

На правах рукописи

ОБЕЙДАТ АТЕФ АХМЕД

**УПРАВЛЕНИЕ ДОСТУПОМ К ОБЩИМ РЕСУРСАМ
В ПИРИНГОВЫХ СИСТЕМАХ**

Специальности: 05.13.13 - Телекоммуникационные системы и
компьютерные сети

А В Т О Р Е Ф Е Р А Т
диссертации на соискание ученой степени
кандидата технических наук

Санкт-Петербург

2009

Работа выполнена в Государственном образовательном учреждении высшего профессионального образования «Новосибирский государственный технический университет» на кафедре «вычислительной техники»

Научный руководитель:

Заслуженный деятель науки РФ,
Заслуженный работник ВШ РФ,
д.т.н., профессор Губарев Василий Василевич

Официальные оппоненты:

доктор технических наук, профессор Яковлев Сергей Алексеевич
кандидат технических наук, доцент Данилов Виталий Иванович

Ведущая организация: ГОУ ВПО «Сибирский государственный университет телекоммуникаций и информатики », г.Новосибирск

Защита состоится 23 декабря 2009г. в 15.00 часов на заседании совета по защите докторских и кандидатских диссертаций Д 212.238.01 Санкт-Петербургского государственного электротехнического университета им. В.И.Ульянова (Ленина) по адресу: 197376, Санкт-Петербург, ул. Проф. Попова, 5.

С диссертацией можно ознакомиться в библиотеке университета.

Автореферат разослан "19 " ноября 2009г.

Ученый секретарь
совета по защите докторских и
кандидатских диссертаций

Пантелеев М.Г.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Пиринговые сети (от англ. peer-to-peer (P2P) – ровня с ровней, равный с равным (P2P) или пользователь – пользователь (П2П)) – это среда, основанная на равноправии всех участников, среди которых нет специально выделенных серверов данных или приложений, а каждый его участник – узел (peer) является как клиентом (получателем информации), так и сервером (поставщиком информации). Иными словами P2P-сети являются распределенной дуальной системой (средой), содержащей взаимосвязанные узлы, способные к самоорганизации в сетевую технологию с целью разделения различных ресурсов (контентов, циклов процессоров, устройств хранения, полос пропускания и т.п.), самоадаптирующихся к отказам и переменному числу узлов, поддерживая при этом приемлемый уровень связности и производительности без необходимости в посредниках или поддержке глобального центрального сервера. Они изначально ориентированы на нестабильность, переменное количество узлов в сети, устойчивость к отказам и самоадаптацию. Их дуальность (двойственность) связана с тем, что, с одной стороны, это совокупность распределенных взаимосвязанных узлов, абсолютно равноправных с точки зрения функциональности и выполняемых задач, необязательно однородных, с другой, совокупность приложений, которые используют одни и те же объекты (ресурсы, диски, циклы процессоров, контент и т.п.).

В настоящее время резко возрос интерес к пиринговым сетям (П2П), в которых, посредством Интернета в качестве средств (среды) обмена неизменяемых объектов, используются такие системы как Kazaa, Gnutella и BitTorrent. Недавние исследования показывают, что в течение последних лет 70% трафика Интернета связаны с деятельностью пиринговых сетей. Это обусловлено тем, что использование пиринговых сетей для управления ресурсом даёт преимущества в виде уменьшения стоимости, хорошей масштабируемости и поддержки специальных сетей. Наряду с таким развитием появилась потребность в приложениях, позволяющих сотрудничать друг с другом большому количеству пользователей. До сих пор большинство решений связано с моделью «клиент-сервер», в которой сервер контролирует обновление общих данных приложения. Эта модель позволяет легко поддерживать одновременно доступ к ресурсам и аутентичность данных приложений, однако сервер становится точкой отказа системы и «узким местом», которое может препятствовать тому, чтобы приложение было доступно большому количеству пользователей.

Более перспективным является развитие децентрализованных сетей. Однако в них появляются такие проблемы как взаимное исключение, тиражирование и обеспечение аутентичности (тождественности) дубликатов динамически изменяемых объектов (файлов или вычислительных ресурсов). Эти проблемы особенно обостряются в силу следующих обстоятельств. 1) В связи с желанием выполнять с объектами исключаящие друг друга одновременные действия, например, чтение и запись, изменение содержания

оригинала объекта одним пользователем и одновременное чтение его дубликата другим. 2) Из-за непредсказуемой изменчивости объектов, структуры сети, необходимости при этом обеспечения нужного тиражирования дубликатов и их аутентичности (тождественности) объекту.

Основные вопросы проблемы взаимного исключения одновременного доступа к общим ресурсам в вычислительных системах и сетях рассматриваются в работах L. Lamport. , R. Agrawala, G. Ricart , A. Agrawala, M. Singhal, K. Raymond, M. Maekawa и др. В результате для централизованных вычислительных сетей предложено множество решений этой проблемы. Однако специфика динамично изменяемых децентрализованных пиринговых сетей не всегда позволяет эффективно использовать известные решения из-за различий в базовых моделях сетей, их строения и функционирования, отсутствия централизации, быстрой непредсказуемой изменчивости оригинальных объектов, требующих изменения дубликатов.

Поэтому, несмотря на большое количество существующих решений взаимного исключения, текущего тиражирования и обеспечения при этом аутентичности копий, в традиционных средствах параллельных вычислений, описанных в литературе, децентрализованность, изменчивость и переменная масштабируемость пиринговых сетей вынуждают искать новые решения.

В связи с этим тема работы является актуальной и практически важной.

Целью работы является разработка и исследование модельного, алгоритмического и программного обеспечения эффективного бесконфликтного доступа в динамических пиринговых сетях к общим изменяемым объектам, т.е. доступа с гарантированным исключением одновременного выполнения противоречивых запросов на работу с одним и тем же объектом и обеспечением множественности и аутентичности его дубликатов.

В соответствии с поставленной целью решаются следующие **задачи исследования:**

1. Выбрать класс семейства пиринговых сетей, как базовую основу для разработки и исследования алгоритмов, рассматриваемых в диссертации, и создать для него системную модель. Выполнить анализ архитектурных особенностей выбранного семейства пиринговых сетей, провести систематизацию и исследование существующих методов и алгоритмов предоставления распределенного доступа к объектам в них.

2. Разработать бесконфликтные масштабируемые отказоустойчивые алгоритмы, обеспечивающие эффективный доступ к совместно используемым изменяемым объектам в динамических пиринговых сетях выбранного класса, быструю тиражируемость и аутентичность дубликатов непредсказуемо изменяемых объектов.

3. Создать средства для оценки показателей качества предложенных алгоритмов и провести сравнительное исследование алгоритмов.

4. Опробовать разработанные алгоритмы доступа к дубликатам объектов

Методы исследования. Для решения задач, поставленных в диссертации, используются методы анализа эффективности функционирования распределенных вычислительных систем, элементов SWOT-анализа, математической статистики, имитационного моделирования.

Научная новизна полученных в диссертационной работе результатов состоит в следующем.

1. Создана формализованная системная модель класса рассматриваемых сетей.

