгоритмы передачи данных в одноранговых сетях могут применяться в разработке любого ПО, связанного с передачей данных в децентрализованных сетях, в частности, в децентрализованных социальных сетях.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Проведенцев Е. С., Децентрализованный подход в разработке онлайновых социальных сетей / Е. С. Проведенцев // IX международная конференция «Научный поиск в современном мире» – Махачкала: Апробация, 2015. – с. 24-28.