

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ  
УНИВЕРСИТЕТ»

На правах рукописи



**ХУССЕЙН Али**

**СПЕЦИАЛЬНОЕ МАТЕМАТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ  
ПРОЦЕССОВ УПРАВЛЕНИЯ ЭНЕРГОЭФФЕКТИВНЫМ  
ПЕРЕМЕЩЕНИЕМ ДАННЫХ В БЕСПРОВОДНЫХ  
ИНФОРМАЦИОННЫХ СИСТЕМАХ**

Специальность: 2.3.5. Математическое и программное обеспечение  
вычислительных систем, комплексов и  
компьютерных сетей

**Диссертация**

на соискание учёной степени кандидата технических наук

Научный руководитель:  
д.т.н., доцент Рындин Никита Адександрович

Воронеж – 2026

## Оглавление

<b>Введение .....</b>	<b>5</b>
<b>1. Особенности разработки специального энергоэффективного программного обеспечения перемещения потоков данных в беспроводных информационных системах .....</b>	<b>15</b>
1.1. Проблемы управления беспроводными сенсорными сетями с точки зрения их энергоэффективности.....	15
1.2. Повышение энергоэффективности управления потоками данных в беспроводных сенсорных сетях.....	24
1.2.1. Поддержание работоспособности сети в течение длительного периода времени.....	25
1.2.2. Состояние проблемы.....	26
1.2.3. К задаче управления потоками данных для повышения энергоэффективности.....	27
1.2.4. Направление исследования .....	28
1.2.5. Теория скрытых марковских моделей .....	29
1.2.6. Основные переменные.....	29
1.2.7. Параметры модели .....	30
1.3. Подходы к разработке математического и программного обеспечения управления данными в специальных задачах.....	31
1.3.1. Проблема обнаружения аномальных данных в системе автоматизации измерения мощности .....	31
1.3.2. Система динамического мониторинга потерь больших объемов данных в мобильной сети на базе Интернета вещей.....	33
1.4. Постановка задач работы.....	36
<b>2. Интеграция данных для беспроводных сенсорных сетей с позиций модели энергоэффективности .....</b>	<b>39</b>
2.1. Проблема проектирования энергоэффективных маршрутов в беспроводных сенсорных сетях.....	39
2.2. Усовершенствованная энергоэффективная технология передачи данных в сенсорных сетях.....	39
2.3. Метод исследования.....	45
2.4. Результаты экспериментов и их анализ .....	47
2.5. Выводы .....	61
Литература к главе 2 .....	63
<b>3. Повышение энергоэффективности протокола управления потоками данных со скрытой марковской моделью в беспроводных сенсорных сетях.....</b>	<b>66</b>
3.1. Предлагаемая архитектура .....	66

3.1.1. Общая схема системы.....	66
3.1.2. Примерная топология сети.....	68
3.1.3. Эмпирическая оценка исходных параметров НММ.....	70
3.2. Стохастическое решение о маршруте.....	71
3.2.1. Предложенная полностью связанная модель.....	71
3.2.2. Обучение стохастической модели управления потоками данных.....	74
3.2.3. Процесс принятия решения об управлении потоками данных.....	74
3.3. Анализ результатов моделирования.....	76
3.3.1. Настройка параметров.....	76
3.3.2. Оценка вероятности управления потоками данных.....	77
3.4. Сравнение производительности.....	81
3.5. Особенности результатов.....	87
3.6. Расчет параметров стохастического маршрута (пример).....	88
3.9. Прямой расчет вероятностей.....	91
3.7. Выводы к главе 3.....	94
Литература к главе 3.....	95

<b>4. Математическое и программное обеспечение специального управления данными в прикладных задачах.....</b>	<b>100</b>
4.1. Обнаружение аномальных данных в электроэнергетической системе.....	100
4.1.1. Метод обнаружения аномальных данных в системе автоматизации измерения электроэнергии.....	100
4.1.2. Обработка данных системы автоматизации учета электроэнергии.....	103
4.1.3. Обнаружение признаков аномальных данных.....	105
4.1.4. Реализация обнаружения аномальных данных.....	106
4.2. Экспериментальное исследование метода обнаружения аномальных данных в электроэнергетической системе.....	109
4.2.1. Содержание эксперимента.....	110
4.2.2. Подготовка к эксперименту.....	110
4.2.3. Результаты эксперимента.....	112
4.3. Система динамического мониторинга потерь больших объемов данных в мобильной сети на базе Интернета вещей.....	116
4.3.1. Разработка архитектуры системы динамического мониторинга потерь больших объемов данных в мобильной сети на базе Интернета вещей.....	116
4.3.2. Разработка программного обеспечения системы динамического мониторинга потерь больших объемов данных в мобильной сети.....	121
4.4. Экспериментальное исследование системы динамического	

мониторинга потерь больших данных в мобильной сети.....	125
4.4.1. Архитектура платформы эксперимента.....	126
4.4.2. Результаты экспериментов и их анализ.....	128
4.5. Выводы к главе 4.....	134
Литература к главе 4.....	136
<b>Заключение.....</b>	<b>139</b>
<b>Список использованных источников.....</b>	<b>141</b>

## **Введение**

**Актуальность темы.** За последние годы беспроводные информационные системы и сенсорные данные в них стали существенно важным объектом в связи с интеллектуализацией их обработки и широким распространением концепции и технологий «Умный дом». Обработка сенсорных данных от множества активных устройств и многочисленных датчиков объединена с облачными сервисами, что породило проблему поиска архитектурных и структурных решений, направленных на снижение энергопотребления всей системы в целом и подсистемы управления данными в частности. Актуальной является задача управления энергоэффективного перемещения данных в распределенных системах. Большой вклад в разработку методов и средств управления данными внесли Андреев А.Н., Заболотный И.Г., Ковалев И.В., Кравец О.Я., Cobb M., Fu X., Kalkha H., Liu S. и другие.