2. Систематизированы и проанализированы существующие алгоритмы взаимного исключения одновременного доступа к общим объектам, тиражирования и обеспечения аутентичности дубликатов объектов в динамических пиринговых сетях.

3. Разработаны новые бесконфликтные алгоритмы доступа к общим ресурсам, обеспечивающие взаимное исключение одновременного доступа к запрашиваемому объекту: алгоритм «узел-координатор» (У2К) и древовидный алгоритм взаимного исключения (ДАВИ).

4. Разработаны алгоритмы тиражирования и обеспечения аутентичности (непротиворечивости) дубликатов динамически изменяемых объектов.

5. Разработано программное обеспечение для имитационного исследования предложенных алгоритмов по основным показателям их качества, проведено апробирование и исследование алгоритмов.

Практическая ценность работы заключается в том, что:

1. Использование предложенных алгоритмов и методов позволяет разработчикам при написании P2P-приложений не беспокоиться о поддержании низкоуровневых механизмов защиты от одновременного доступа к совместно используемым объектам и об обеспечении аутентичности их дубликатов.

2. На основе созданных алгоритмов возможна разработка большого количества приложений, названных «совместным окружением», которые позволяют различным пользователям иметь общий доступ к информации и изменять ее в режиме реального времени в распределенном виде, в частности перенос приложений, применяющихся на модели «клиент-сервер», в P2P системах. Примерами является партнерское текстовое редактирование, видео конференц-связь, он-лайн игры, он-лайн аукционы, образовательные приложения и т.д.

3. Предложенные алгоритмы взаимного исключения имеют меньшую сложность. Результаты имитации алгоритмов выявили, что они обладают намного большей масштабируемостью, требуют в среднем в 3-4 раза меньшего количества служебных сообщений.

4. Созданное на основе системы Free-pastry приложение позволяет обеспечить функционирование и тестирование предложенных алгоритмов.

Реализация работы. Созданный программный пакет, реализующий алгоритмы управления доступом к объектам в пиринговых сетях, был принят к

внедрению в компаниях, занимающихся разработкой и внедрением программных систем, – ООО «Инфо-Стрим» и ООО «ИДЕЛЬ» для решения задач исследования моделей проектируемых приложений. Результаты работы также используются в учебном процессе Новосибирского государственного технического университета (НГТУ), что подтверждено соответствующими документами, копии которых приведены в приложении к диссертации.

Апробация работы. Основные положения диссертационной работы докладывались и обсуждались на: II Всероссийском смотре научных и творческих работ иностранных студентов и аспирантов, 28- 29 апреля, 2008 г., Томск, Россия; Международной конференции «The Third International Forum on Strategic Technologies (IFOST)», 23-29 июня, 2008 г., Новосибирск, Россия; Международной конференции «The 10th International Workshop on Computer Science and Information Technologies (CSIT)», 15-17 сентября, 2008г., Antalya, Turkey; Российской научной конференции «НАУКА, ТЕХНОЛОГИИ, ИННОВАЦИИ», 4-7 декабря, 2008г., Новосибирск, Россия; XV международной научно-технической конференции студентов и аспирантов «РАДИОЭЛЕКТРОНИКА, ЭЛЕКТРОТЕХНИКА И ЭНЕРГЕТИКА»// 26 - 27 февраля, 2009г., Москва, Россия; Всероссийской конференции «Винеровские чтения 2009» , 11 - 16 марта, 2009 г., Иркутск, Россия; Международной конференции «International Siberian Conference on Control and Communications (SIBCON-2009)», 27–28 марта , 2009 г., Томск, Россия; Седьмой всероссийской научно-практической конференции студентов, аспирантов и молодых ученых с международным участием «Молодежь и современные информационные технологии», 24-27 февраля, 2009г., ТПУ, г. Томск, Россия; II Всероссийской научно-практической конференции «Научная инициатива иностранных студентов и аспирантов российских вузов», 21- 22 мая, 2009г., ТПУ, г.Томск, Россия; Международной конференции «The Forth International Forum on Strategic Technologies (IFOST 2009)», 21-23 октября, 2009 г., Ho Chi Minh city, Vietnam.

На защиту выносятся следующие результаты

1. Эффективные алгоритмы бесконфликтного доступа к запрашиваемому объекту в децентрализованных пиринговых системах: алгоритм «узел-координатор» (У2К) и древовидный алгоритм взаимного исключения (ДАВИ).
2. Алгоритмы тиражирования динамично получаемых дубликатов непредсказуемо изменяемых объектов с обеспечением их аутентичности.
3. Программное обеспечение для проведения имитационного исследования предложенных алгоритмов по выбранным показателям качества и результаты их исследования.
4. Программное обеспечение для работы в пиринговых сетях согласно предложенным алгоритмам.

Публикации. Основные теоретические и практические результаты диссертации опубликованы в 15 работах, среди которых 2 публикации в ведущих рецензируемых изданиях, рекомендованных в действующем перечне ВАК, а также 13 докладов, перечисленных в конце автореферата.

Структура и объем работы. Диссертация состоит из введения, шести глав, заключения, списка литературы, включающего 105 наименований, из которых 88 иностранных, и двух приложений. Текст диссертации изложен на 171 странице, включает 36 рисунков и 8 таблиц. В приложении содержатся копии документов о внедрении и дипломов за участие в научно-технических конференциях.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность темы диссертационной работы, формулируются цель и задачи исследования, указываются методы исследования, приводятся сведения об основных результатах работы, их практической ценности, внедрении и апробации.

В первой главе определяются основные понятия, используемые в диссертации; приводится обзор литературных источников и анализ имеющихся по теме диссертации результатов.

Изложен обзор клиентских информационно-вычислительных средств с особым выделением пиринговых приложений для совместного доступа к общим ресурсам. Дается классификация пиринговых систем.

Описывается задача взаимного исключения в P2P-сетях – исключения одновременного доступа к одним и тем же файлам (а в более общей постановке - к любым другим объектам (ресурсам) разных пользователей), т.е. одновременный допуск с критическими операциями более одного пользователя к одному и тому же объекту сети. Части файлов, которые могут подвергаться критическим операциям, оформляются в виде критических секций (КС) областей объектов соответствующих узлов. Выявлены недостатки и слабые стороны существующих алгоритмов.

Описываются показатели качества задач взаимного исключения в P2P системах и их характеристики. В качестве таких показателей выбраны: *число сообщений*, пересылаемых друг другу узлами во время использования критической секции; *время T_0 ответа на запрос* – интервал времени, в течение которого запрос ждет разрешение на работу с критической секцией после отправки сообщения ЗАПРОС; *время T_{3C} задержки синхронизации* – временной интервал между двумя успешными обращениями к критической секции; *средняя задержка передачи сообщений \bar{T}_n* (они показывают, как скоро запросный узел может войти в критическую секцию); *пропускная способность L* – число запросов, обработанных за единицу времени.

Описываются задачи тиражирования и аутентичности данных. Тиражирование данных заключается в сохранении многочисленных копий объектов, называемых дубликатами, на отдельных узлах сети. Тиражирование повышает надежность (отказоустойчивость) сети, способствует увеличению доступности объектов и повышению эффективности (производительности) сети. Аутентичность (тождественность) данных означает эквивалентность всех копий оригиналу. Выявлены недостатки и слабые стороны существующих решений обеспечения аутентичности и тиражирования данных.

Предлагается формализованная системная модель класса динамических пиринговых сетей, на основе которой будут рассматриваться разрабатываемые

далее алгоритмы. Основные положения модели. Система состоит из равноправных узлов p , каждый из которых может быть клиентом (получателем) и сервером (поставщиком) ресурсов. Каждому доступному пользователям объекту R_j (реальному файлу или вычислительному ресурсу) соответствует определенный набор узлов, содержащих его точные копии (дубликаты). Взаимодействие узлов между собой осуществляется согласно правилам работы P2P-сети, которые предписывают любому узлу посылать сообщения другим узлам. Дубликаты объекта всегда доступны, но их внутреннее состояние может быть случайно сброшено вследствие сбоя. После обнаружения сбоя дубликат вновь восстанавливается в сети с прежним значением идентификатора узла. Количество клиентов сети непредсказуемо и может быть очень большим. Предполагается, что клиенты не настроены злонамеренно. В сети может быть высокая **динамичность** (текучесть клиентов) – узлы могут входить и покидать сеть в любое время; клиенты и узлы, содержащие дубликаты, взаимодействуют друг с другом посредством *сообщений*, посылаемых через ненадежные каналы. Каждый объект R_j имеет уникальный идентификатор k_j , который используется как хэш текстового имени объекта. Хэш может быть вычислен с помощью конфликтоустойчивой хэш-функции, которая гарантирует равномерное распределение идентификаторов объекта. Каждый узел p_i может определять идентификатор объекта R_j . Любой пользователь, присоединенный к Интернету, может выступать в качестве узла после установки соответствующего программного обеспечения. Набор узлов формирует в Интернете оверлейную сеть. Узел для пользователя выступает в качестве точки доступа. Каждый узел может предоставлять в сети место для хранения информации, участвовать в маршрутизации запросов в пределах сети, предоставлять различные услуги другим узлам сети и/или получать доступ к объектам.

Во второй главе исследованы существующие решения взаимного исключения и описаны разработанные алгоритмы бесконфликтного доступа (взаимного исключения) к общим объектам для динамических пиринговых сетей. Были рассмотрены следующие алгоритмы: Сигма-протокол (Sigma), сквозной протокол (E2E) и несквозной протокол (Non-E2E). Проведено аналитическое исследование их свойств. Результаты сформулированы в виде теорем (C – число оверлейных операций, с помощью которых сообщение может быть послано от одного узла к другому, n – число дубликатов объектов).

Теорема 2.1 (Сигма-протокол). Наибольшее число сообщений в Сигма-протоколе не превышает $5Cn$.

Теорема 2.2 (E2E). Наибольшее число сообщений в End-2-End-протоколе не превышает $5Cn$.

Теорема 2.3 (NonE2E). Наибольшее число сообщений в Non-End-2-End-протоколе не превышает $4Cn$.

Проведен SWOT-анализ (сильные и слабые стороны, возможности и угрозы) существующих решений взаимного исключения в пиринговых сетях.

Выявлено, что алгоритмы Sigma и E2E не удовлетворяют требованиям отказоустойчивости и упорядоченности запросов, а Non-E2E не удовлетворяет требованию упорядоченности запросов.

Были предложены два новых алгоритма взаимного исключения, свободные от перечисленных недостатков, - алгоритм взаимного исключения «Узел-Координатор» (У2К) и древовидной алгоритм взаимного исключения (ДАВИ).

Алгоритм У2К основан на организации набора узлов и передаче сообщений только между запросными узлами и координатором (вместо передачи сообщения между запросным узлом и всеми узлами, хранящими дубликаты, как это было в существующих решениях, см. рис. 1).

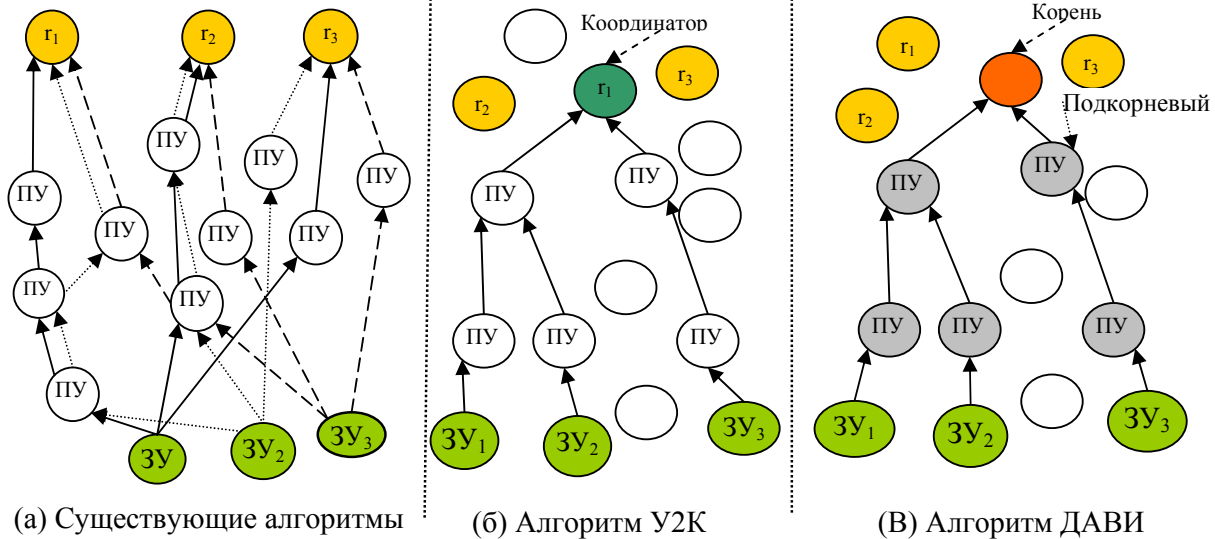


Рис.1. Примеры реализации трех запросных операций предложенного и существующих алгоритмов; r_i – узел, содержащий дубликат r_i ; ПУ – промежуточный узел (подкорневой узел); $ЗУ_i$ – i -ый запросный узел.

Блок-схема алгоритма приведена на рис. 2. Входными данными для алгоритма являются: R_j – объект с идентификатором j ; c_j – идентификатор клиента, владеющего объектом R_j ; t_{cj} – время получения полномочий (разрешения) на доступ к объекту R_j ; t_{ω} – временная метка, переменная состояния, всегда хранящаяся в p_i ; Q_j – полная очередь запросов, изначально пустая, Q_j используется только координатором; pQ_j – частичный перечень ожидающих запросов, который содержит сведения о предыдущих запросах узлов, упорядоченных по времени; pQ_j используется запросным узлом; s – идентификатор источника сообщения запросного узла; d – идентификатор узла доставки сообщения; Тип – тип сообщения; M – сообщения между узлами, которые обычно содержат: $\{s; d; t_{\omega}; pQ_j; (c_j, t_{cj}); Tun\}$.