Одной из актуальных предметных областей задач управления данными является оптимальное перемещение данных в беспроводных информационных системах. Интерес представляет и снижение энергопотребления за счет удачной кластеризации данных. С точки зрения интуитивного понимания, чем лучше мониторинг данных, тем меньше их потерь может происходить в информационной системе.

Таким образом, актуальность темы диссертационного исследования продиктована необходимостью разработки специальных средств математического обеспечения процессов управления энергоэффективным перемещением данных в беспроводных информационных системах на основе динамической реконфигурации и кластеризации.

Тематика диссертационной работы соответствует научному направлению ФГБОУ ВО «Воронежский государственный технический университет» «Вычислительные комплексы и проблемно-ориентированные системы управления».

**Целью работы** является разработка моделей и алгоритмов работы средств специального математического обеспечения процессов управления энергоэффективным перемещением данных в беспроводных информационных системах.

**Задачи исследования.** Для достижения поставленной цели необходимо решить следующие задачи:

1. Проанализировать проблематику моделирования и алгоритмизации процессов управления перемещением потоков данных в беспроводных информационных системах.

2. Предложить модификацию стратегии управления данными, позволяющую повысить энергоэффективность маршрутизации в крупномасштабных беспроводных сенсорных сетях на основе агентного подхода и фреймворка динамической реконфигурации для технологии слияния на основе дерева.

3. Разработать модель снижения энергопотребления за счет использования вероятностных методов и инструментов кластеризации.

4. Разработать алгоритм повышения показателей энергоэффективной связи за счет сокращения общего времени, необходимого для завершения процесса.

5. Разработать структуру программного комплекса динамического мониторинга потерь больших объемов данных с многоуровневым сбором информации (уровни кластера, производительности узлов и работы заданий), повышающую максимальную точность мониторинга.

**Объект исследования:** процессы управления энергоэффективным перемещением потоков данных в беспроводных информационных системах.

**Предмет исследования:** средства математического и программного управления процессами анализа моделей и алгоритмов управления перемещением потоков данных в беспроводных информационных системах на

основе агентного подхода и фреймворка динамической реконфигурации для технологии слияния.

**Методы исследования.** При решении поставленных в диссертации задач использовались методы теории вероятностей, теории принятия решений, методы кластерного анализа, а также методы объектно-ориентированного программирования.

**Тематика работы** соответствует следующим пунктам паспорта специальности 2.3.5 «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей»: п. 4. «Интеллектуальные системы машинного обучения, управления базами данных и знаний, инструментальные средства разработки цифровых продуктов»; п. 9. «Модели, методы, алгоритмы, облачные технологии и программная инфраструктура организации глобально распределенной обработки данных».

**Научная новизна работы.** В диссертации получены следующие результаты, характеризующиеся научной новизной:

1. Модификация стратегии управления данными, отличающаяся применением агентного подхода и фреймворка динамической реконфигурации для технологии слияния на основе дерева, позволяющая повысить энергоэффективность маршрутизации в крупномасштабных беспроводных сенсорных сетях.

2. Модель снижения энергопотребления процессов перемещения данных в беспроводных информационных системах, отличающаяся использованием скрытых марковских моделей и технологий кластеризации и обеспечивающая повышение эффективности узла слияния по мере увеличения количества кластеризуемых узлов, а также их индивидуальных уровней энергии.

3. Алгоритм улучшения показателей энергоэффективной связи объектов беспроводных информационных систем, отличающийся автоматическим выбором гетерогенных факторов ранжирования узла слияния – кан-

дидата с наибольшей реализуемостью, и обеспечивающий сокращение общего времени, необходимого для завершения процесса кластеризации.

4. Гибридный итерационный алгоритм поиска аномальных данных, сочетающий применение матрицы сходства аномальных признаков данных и кластерный анализ с использованием обучения iForest, обеспечивающий улучшение точности обнаружения в среднем на 22%.

5. Структура программного комплекса динамического мониторинга потерь больших объемов данных с многоуровневым сбором информации (уровни кластера, производительности узлов и работы заданий), обеспечивающая максимальную точность мониторинга до 98%.

**Теоретическая и практическая значимость исследования** заключается в развитии средств разработки специального математического и программного обеспечения процессов управления перемещением данных в беспроводных информационных системах на основе агентного подхода и фреймворка динамической реконфигурации для технологии слияния.

Теоретические результаты работы могут быть использованы в проектных и научно-исследовательских организациях, занимающихся проектированием энергоэффективных программных систем управления потоками данных.

#### **Положения, выносимые на защиту**

1. Модификация стратегии управления данными позволяет повысить энергоэффективность маршрутизации в крупномасштабных беспроводных сенсорных сетях.

2. Модель снижения энергопотребления процессов перемещения данных в беспроводных информационных системах обеспечивает повышение эффективности узла слияния по мере увеличения количества кластеризуемых узлов.

3. Алгоритм улучшения показателей энергоэффективной связи объектов беспроводных информационных систем обеспечивает сокращение

общего времени, необходимого для завершения процесса кластеризации.

4. Гибридный итерационный алгоритм поиска аномальных данных на основе кластерного анализа с использованием алгоритма обучения iForest обеспечивает улучшение точности обнаружения в среднем на 22%.

5. Структура программного комплекса динамического мониторинга потерь больших объемов данных обеспечивает максимальную точность мониторинга до 98%.

**Результаты внедрения.** Основные результаты внедрены в АО «Научно-исследовательский институт вычислительных комплексов им. М. А. Карцева» (г. Москва) при проектировании систем управления перемещением потоков данных в беспроводных информационных системах, в учебный процесс Воронежского государственного технического университета в рамках дисциплин: «Вычислительные машины, системы и сети», «Информационные сети и телекоммуникационные технологии», а также в рамках курсового и дипломного проектирования.

**Апробация работы.** Основные положения диссертационной работы докладывались и обсуждались на следующих конференциях: Международной научно-практической конференции «Интеллектуальные информационные системы» (Воронеж, 2021); Международной молодежной научной школе «Оптимизация и моделирование в автоматизированных системах» (Воронеж, 2022); VI Всероссийской научно-практической конференции «Информационные технологии в экономике и управлении» (Махачкала, 2024); XXX-th - XXXI-th International Open Science Conference «Modern informatization problems» (Yelm, WA, USA, 2025, 2026), а также на научных семинарах кафедры искусственного интеллекта и цифровых технологий ВГТУ (2023-2026 гг.).

Обоснованность и достоверность результатов работы обусловлена корректным использованием теоретических методов исследования и подтверждена сравнительным анализом данных вычислительных и натуральных

экспериментов.

**Публикации.** По результатам диссертационного исследования опубликовано 16 научных работ, в том числе 6 – в изданиях, рекомендованных ВАК РФ (из них 1 – в издании, индексируемом в WoS и одно свидетельство о регистрации программы для ЭВМ). В работах, опубликованных в соавторстве и приведенных в конце автореферата, лично автором получены следующие результаты: [4, 12, 14, 16] - модификация стратегии управления данными, отличающаяся применением агентного подхода и фреймворка динамической реконфигурации для технологии слияния на основе дерева; [1, 7, 13] - модель снижения энергопотребления процессов перемещения данных в беспроводных информационных системах, отличающаяся использованием скрытых марковских моделей и технологий кластеризации; [5, 8] - алгоритм улучшения показателей энергоэффективной связи объектов беспроводных информационных систем, отличающийся автоматическим выбором гетерогенных факторов ранжирования узла слияния – кандидата с наибольшей реализуемостью; [2] - гибридный итерационный алгоритм поиска аномальных данных, сочетающий применение матрицы сходства аномальных признаков данных и кластерный анализ с использованием обучения iForest, обеспечивающий улучшение точности обнаружения в среднем на 22%; [3, 6] - структура программного комплекса динамического мониторинга потерь больших объемов данных с многоуровневым сбором информации.

**Структура и объем работы.** Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы из 105 наименований. Работа изложена на 152 страницах.

### **Основное содержание работы**

**Во введении** обоснована актуальность исследования, сформулированы его цель и задачи, научная новизна и практическая значимость полу-

ченных результатов, приведены сведения об апробации и внедрении работы.

**В первой главе** исследуются проблемы управления большими данными облачных сервисов на основе многостадийных алгоритмов и динамического перераспределения данных. Отмечено, что повысить эффективность такого управления можно путем применения алгоритма расширения хранилища больших сервисов в различных облачных зонах, алгоритм компоновки больших сервисов, отличающийся учетом качества данных и определением набора формальных понятий, выбора архитектуры динамической системы распределения данных, интеграции принятия решений в большие данные, архитектуры программной системы оптимизации больших данных от датчиков в Интернете вещей. Результат анализа потребовал формализации данных задач, а также алгоритмизации их решения с учетом особенностей. Сформулирована цель и задачи исследования.

**Вторая глава** посвящена задаче рациональной компоновки больших облачных сервисов. В качестве модели абстрагирования больших данных и сокрытия их сложности в больших сервисах (BS) рассматриваются как управляемая интеграция массивной, сложной серии разнородных сервисов, ориентированных на большие данные. Такая сложная и масштабная сервисная экосистема способна обрабатывать огромные объемы данных и предлагать их в качестве сервисов по требованию клиентов.

Будем считать, что качество BS зависит не только от традиционных параметров QoS (например, надежности, доступности, стоимости, безопасности и т.д.), но и от качества источников данных (QoD), используемых компонентами BS. Фактически, оценка полноты, точности и своевременности этих источников данных является важным шагом для принятия решения о способности сервиса-кандидата участвовать в больших облачных сервисах (BSCo), даже если он отличается высоким качеством обслуживания.

Представлен алгоритм расширения хранилища больших сервисов в различных облачных зонах, отличающийся представлением в виде семейства решеток и использованием сходства по Жакарду экземпляров сервисов и источников данных и обеспечивающий оценку близости формальных концепций, которые объединяют эти сервисы и источники данных.

**Третья глава** посвящена алгоритмизации управления распределением больших данных интернета вещей.

В соответствии с массовостью, пространственно–временной корреляцией, дисбалансом доступа и постоянной изменчивостью информации в Интернете вещей, для адаптации к ней необходим механизм распределения данных во временной области.

На этапе инициализации узел управления распределяет точки данных по каждому узлу хранилища в соответствии с алгоритмом статического распределения и инициализирует глобальную таблицу распределения данных. На этапе эксплуатации узел хранения запускает модуль адаптивной обратной связи по нагрузке и регулирует временную область обратной связи в режиме реального времени в соответствии с нагрузкой, создаваемой обновлением данных. Узел управления запускает модуль динамического распределения данных, отслеживает информацию о загрузке, передаваемую каждым узлом хранения, в режиме реального времени и регулирует распределение данных по каждому узлу хранения в режиме реального времени. Порт сбора данных синхронизируется с узлом управления для обеспечения согласованности глобального обновления распределения данных.

Ключом к регулированию нагрузки на перегруженный узел хранения является сокращение миграции данных. В системе стоимость миграции каждой точки данных одинакова, поэтому чем меньше количество перенесенных точек данных, тем ниже стоимость процесса настройки. В соответствии с «жадной» идеей, выберем точку данных в перегруженном узле

хранения с максимальной нагрузкой и изменим точку на узел хранения с минимальной нагрузкой после того, как он перенесет нагрузку..

Предложена архитектура динамической системы распределения данных, отличающаяся использованием «жадного» алгоритма сокращения миграции данных с динамическим выбором точки данных в перегруженном узле хранения с максимальной нагрузкой и обеспечивающий регулирование распределения данных по каждому узлу хранения в режиме реального времени..