Сначала запросные узлы посылают запрос координатору v_j (на рис. 1а – узел p_i), чтобы получить доступ к объекту R_j . В качестве координатора выступает узел сети с идентификатором, наиболее близким к хешированному идентификатору k_j объекта R_j . При получении запроса на использование объекта R_j координатор v_j посылает ответ запросному узлу (на рис. 1б – узел v_j , содержащий дубликат r_1). Ответ содержит информацию о текущем владельце, работающим с объектом, и очереди ожидающих узлов. После завершения

работы с объектом узел посылает координатору сообщение ОСВОБОЖДЕНИЕ. Когда координатор v_j получает сообщение ОСВОБОЖДЕНИЕ от текущего владельца объекта, он уведомляет об этом все запросные узлы набора, и, если только что завершивший работу с объектом узел производил модификацию объекта, сообщает им о новой версии объекта, а при необходимости тиражирует ее.

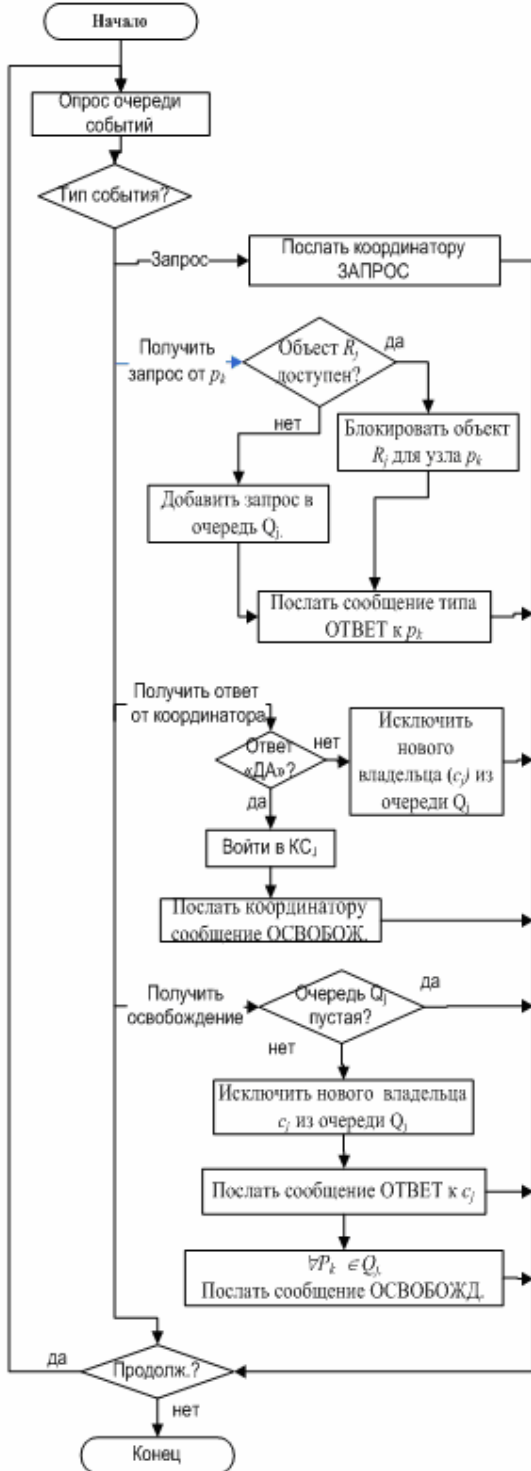


Рис. 2. Блок-схема алгоритма У2К

Полномочие доступа к объекту предоставляется при одновременном соблюдении следующих условий: (а) допустимо множество операций чтения в одно и то же время; (б) разрешается только одна операция записи в одно и то же время; (с) отсутствие (запрет) пар операций «чтение-запись» в одно и то же время.

Был проведен анализ производительности алгоритм У2К. Результаты сформулированы в виде нескольких теорем.

Теорема 2.4. *Наибольшее число оверлейных операций, реализуемых в алгоритме, не превышает $(m+2)C$ (m – число запросных узлов в очереди, C – число операций оверлея для каждого сообщения).*

Теорема 2.5. *Среднее время ответа \bar{T}_o в У2К алгоритме не больше $2\bar{t}$, где \bar{t} – среднее время прохождения сообщения от одного узла к другому.*

Теорема 2.6 *Среднее время задержки синхронизации \bar{T}_{zc} в У2К алгоритме составляет не более чем $2\bar{t} + \bar{t}_{kc}$, где \bar{t} – среднее время прохождения сообщения от одного узла к другому и \bar{t}_{kc} – среднее время работы с критической секцией.*

Теорема 2.7. *Максимальный параллелизм алгоритма – n операций типа чтения или одна операция типа запись, где n – число дубликатов объекта.*

Основная идея алгоритма

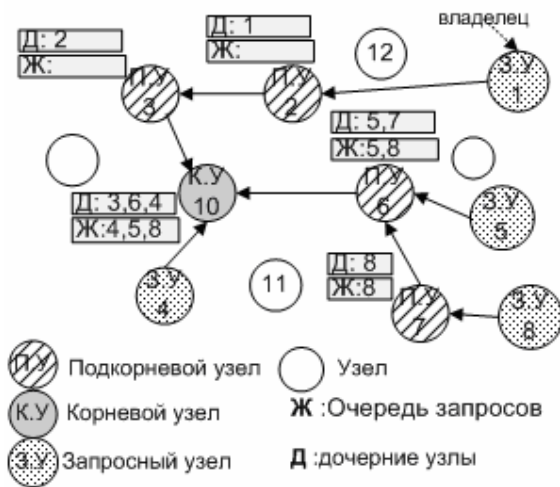


Рис. 3. Дерево узлов в ДАВИ

ДАВИ - построить дерево, образованное объединением только тех узлов, которые расположены на пути от запросного узла до корневого (рис. 1,3). В качестве корневого узла (КУ) дерева выбирается узел, идентификатор которого близок или равен идентификатору запрашиваемого объекта (хешированному ключу). Точки ветвления дерева – это промежуточные узлы (ПУ) на пути от запросных узлов (ЗУ) до корневого. Назовем их подкорневыми.

Любой узел, желающий получить доступ к объекту, должен соединиться с деревом. Тогда запросные узлы будут листьями дерева. Сеть может содержать множество деревьев – лес деревьев. Каждое дерево соответствует одному объекту. Подкорневые узлы поддерживают частичную очередь запросных узлов и список дочерних элементов. Дочерние элементы для узла p_i – это ближайшие к p_i узлы на пути следования запроса от запросного узла p_s к корневому. Очередь запросов содержит все запросы, которые проходят через этот узел, а в список очереди дочерних элементов включаются сведения только об их источниках сообщений (т.е. о предыдущих узлах пути). Корневой узел содержит полную очередь запросов Q_j и принимает решение относительно доступа к запрашиваемому объекту. Блок-схема алгоритма ДАВИ приведена на рис. 4.