**В главе 4** проанализированы особенности интеграции больших данных в системы принятия решений.

Основной вопрос, который рассматривается, заключается в следующем: какие аспекты следует учитывать при разработке проекта с использованием больших данных, направленного на решение проблемы принятия решений в организации?

Модель BD-Da использует на три уровня, которые необходимо учитывать при разработке проекта с использованием больших данных, направленного на решение проблемы принятия решений в организации. Эти уровни - уровень данных, анализ данных, принятие решений.

Модель BD-Da представляет концепцию больших данных, основанную на шести концепциях, а именно: наборы данных с новыми характеристиками, жизненный цикл анализа данных, технологии, аналитические методы, понимание и принятие решений.

Расширенная модель BD-Da учитывает источники данных, что позволяет определять и представлять потенциальные альтернативы источников, предоставляющих данные.

Представлена графическая модель BD-Da интеграции принятия решений в большие данные, отличающаяся использованием наборов данных с новыми характеристиками, жизненного цикла анализа данных, технологий, аналитических методов, понимания и принятия решений и обеспечи-

вающая выделение трех уровней больших данных, которые необходимо учитывать при разработке проекта больших данных: данных, анализа и принятия решений.

**Пятая глава** описывает программные проекты управления большими данными.

Для повышения стабильности операций интеллектуального анализа гетерогенных больших данных в среде сложных атрибутов, таких как анализ и очистка данных, разработан алгоритм интеллектуальной кластеризации гетерогенных больших данных. Метод очистки данных применяется для очистки пространства параметров в среде сложных атрибутов, и вводится обычный термин кластеризации в разреженном подпространстве для устранения нерелевантной и избыточной информации из разнородных больших данных, и получается интеллектуальный индекс кластеризации разнородных больших данных. После измерения результатов кластеризации завершается разработка алгоритма интеллектуальной кластеризации гетерогенных больших данных в среде сложных атрибутов. Результаты экспериментов показывают, что алгоритм интеллектуальной кластеризации гетерогенных больших данных в среде сложных атрибутов обладает высокой стабильностью в процессе анализа и очистки данных.

Представлена архитектура программной системы оптимизации больших данных от датчиков в Интернете вещей, отличающаяся итерационным распределением больших данных на основе упорядочивания объектов и ссылок и реализующая уменьшение доли дубликатов и несоответствий в наборе данных в среднем на 12%.

**В заключении** представлены основные результаты работы, а также рекомендации и перспективы дальнейшей разработки темы.

# **1. Особенности разработки специального энергоэффективного программного обеспечения перемещения потоков данных в беспроводных информационных системах**

## **1.1. Проблемы управления беспроводными сенсорными сетями с точки зрения их энергоэффективности**

Беспроводные сенсорные сети (WSN) как компонент компьютерных сетей в настоящее время принадлежат одной из активно развивающихся областей знаний. Беспроводные сенсорные сети интегрируют механизмы идентификации, обработки и ретрансляции информации [2.1]. Приложения реального времени в области электроники, промышленной автоматизации, бизнеса, здравоохранения, автомобилестроения, авионики и др. используют WSN для повышения эффективности функционирования. Все эти отрасли используют WSN, интегрируя их во встроенные системы реального времени. По сравнению с проводными сетями беспроводные сенсорные сети (WSN) [2.2] обладают рядом преимуществ, включая отсутствие проводов и кабельной разводки, меньшую стоимость установки, меньший размер и возможность повторного использования, а также многими другими преимуществами, позволяющими удовлетворить растущие требования широкого спектра устройств из приложений.

WSN потенциально могут использоваться в самых разных контекстах. Однако необходимо решить ряд проблем, таких как потребление энергии, вычислительная мощность, срок службы, пропускная способность беспроводных сетей, объем памяти, которые необходимо решить. Для устранения выявленных проблем и продления срока службы беспроводной сенсорной сети было представлено и разработано несколько стратегий, касающихся маршрутизации [2.3], управления питанием и агрегации данных. Эти стратегии были разработаны для решения проблем, выявленных в WSN. Сети WSN становятся все более эффективными в использовании

энергии благодаря улучшенной маршрутизации и объединению данных. Это упрощает использование сетей WSN для приложений, которые все еще находятся в стадии разработки.

Беспроводная сенсорная сеть, также называемая WSN, представляет собой группу датчиков, которые подключаются на специальной основе. Каждый датчик представляет собой дешевое и компактное электрическое устройство, которое работает с другими датчиками для сбора информации.

Автономные по питанию WSN могут независимо от принадлежности какой-либо инфраструктуре взаимодействовать друг с другом в части обмена данными. Ниже рассмотрены три основные части сенсорных узла как компоненты его архитектуры. Эти компоненты включают в себя подсистему считывания, подсистему обработки и модуль беспроводной связи. Подсистема считывания отвечает за определение различных параметров окружающей среды, связанных с процессом отбора сигналов. Подсистема обработки отвечает за сбор данных, их обработку и хранение. Модуль беспроводной связи отвечает за передачу и прием сигналов. Процедура определения местоположения полностью выполняется этими подсистемами, каждая из которых отвечает за определенную ее часть. Каждый сенсорный узел имеет источник питания в виде батареи, и этот источник питания может обеспечивать только определенное количество энергии.