Входящими данными алгоритма являются: Id_i - идентификатор узла i ; c_j - идентификатор клиента, владеющего объектом R_j ; Q_j - полная очередь запросов; pQ_j - перечень ожидающих запросов, который содержит сведения о запросах всех клиентов, упорядоченных по времени; LC_j - список дочерних элементов, соединяемых через данный узел; R_j - идентификатор объекта j ; t_{so} - временная отметка, переменная состояния, всегда хранящаяся в p_i ; s - идентификатор источника сообщения, запросного узла p_s ; u - идентификатор предыдущего узла p_{i-1} для промежуточного узла p_i на пути от запросного узла p_s к корневому; r - идентификатор следующего узла p_{i+1} для промежуточного узла p_i на пути от запросного узла p_s к корневому; тип сообщения; d - идентификатор узла p_d , которому посылается сообщения; M - сообщение между узлами, содержащее: $\{s, u, r, d, t_{so}, pQ_j, Tun\}$.

Любой узел, желающий получить доступ к объекту, посылает запрос только корневому узлу для того, чтобы минимизировать загрузку сети служебным трафиком и исключить нагрузку на другие дубликаты объекта. Все узлы дерева отслеживают состояние своего родителя (не вышел ли он из строя) для удовлетворения требований децентрализации и отказоустойчивости.



Рис. 4. Блок-схема алгоритма древовидного взаимного исключения

Когда любой дочерний узел обнаруживает, что его родитель выходит из строя, он сразу начинает операцию восстановления дерева. Дочерний узел узнает, что его родитель вышел из строя, когда от родителя перестают поступать тактовые сообщения. На всем протяжении маршрута промежуточные узлы хранят реестры дочерних элементов и запросных узлов. Только корневой узел принимает решение относительно доступа к объекту в соответствии со своей информацией. Кроме того, только он формирует в каждом такте полную очередь запросов, сообщая ее затем всем своим дочерним, запросным и промежуточным узлам. Если объект занят, корневой узел посылает ОТВЕТ «НЕТ» на запросы всех узлов. Когда объект свободен или когда корневым узлом (КУ) получено сообщение ОСВОБОЖДЕНИЕ, корневой узел посылает ОТВЕТ «ДА» следующему узлу, запрос которого стоит первым в очереди. ОТВЕТ «ДА» содержит разрешение на использование объекта. По его получении запросный узел получает доступ к совместно используемому объекту путем входа в его КС. Другие узлы (кроме узла, который получил разрешение на доступ к объекту) просто ждут получения разрешения от корневого узла. Когда узел заканчивает работу в КС, он посылает сообщение ОСВОБОЖДЕНИЕ корневому узлу, перед тем как снова

запросить доступ к какому-либо объекту. Очередь запросов любого узла сети содержит все запросы, которые проходят через этот узел, а список дочерних элементов содержит сведения о предыдущих узлах пути. Корневой узел содержит полную очередь запросов Q_j и принимает решение относительно доступа к запрашиваемому объекту.

Был проведен анализ производительности алгоритма ДАВИ. Результаты сформулированы в виде нескольких теорем.

Теорема 2.11. *Наибольшее число операций в ДАВИ не превышает 4С.*

Теорема 2.12. *Среднее время ответа \bar{T}_o в ДАВИ удовлетворяет неравенству $\bar{T}_o \leq 2\bar{t}$, где \bar{t} - среднее время прохождения сообщения от одного узла к другому.*

Теорема 2.13. *Средняя задержка синхронизации \bar{T}_{zc} в алгоритме ДАВИ составляет не более чем $2\bar{t} + \bar{t}_{kc}$, где \bar{t} - среднее время прохождения сообщения от одного узла к другому и \bar{t}_{kc} – среднее время работы в критической секции.*

Показано, как алгоритмы У2К и ДАВИ могут справляться с ошибками, возникающими в сети (любой узел сети может отказать, а узлы взаимодействуют посредством сообщений, передаваемых по ненадежным каналам). Для обоих алгоритмов рассмотрены такие ошибки как неисправность узла, присутствующего в настоящий момент в КС; неисправность любого узла в сети, а также неисправность в координаторе для У2К и неисправность в корневом узле или промежуточных подкорневых узлах для ДАВИ.

Показано, что алгоритмы У2К и ДАВИ обладает следующими свойствами: упорядоченности, аутентичности, децентрализации, меньшей степенью загрузки сети служебным трафиком, масштабируемости, отказоустойчивости.

В работе была доказана корректность алгоритмов У2К и ДАВИ, т.е. соблюдение ими требований безопасности, живучести и упорядоченности (справедливости).

В третьей главе выполнен обзор существующих алгоритмов тиражирования данных и обеспечения аутентичности дубликатов объекта в динамических пиринговых сетях. Рассмотрены два существующих алгоритма – «использование нескольких хэш-функций» и «наборы преемников и лист-наборы узлов». Показано, что оба алгоритма имеют общий недостаток - они не обеспечивают аутентичность дубликатов объекта, когда любой из узлов, который хранит дубликаты, выходит из системы и возвращается в систему, а за это время другой узел выполнил или выполняет операцию записи или изменения объекта. Кроме того, в алгоритме «использование нескольких хэш-функций» требуется инверсия хэш-функции, а алгоритм «наборы преемников и лист-наборы узлов» создает большую нагрузку на сеть, причем она является неравномерной, что приводит к узким местам.

С целью исправления отмеченных недостатков были предложены два новых алгоритма. На рис 5, 6 приведены их блок-схемы.

Первый предложенный алгоритм подходит для службы рассылки/получения данных на условиях подписки и имеет название –

алгоритм тиражирования с использованием опубликования информации о местоположении всех дубликатов у координатора (ТОДК).