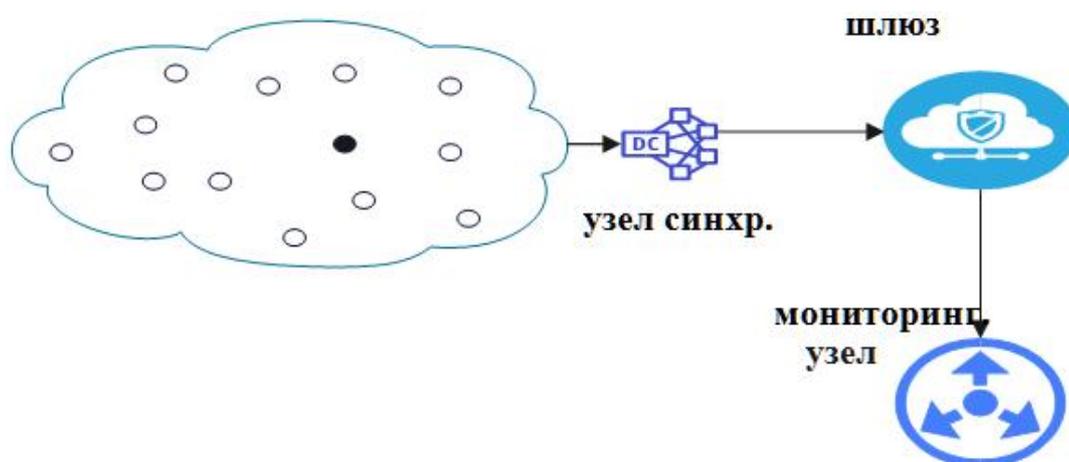


Рис. 1.1. Типовая архитектура WSN

Типичная архитектура WSN показана на рис. 1.1. Эта архитектура состоит из нескольких отдельных наборов сенсорных узлов, расположенных в различных частях мира. При поддержке единой операционной системы каждый сенсорный узел, составляющий архитектуру, реализует несколько различных интерфейсов прикладного программирования (API) [2.5]. Последние тенденции исследований подчеркивают тот факт, что все сенсорные узлы, принадлежащие WSN, расположены в определенном регионе развертывания и, будучи сгруппированы вместе, образуют кластер. Это является центральным принципом современных тенденций исследований. Невозможно создать полнофункциональный кластер без головного кластера, который представляет собой совокупность ресурсов, обладающих большей мощностью и эффективностью в вычислительном отношении.

В обязанности каждого руководителя кластера входит управление и сбор агрегированных данных с сенсорных узлов, связанных с его кластером. Эти сенсорные узлы выполняют процедуру сбора данных для выборки полученного сигнала. Затем полученные аналоговые значения преобразуются в цифровые значения для дальнейших вычислений и предварительной обработки. Затем информация обрабатывается каждым головным узлом кластера и отправляется на узел-приемник для последующей обработки. Принимающий узел или электронное устройство анализирует собранные данные и передает обработанные данные на станцию мониторинга через шлюзовой узел. Аналого-цифровой преобразователь - это компонент, встроенный в архитектуру каждого сенсорного узла. Основная функция этого компонента заключается в преобразовании аналоговых данных датчика в цифровые значения.

Сенсорный узел - это устройство, которое работает на основе электрических и электронных характеристик и отвечает за обнаружение неко-

торых наиболее важных физических свойств, таких как дым, влажность, тепло, давление, движение, температура и влажность, среди прочего [2.6]. Сенсорные узлы распределены по всему региону, который необходимо обследовать для сбора необходимых данных. Узлы часто размещаются либо случайным образом (когда местоположение сети неизвестно), либо, в большинстве случаев, равномерно [2.7].

MZDF (технология многозонного объединения данных): Этот этап является последовательным расширением предыдущего этапа, называемого этапом PEEDF. При создании MZDF был использован подход аналитического моделирования [8]. Стратегия многозонного объединения данных оказалась простой в реализации и высокоэффективной и, следовательно, была предложена в результате этой работы (MZDF). Концепция зон позволяет MZDF еще больше увеличить возможности объединения данных, что является одной из его основных функций. Зоны используются для того, чтобы изолировать узлы внутри сети, которые либо принимали участие в процессе объединения данных, либо не принимали в нем участия. Это то, что может сделать MZDF, поскольку у него есть возможность значительно расширить свои возможности по объединению данных. Было объяснено, как новый подход к топологическому проектированию может быть использован для повышения эффективности объединения данных с особым вниманием к требованиям, предъявляемым к системе в связи с потребностью крупномасштабных приложений в высокой производительности. В этой работе используется энергетическая модель, основанная на пороговых значениях, с целью упрощения процесса проектирования для передачи данных. Это было одной из основных причин использования данной модели.

Общая стратегия, которая использовалась в этом исследовании, направлена на совершенствование стратегий объединения данных, что важно, когда речь заходит об использовании крупномасштабных сенсорных

данных, собранных из многих доступных зон зондирования. Это была основная цель, на достижение которой было направлено данное исследование. При сравнении с LEACH с точки зрения показателей производительности, таких как долговечность сети, результаты моделирования показывают, что предложенный механизм слияния работает намного лучше, чем LEACH.

TBFT расшифровывается как метод слияния на основе дерева. На этом заключительном этапе внедрения метода ключевой целью является достижение оптимизации при объединении данных в WSN. Метод TBFT повышает эффективность объединения данных, дополняя его базовым и новым древовидным механизмом, который отличается от других существующих в настоящее время методов. На данном этапе предлагаемой работы был применен аналитический подход. В соответствии со стратегией, метод объединения данных на основе дерева реализован с использованием дополнительного модуля, известного как core fusion node. Этот дополнительный модуль гарантирует, что узел fusion объединит данные энергосберегающим способом, чтобы исключить избыточную информацию перед отправкой на базовую станцию в течение нескольких переходов. Предлагаемая сетевая архитектура включает в себя новый тип узла, называемый узлом-концентратором Fusion (HFN). HFN требуются, учитывая, что предлагаемая система имеет разработанную древовидную топологию.

Теоретические расчеты, разработанные в [2.9], позволяют эффективно модифицировать топологию системы внутри кластера, чтобы уменьшить количество энергии, используемой для обмена данными. Это было достигнуто за счет уменьшения количества узлов, подключенных друг к другу. Это, в свою очередь, продлевает срок службы системы и обеспечивает приемлемый уровень производительности для передачи информации.

Предварительное исследование показало, что прогнозируемый срок службы сенсорного узла, работающего от батареи, может быть значитель-

но увеличен, если применять рекомендуемые расчеты. Узел, о котором идет речь, питается от аккумулятора.