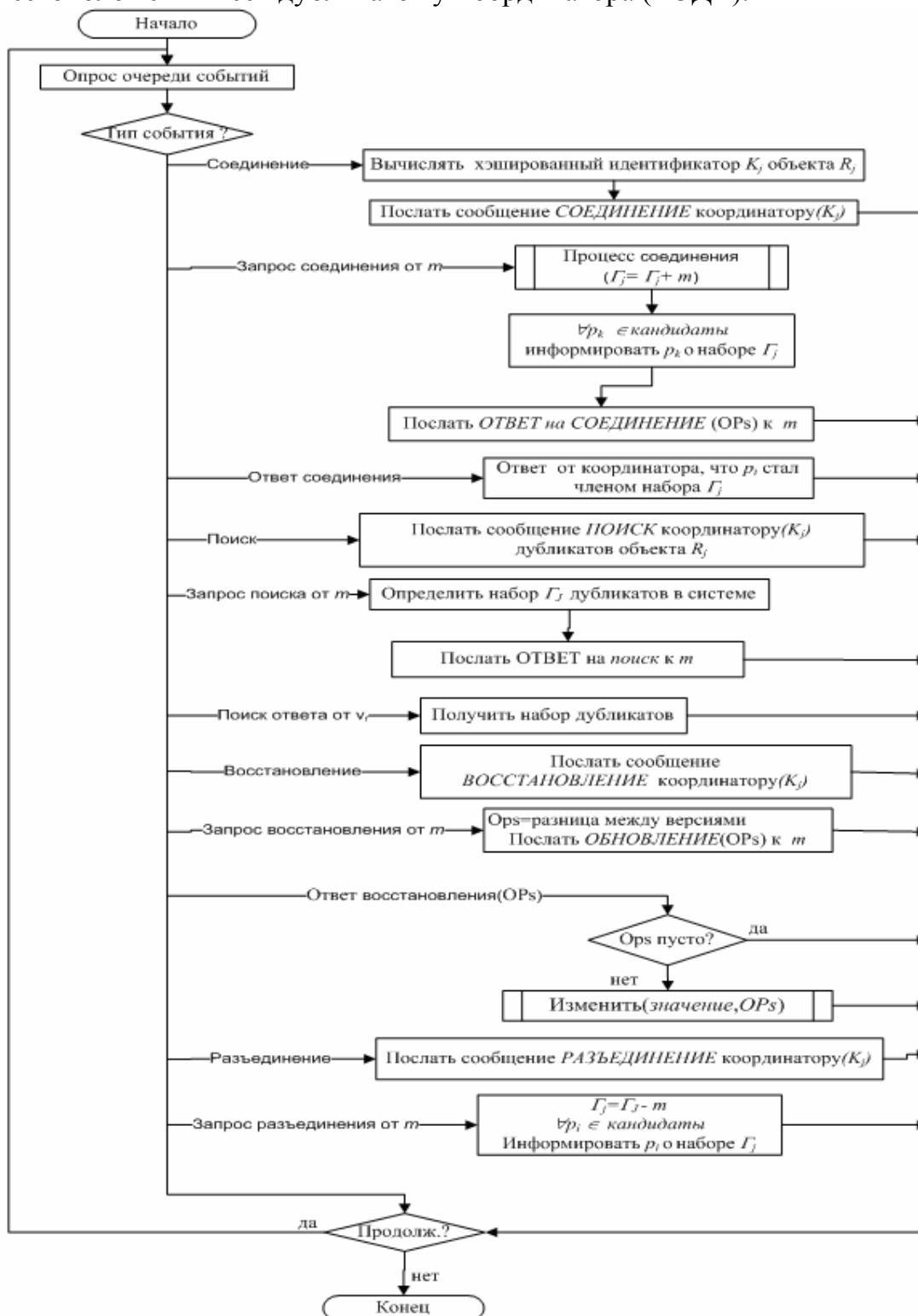
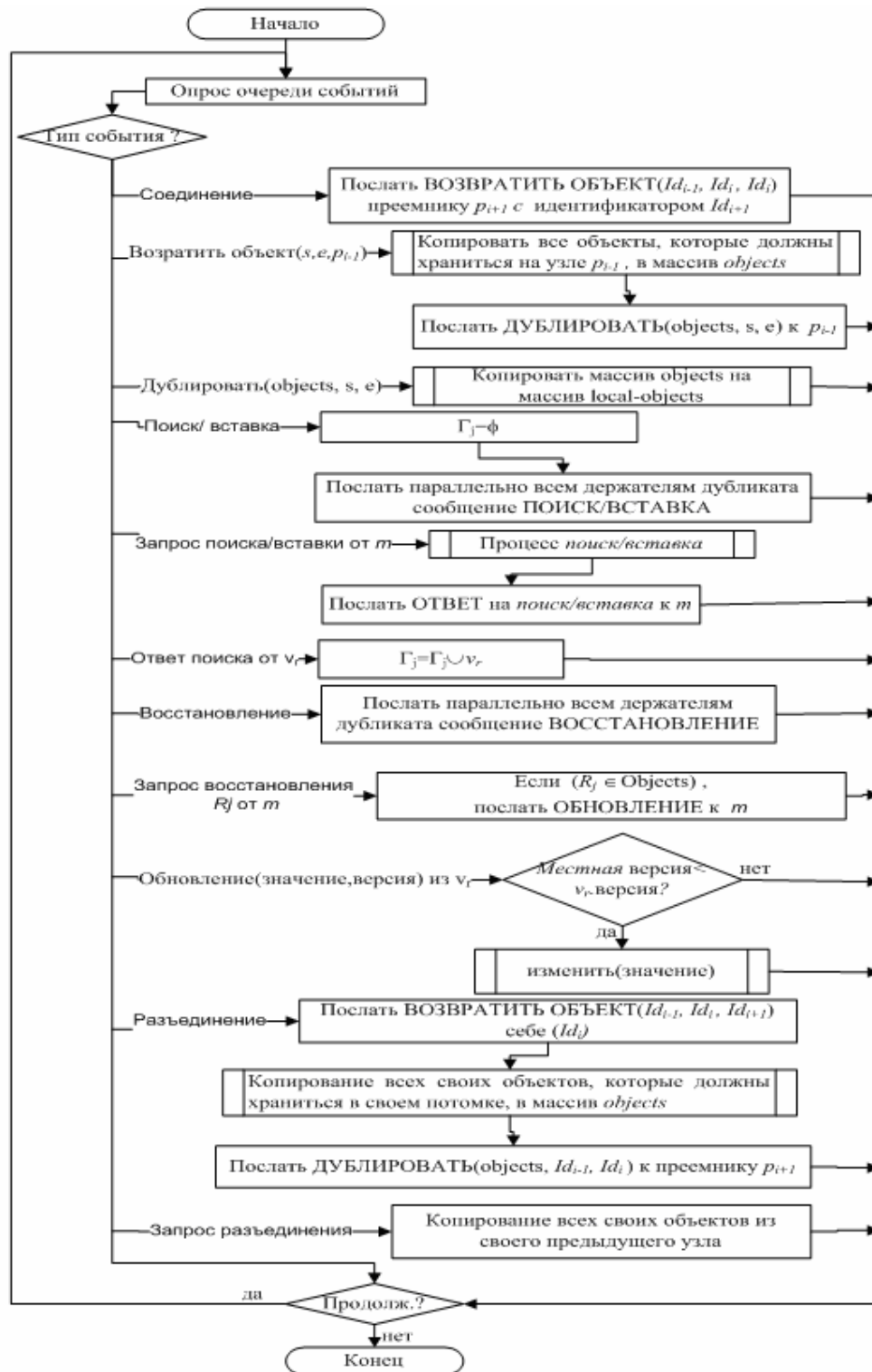


Рис.5. Блок-схема алгоритма ТОДК (узел p_i)

Второй алгоритм подходит для службы распространения данных и называется – алгоритм тиражирования с использованием сцепленной хэш-функции (ТСХФ)(см. рис. 6).

Рис.6. Блок-схема алгоритма ТСХФ (узел p_i)

Основная идея алгоритма **ТОДЖ** – продолжение идеи метода лист-набора. Вместо размещения дубликатов как в лист-наборе узлов предлагается публиковать информацию о владельцах (местоположении дубликатов) в координаторе, которым является узел с идентификатором, наиболее близким к идентификатору объекта.

Входящими данными для алгоритма ТОДК являются: R_j – объект с идентификатором Id_j ; K_j – идентификатор (ключ) объекта R_j (узел сети с идентификатором, наиболее близким к идентификатору K_j объекта R_j , является координатором); OPs – список модификаций между двумя версиями дубликатов объекта; Γ_j – набор узлов, содержащих дубликаты объекта R_j ; s – идентификатор источника сообщения запросного узла; d – идентификатор узла доставки сообщения; Tun – тип сообщения; M – сообщения между узлами, которые содержат: $\{s; d; Tun\}$.