Исследование также показало, что соблюдение рекомендаций OSMRA может увеличить срок службы систем. Это достигается за счет снижения активности того компонента сенсорного блока, работа которого требует наибольшего расхода энергии. В исследовании [2.10] был предложен протокол маршрутизации на основе соответствия требованиям (FRP) в качестве метода для поддержки более эффективной передачи информации при меньшем потреблении энергии. Это было достигнуто с помощью сетевого протокола. Возможно уменьшить количество энергии, используемой сенсорными узлами, если выбрать узлы с наименьшим количеством скачков и расстоянием между ними. Вместо более трудоемких методов используется вычисление родословной, чтобы обнулить узел, который, с точки зрения проектирования, наиболее подходит для сети в целом. Концепция, разработанная в [2.11], заключается в том, что узлу, получившему наибольший объем информации, в первую очередь необходимо перейти к каналу, а затем он отправляет полученную информацию в приемник с помощью команды ожидания. Это пример сценария, в котором узлу в первую очередь необходимо подключиться к каналу. Информация передается с использованием различных методов, чтобы снизить энергопотребление при сохранении высокого коэффициента доставки пакетов. Как следствие, процесс определения точного момента, в который происходит сбой, приводит к значительному увеличению количества информации, которая может быть отправлена.

В [2.12] использовался протокол, известный как Ad Hoc on Demand Multipath Routing, для определения множества различных путей, которые могут быть использованы для передачи информации от исходного узла к узлу назначения. Для успешного выполнения этой задачи два узла сети были связаны друг с другом. Предложенная работа эффективно обеспечи-

вала продуктивную передачу данных с низким энергопотреблением, в то время как принимающий узел (базовая станция) находился в статическом состоянии. Это было окончательно завершено после того, как мы убедились, что все остальные соседние узлы находятся в состоянии готовности. В [2.13] исследовано, как базовая станция может помочь в прогрессивной маршрутизации на основе групповых соединений, используя LEACH, которое также может заменить сбор информации в процессе настройки. Предпосылка revolution включает в себя компонент, известный как "сбор голов кластера". Эта часть отвечает за получение информации от соседних узлов и отправку ее на разделительную базовую станцию.

В [2.14] предложен типичный протокол маршрутизации с дополнительной ориентацией, который назван WRP. С помощью этого подхода будет увеличена долговечность системы. В WRP часто считается, что узел-приемник обладает большими возможностями, чем другие узлы, и в значительной степени отвечает за накопление прогнозируемого времени пробуждения каждого узла. Несмотря на то, что некоторые сенсорные узлы находятся на расстоянии нескольких шагов от приемного узла, это позволяет центру приема проверить время срабатывания всех сенсорных узлов.

При разрыве большого количества соединений проанализированные данные используются для поиска альтернативного пути, который не приводит к перегрузке. В [2.15] представлено подробное описание эффективного традиционного протокола передачи информации об энергопотреблении на удаленные сенсорные устройства. С помощью протокола EAGRP (Energy Aware Geographic Routing Protocol) увеличивается срок службы системы. Данные об энергопотреблении и местоположении доступны пользователям. Для решения задач маршрутизации пакетов от источников к приемниками эти данные используются в протоколе EAGRP. Это приводит к сокращению срока службы датчика, а также системы в целом. Тем не менее, все еще возможно достичь более высокой скорости доставки па-

кетов при небольшом снижении энергоэффективности. Это связано с тем, что в качестве нормы, которой следует придерживаться, предлагается эффективная и энергосберегающая стратегия управления дистанционными сенсорными системами. Руководящая конфигурация EAGRP учитывает два фактора в процессе принятия решений: одним из них являются регионы, а другим - различные уровни энергопотребления узлов. Результаты измерений выполнения распределяются по переменному числу узлов, и каждый узел знает о площади поверхности и уровне энергопотребления окружающих его узлов. Анализ показывает, что предлагаемый подход не только сокращает время, но и повышает скорость доставки пакетов. Кроме того, достигается оптимальное использование имеющейся энергии.

Исследование было проведено путем внедрения нового решения проблемы эффективной передачи пакетов данных по разрозненным сенсорным сетям экологически безопасным способом. Система достаточно чувствительна и настроена, что позволяет использовать сенсорные системы и различные типы сенсорных узлов, которые способны определять энергию по-своему. В [2.16] предложено соглашение о доверительном управлении, которое базируется на регионе и защищает WSN от атак на маршрутизацию, одновременно обеспечивая поддержку крупномасштабных организаций WSN. Было показано, что предлагаемая конфигурация способна правильно различать дружественные и враждебные узлы, сохраняя при этом стратегическую дистанцию от последних. Кроме того, предложенная схема была реализована в сенсорных узлах, которые считаются лучшими в своем классе, чтобы служить надежным испытательным полигоном. Эта работа связана с затратами на определение использования средств измерения и уроками, извлеченными при разработке, внедрении и анализе гарантированных результатов.

Они представили новый стандарт непрерывной маршрутизации с двумя переходами (траекториями), основанный на учете энергопотребле-

ния. Дорожка позволяет повысить эффективность текущего выполнения метода за счет сокращения числа прямых альтернатив для сброса посылок, как это предусмотрено в настоящем соглашении. Стратегия основана на идее использования данных о двух соседях по прыжку и инструмента контроля силы. Это соответствует концепции использования двух соседей по прыжку. В то время как первый используется для повышения энергоэффективности операций, второй используется для повышения качества совместной передачи данных при одновременном снижении количества задерживаемых пакетов данных. Обе эти цели достигаются одновременно. Еще одно вычисление, представляющее собой улучшенную версию стандарта LEACH, называется вычислением LEACHTLCH. Этот расчет был разработан в [2.18]. Предполагается, что результаты этого расчета приведут к снижению энергопотребления всей системы в целом, а также к увеличению срока службы системы. Этап воспроизведения используется для создания копии нового расчета в Matlab. Результаты этого копирования показывают, что система потребляет меньше энергии и работает дольше, чем все, что может предложить протокол LEACH.

Можно с уверенностью предположить, что большая часть исследований, проведенных в прошлом, была посвящена моделированию агрегации данных, а не подходам к объединению данных. Это объясняется тем, что моделирование агрегации данных проще выполнять. Как прямой результат, такого рода подходы используются только с методами, основанными на слиянии.

Использование заранее определенной информации при принятии решений сенсорная сеть, обслуживающая приложения, которые зависят как от времени, так и от выполнения задач, использует данные, которые были запланированы, а также данные, собранные во время выполнения. С другой стороны, существующие исследования используют только статические знания, которые были заранее оговорены при маркетинге.

Это имеет более слабую связь с энергоэффективностью. Несмотря на то, что было проведено большое количество исследований в области повышения энергоэффективности сенсорных сетей, к сожалению, эти исследования не проводились с учетом технологии объединения данных. Несмотря на то, что было проведено множество исследований о том, как заставить сенсорные сети потреблять меньше энергии, это все еще актуально.

В последнее время все меньше внимания уделяется вычислительной сложности задач. Подавляющее большинство систем, представленных сегодня на рынке, предъявляют высокие требования к производительности своих коммуникационных возможностей; однако они делают это за счет увеличения сложности вычислений. Подходы, использующие ANN, GA или SVM, предполагают более широкое использование итераций с большими значениями периодов, что затрудняет реализацию этих стратегий на сенсорном узле с ограниченными ресурсами.

Таким образом, темы, которые были рассмотрены до этого момента, представляют собой основные пробелы в исследованиях, которые в настоящее время присутствуют в связи с совершенствованием процесса объединения данных с целью повышения энергоэффективности беспроводных сенсорных сетей.

## **1.2. Повышение энергоэффективности управления потоками данных в беспроводных сенсорных сетях**

В наши дни, с быстрым развитием микроэлектроники, Интернета вещей (IoT) и беспроводных сенсорных сетей (WSN), исследователи постоянно сталкиваются со многими проблемами, среди которых продление срока службы сети, оптимизация использования ресурсов, адаптация существующих алгоритмов управления потоками данных и так далее [3.1]. В случае WSN, развернутых на большом количестве микроэлектронных уст-

ройств [3.2, 3.3], они, как правило, сталкиваются с множеством ограниченных факторов, таких как ограниченные жесткие диски и физические ресурсы [3.4], ограничения топологии, вычислительная мощность, возможности хранения, связь дальность действия и энергопотребление [3.5, 3.6]. Они работают совместно для передачи данных [3.7, 3.8] из места развертывания в конкретное место назначения [3.9, 3.10], используя различные инновационные и конкурентоспособные протоколы управления потоками данных [3.11, 3.12].

### ***1.2.1. Поддержание работоспособности сети в течение длительного периода времени***

Поддержание работоспособности сети в течение длительного периода времени имеет решающее значение для производителей WSN [3.13]. Технологические достижения делают датчики доступными для многих областей применения, таких как окружающая среда, здравоохранение, транспорт, безопасность и армия [3.14, 3.15]. Эти небольшие устройства собирают данные и пересылают их по всей сети с помощью алгоритмов управления потоками данных, которые в основном используют возможности WSN и должны быть разработаны таким образом, чтобы оптимизировать ресурсы.

В работе используется стохастический метод, называемый скрытой Марковской моделью (НММ), для разработки нового вероятностного метода управления потоками данных. Получены результаты:

- определен новый метод оптимальной вероятностной управления потоками данных;
- разработан математический формализм НММ для WSNs;
- проведено моделирование НММ с использованием функций управления потоками данных;
- сформировано экспериментальное предложение в виде таблицы для

реализации новой модели стохастической управления потоками данных.

В литературе широко рассматривается проблема энергосбережения, направленная на увеличение срока доступности беспроводной сенсорной сети. В этой части некоторые из них представлены и сопоставлены в свете мотивации данной работы.

### ***1.2.2. Состояние проблемы***

Для решения проблемы энергопотребления в беспроводной сенсорной сети ранее были использованы методы кластеризации [13.6, 3.17]. Авторы в [3.18] предложили модель нечеткой логики, основанную на кластерном подходе, для улучшения энергосбережения с помощью протоколов управления потоками данных. Метод сжатия был использован [3.12] для оптимизации срока службы сети в период передачи. Этот метод увеличил коэффициент экономии энергопотребления и потреблял половину энергии по сравнению с другими.

Оптимизация роя дискретных частиц (DSPO) [3.19] была реализована в рамках подхода коммивояжера. Авторы модифицировали алгоритм для оптимизации топологии WSN и увеличили срок службы сети как минимум на 15% по сравнению с аналогичным периодом для немодифицированного алгоритма. В [3.20] метод прогнозирования, основанный на сокращении объема данных, был использован для предложения нового протокола управления потоками данных для повышения энергоэффективности в WSN. Кроме того, в WSN была использована скрытая марковская модель для обнаружения неисправных узлов и их замены в [3.21]. Полученные результаты демонстрируют улучшение реактивности сети.

В [3.22, 3.23], посвященных обнаружению атаки через уязвимость и предотвращению вредоносной активности, уничтожению данных с применением метода Витерби для определения оптимальной скрытой последовательности управления потоками данных, была разработана безопасная

стратегия на основе НММ для блокирования атаки уязвимости в WSN. В [3.24] на основе метода НММ была разработана методика обнаружения неисправных узлов, описывающая тенденцию к увеличению использования сетевых ресурсов определенными узлами.

В [3.25] использован формализм НММ для моделирования качества данных в WSN. Предложенный технологический метод был использован для повышения целостности данных и восстановления пропущенных значений.