Идея алгоритма ТСХФ заключается в том, что для определения местоположения первого дубликата производится рекуррентным хешированием имени объекта с помощью хэш-функции $sha-1$. Местоположение (ключ) для каждого дубликата вычисляется путем хеширования ключа предыдущего дубликата. Формально это означает, что если мы имеем R_i – объект “X”, то хэш-функция первого дубликата будет $K_{i1} = hash(“X”)$, а ключи для всех остальных дубликатов будут вычисляться следующим образом: $K_{ji} = hash(K_{j,i-1})$, где $i=2, \dots, n$, n – коэффициент дублирования.

Входящими данными для алгоритма ТСХФ являются: R_j – объект с идентификатором Id_j ; Γ_j – набор узлов, содержащих дубликаты объекта R_j ; $Objects(O)$ – двумерный (f, n) -массив, где f – это количество объектов, n – количество дубликатов; s – идентификатор источника сообщения запросного узла; d – идентификатор узла доставки сообщения; Tun – тип сообщения; M – сообщения между узлами, которые содержат: $\{s; d; O; Tun\}$.

Был выполнен анализ сложности алгоритмов тиражирования (существующих и предложенных) по числу сообщений (n – степень тиражирования дубликатов объекта, C – максимальное число операций оверлея для каждого сообщения, $C^*=1$, т.к. сообщения между узлом и его близкими узлами требуется 1 операция). Результаты приведены в табл. 1.

Таблица 1

Сравнение алгоритмов по сложности операций

Алгоритм \ Операция	Набор различных хэш-функций	Наборы преемников и лист-наборы	ТОДК	ТСХФ
Поиск / вставка	$\leq Cn$	$\leq C(n+1)$	$\leq C$	$\leq 2C$
Соединение / разъединение	$\leq C$	$\leq Cn$	$\leq nC^* + C$	$\leq C^*$
Восстановление	-	-	$\leq 2C$	$\leq C$
Обновление	-	-	$\leq C(n+1)$	$\leq Cn$
Итого (все операция)	$\leq C(n+1)$	$\leq C(2n+1)$	$\leq nC^* + C(n+5)$	$\leq C(n+3) + C^*$
(без новых операций)	$\leq C(n+1)$	$\leq C(2n+1)$	$\leq nC^* + 2C$	$\leq 2C + C^*$

Из нее видно, что производительность предлагаемых алгоритмов по числу всех операций лучше, чем у существующих алгоритмов. С другой стороны, предложенные алгоритмы содержат новые операции “восстановление” и “обновление”. Эти операции отсутствуют в существующих методах, поскольку считается, что объект не изменяемый.

Для достижения n степеней тиражирования во всех алгоритмах требуется тиражировать новый дубликат вместо каждого выходящего из системы. В табл. 2 приведено сравнение алгоритмов по количеству операции тиражирования нового дубликата. Видно, что алгоритм ТСХФ всегда сохраняет степень тиражирования, как и существующие методы, а алгоритму ТСХФ не требуется дополнительных шагов для сохранения степени тиражирования.

Таблица 2

Сравнение алгоритмов по количеству операций на тиражирование ($C^*=1$, т.к. для сообщения между узлом и его ближайшим соседом требуется 1 операция оверлея)

Набор различных хэш-функции	Наборы преемников и лист-наборы	ТОДК	ТСХФ
$\leq C^{(1)}$	$\leq Cn$	$\leq nC^* + C$	0

(1) Требуется инверсия хэш-функции

Четвертая глава посвящена описанию созданного алгоритмического и программного обеспечения, пригодного для имитации функционирования и исследования качества разработанных алгоритмов.

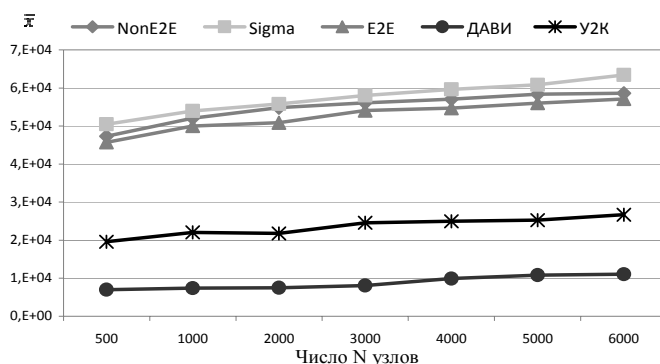
В качестве рабочей версии выбрана полностью децентрализованная пиринговая модель Pastry и её реализация FreePastry. В качестве языка программирования был выбран язык Java, так как именно он используется при реализации FreePastry.

Описывается структура программного пакета, разработанного для исследования P2P сетей. Приводится описание разработанных подсистем редактирования, а также подсистемы прогона модели.

Пятая глава. Описываются схемы экспериментов и приведены результаты их проведенная с указанием достоверности. Сравнение проводится по производительности алгоритмов (число сообщений, время ответа, время задержки синхронизации). На рис. 7–9 показаны примеры результатов проведенных исследований. Также были проведены исследования и сравнения алгоритмов по таким показателям, как пропускная способность, текучесть. Из сравнения результатов исследований следует, что предложенные алгоритмы имеют лучшие характеристики производительности.

Шестая глава. Описывается созданный программный пакет для реальной работы в P2P сетях согласно требованиям, сформулированным в первой главе, и по предложенным во второй и третьей главах алгоритмам.

Созданный программный пакет был использован в компаниях, занимающихся разработкой и внедрением программных систем, – ООО «Инфо-Стрим» и ООО «Идель» на задачах исследования моделей проектируемых приложений, что подтверждается документами, приведенными в приложении к диссертации. Разработанный пакет применяется также в учебном процессе на факультете АВТ НГТУ, кафедра вычислительной техники.



Отношение $\hat{\sigma} / \bar{x}$
 NON-E2E (1.1-2.0)%
 Sigma (0.7-1.7)%
 E2E (1.4-2.4)%
 У2К (2.1-5.0)%
 ДАВИ (4.5-6.6)%

Рис.7. Сравнение предложенных и существующих алгоритмов по среднему количеству сообщений \bar{x} и масштабируемости; $\hat{\sigma}$ - оценка среднеквадратического отклонения количества сообщений

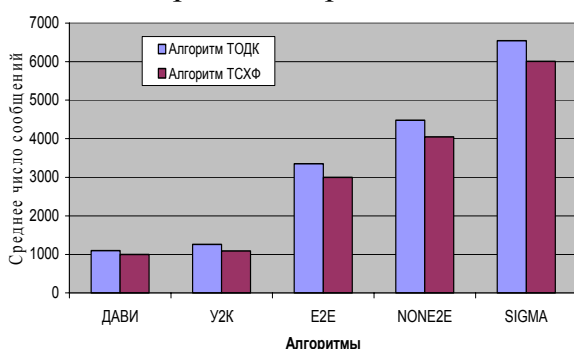


Рис.8. Влияние на нагрузку сети типа алгоритмов тиражирования

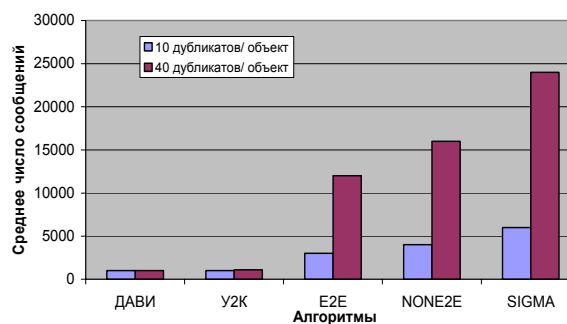


Рис 9. Влияние на нагрузку сети количества дубликатов

В заключении приведены основные результаты и выводы по диссертационной работе.

В приложении представлены копии документов об использовании результатов работы и дипломов участия конференций.

Основные выводы и результаты работы

Таким образом, в диссертационной работе, согласно поставленной цели, было разработано модельное, алгоритмическое и программное обеспечение, пригодное для организации эффективного множественного бесконфликтного доступа к общим изменяемым объектам в динамичных децентрализованных пиринговых сетях, имеющее важное значение в различных областях знаний и народного хозяйства, где применяются подобные сети. А именно.

1. Выполнен обзор децентрализованных файловых пиринговых сетей и дана их классификация. Выявлены требующие решения задачи, характерные для пиринговых сетей. Это задачи взаимного исключения, тиражирования и обеспечения аутентичности дубликатов объектов. Формализована системная модель класса рассматриваемых сетей.

2. Проведена систематизация и дан анализ существующих решений взаимного исключения в пиринговых сетях, рассмотрены их сильные и слабые стороны. Предложены и исследованы алгоритмы взаимного исключения «узел-координатор» (У2К) и древовидный алгоритм взаимного исключения (ДАВИ),

детально описаны их основы и свойства. Рассмотрена возможность обнаружения отказов в сети с помощью предложенных алгоритмов. Проведено аналитическое исследование свойств алгоритмов и показано, что разработанные алгоритмы являются более эффективными по сравнению с существующими.

3. Проведена систематизация и дан анализ существующих методов тиражирования объектов, показаны их недостатки. Разработаны алгоритмы тиражирования ТОДК и ТСХФ, детально описаны их основы, рассмотрены преимущества, показано, что они эффективнее существующих методов и свободны от их недостатков. Решена задача обеспечения аутентичности динамично получаемых дубликатов непредсказуемо изменяемых объектов в этих алгоритмах.

4. Разработан, реализован, апробирован и внедрен программный пакет для исследования предложенных алгоритмов. С его помощью проведено имитационное исследование предложенных алгоритмов, сравнение их с существующими. Показано, что предложенные алгоритмы по всем показателям превосходят известные алгоритмы.

5. Проведена апробация разработанного программного пакета для задач исследования проектируемых вычислительных сетей.

Основные публикации по теме диссертации

Публикации, входящие в Перечень ВАК:

1. Губарев В. В., **Обейдат А. А.** Алгоритм взаимного исключения одновременного доступа пользователей к общим ресурсам в пиринговых системах // Научный Вестник НГТУ, 2009. – № 2 (35). – С. 69–85.

2. Губарев В. В., **Обейдат А. А.** Алгоритм взаимного исключения в пиринговых системах // Известия Томского политехнического университета. – 2009. – Т. 314. – № 5. – С. 32–36.

Другие публикации:

3. **Обейдат А. А.**, Губарев В.В. Обзор алгоритмов распределенного взаимного исключения в динамических пиринговых системах // Сб. научных трудов НГТУ, Новосибирск: НГТУ, 2007. – № 2 (48). – С. 63–68.

4. **Обейдат А. А.**, Губарев В. В. Систематизация и SWOT-анализ протоколов взаимного исключения в пиринговых системах // Сб. научных трудов НГТУ, Новосибирск: НГТУ, 2008. – № 2 (52). – С. 59–64.

5. **Obeidat A. A.**, Gubarev V. V. Developing End-To-End Mutual Exclusion Protocol in Peer-to-Peer Systems (Разработка Сквозного Протокола Взаимного Исключения в Пиринговых Системах) // The Third International Forum on Strategic Technologies (IFOST), Novosibirsk, NSTU, 2008. – P. 282–286.

6. **Obeidat A. A.**, Gubarev V. V., Al-yousef A. A. Decentralized and Fair Mutual Exclusion Protocol in Peer-to-Peer Systems. (Протокол Взаимного исключения с обеспечением децентрализации и упорядоченности в пиринговых системах) // The 10th International Workshop on Computer Science and Information Technologies (CSIT), Antalya, Turkey, Ufa State Aviation Technical University, 2008. — P. 11–16.

7. Губарев В.В., Гужов В.И., Джо К.Х., **Обейдат А.А.** Клиентские информационно-вычислительные средства как объекты изучения // Материалы НПК «Инновации в условиях развития информационно-коммуникационных технологий» (Инфо-2009): Качество – Безопасность – Диагностика. – М.: Изд-во МИЭМ.–2009. – С. 105–107.

8. **Обейдат А. А.** Алгоритм узел – узел взаимного исключения в пиринговых системах// Всероссийская научная конференция студентов, аспирантов и молодых ученых "Наука. Технологии. Инновации". – Новосибирск: Изд-во НГТУ, 2008. – Ч. 1. – С. 202–223.

9. **Обейдат А. А.** Голосование лидера в пиринговых системах // Современные проблемы информатизации в проектировании и информационных системах: Сб. трудов. – Вып. 14. – Воронеж: «Научная книга», 2009. – С. 520–522.

10. **Обейдат А. А.** Управление доступом к объектам изменчивых данных в пиринговых системах // XV международной научно-технической конференции студентов и аспирантов «РАДИОЭЛЕКТРОНИКА, ЭЛЕКТРОТЕХНИКА И ЭНЕРГЕТИКА».– Москва, Издательский дом МЭИ.– 2009. – С. 314–315.

11. **Обейдат А. А.**, Губарев В.В. Назначение лидера в пиринговых системах // Материалы III Всероссийской конференции «Винеровские чтения 2009» [Электронный ресурс]. – Иркутск: ГОУ ВПО ИрГТУ, 2009. – № 24.

12. **Obeidat A. A.**, Gubarev V. V. Leader election in peer-to-peer systems (Выбор лидера в пиринговых системах) // International Siberian Conference on Control and Communications (SIBCON-2009), Томск: ТГУ, 2009. – P.25–32.

13. **Обейдат А. А.** Свойства алгоритма «узел-координатор» взаимного исключения в динамичных пиринговых системах // II Всероссийская научно-практическая конференция «Научная инициатива иностранных студентов и аспирантов российских вузов».– Томск: ТПУ, 2009.– С. 142–145.

14. **Обейдат А. А.** Обеспечение аутентичности дубликатов ресурса в пиринговых системах // II Всероссийская научно-практическая конференция «Научная инициатива иностранных студентов и аспирантов российских вузов».– Томск. ТПУ, 2009.– С. 137–141.

15. **Obeidat A.A.** Replication of Mutable Objects in Peer-to-Peer Systems (Тиражирование изменяемых объектов в пиринговых системах) // The Third International Forum on Strategic Technologies (IFOST 2009), Ho Chi Minh city, Vietnam: Ho Chi Minh City University of Technology.–2009. –P.171 – 175.