Авторы [3.26] использовали НММ с полууправлением для отслеживания мобильных узлов и обучения своей системы на основе измерений точек сетки. Окончательные результаты помогли повысить точность локализации узлов и продемонстрировать полезность метода.

В [3.27-3.29] использовалась скрытая марковская модель для повышения безопасности беспроводной сенсорной сети, предложен оптимизированный метод с учетом энергопотребления, а также разработаны новые алгоритмы и изучена их сложность для решения проблем энергопотребления из-за использования сетевых ресурсов многими факторами. Кроме того, в работах [3.30, 3.31] рассматриваются новые тенденции, связанные с методом совместной управления потоками данных, основанным на схеме кластеризации для сбора энергии в беспроводных сенсорных сетях и для отслеживания наличия мощных устройств Интернета вещей в облачной сети.

### ***1.2.3. К задаче управления потоками данных для повышения энергоэффективности***

Повышение энергоэффективности является решающим фактором при разработке моделей управления потоками данных в беспроводных сенсорных сетях (WSN). Многие схемы управления потоками данных с низким энергопотреблением реализованы для обмена различными форма-

ми собранных датчиками данных по оптимальному пути потоков данных через сеть, что увеличивает срок ее службы и поддерживает высокую масштабируемость WSN. В этой главе предлагается стохастическая модель управления потоками данных с повышением энергопотребления для снижения использования ресурсов узлами в процессе управления потоками данных. Мы стремимся адаптировать стохастический формализм, основанный на скрытых марковских моделях (НММ), к опыту существующих сенсорных сетей и разработать новый оптимальный механизм управления потоками данных, который значительно задействует доступные ресурсы. Между тем, предложенный алгоритм стохастического управления потоками данных обеспечивает общее снижение энергопотребления сети и обеспечивает оптимальную передачу данных. Результаты экспериментов показывают, что предложенный метод эффективен с точки зрения энергопотребления, общего увеличения ресурса и позволяет увеличить срок службы сети как минимум на 19,05% по сравнению с другими существующими методами.

#### ***1.2.4. Направление исследования***

Методы НММ широко использовались в WSN и других [3.32, 3.33]. В работе предлагается новый метод статистического управления потоками данных, основанный на НММ, для наблюдения за поведением существующей сети и изучения его результатов. Мы используем данные, полученные самой сетью, для инициализации предложенной стохастической модели. Наш подход позволяет нам прогнозировать оптимальные кратчайшие пути после нескольких этапов обучения, используя метод максимизации математического ожидания [3.34]. Кроме того, модель является автоадаптивной, использует оставшуюся энергию и динамически выбирает кратчайший путь на основе соответствующей энергии.

Теоретическая математическая часть скрытой марковской модели

представлена в следующем разделе.

### ***1.2.5. Теория скрытых марковских моделей***

Скрытая марковская модель (НММ) - это стохастический метод, который продемонстрировал свою эффективность во многих областях [3.32, 3.35], таких как обработка изображений, распознавание цифровых отпечатков, классификация, сегментация ДНК [3.33, 3.36] и так далее. В нашем контексте формализм НММ [3.37], дающий хорошие результаты в указанных областях, адаптирован для реализации вероятностного протокола управления потоками данных. Насколько известно, этот метод впервые был использован в протоколе управления потоками данных.

Фактически, процессы реального времени производят наблюдения, которые могут быть дискретными, непрерывными, стационарными и изменяться во времени.

Основная трудность заключается в том, чтобы охарактеризовать наблюдения как случайные процессы, которые следует оценивать с помощью четко определенного теоретического подхода. Мы рассматриваем протокол управления потоками данных как дискретное событие, и его функции используются для определения скрытых событий как переменных марковской модели [3.38]. Идея состоит в том, чтобы оценить предположения о скрытых путях с помощью вероятностного скрытого марковского метода. Преимущества этого метода заключаются в том, что он требует меньше ресурсов, позволяет извлекать информацию из существующих параметров работы WSN и прогнозировать оптимальный способ передачи данных [3.39].

### ***1.2.6. Основные переменные***

В табл. 1.1 приведены основные переменные, используемые в работе. Рассмотренные скрытые состояния представляют энергии различных

путей и обозначаются символом E. Все эти обозначения повторяются в тексте.

Таблица 1.1

Обозначения НММ

обозначение	описание
A	Вероятности перехода между состояниями
B	Матрица вероятностей наблюдения
M	Количество символов наблюдения
N	Количество состояний в модели
P	Распределение начальных состояний
S	Различные состояния марковского процесса ( $s_0, s_1, \dots, s_{N-1}$ )
T	Длина последовательности наблюдений
V	Набор возможных наблюдений (0, 1, ..., M-1)
Y	Последовательность наблюдений ( $Y_0, Y_1, \dots, Y_{T-1}$ )
X	Контекстуальный набор скрытых состояний, $\Omega = \{S_i\} = \{E_1, E_2, E_3\}$

### 1.2.7. Параметры модели

Модель НММ, обозначаемая  $l$ , состоит из набора параметров, которые определяются следующим образом.

$$l = (A, B, \Pi) \quad (1.1)$$

где

- Вероятности начального состояния  $\Pi$ :

$$\Pi_i = [P(X_1 = E_i / l)]_{1 \leq i \leq N} \quad (1.2)$$

где

$$\sum_{i=1}^N p_i = 1 \quad (1.3)$$

- Матрица A переходов состояний:

$$A=a_{ij}=[P(X_t=E_i/X_{t-1}=E_i, I)]_{1 \leq i, j \leq N} \quad (1.4)$$

где матрица должна соответствовать следующему предположению:

$$\sum_{j=1}^N a_{ij} = 1 \quad (1.5)$$

Каждый элемент  $a_{ij}$  представляет собой коэффициент перехода между скрытыми состояниями. Эмпирический метод, используемый в этой работе для оценки начальных коэффициентов, описан в [3.40], значение переходных состояний зависит от энергии каждого пути.

- Матрица  $N \times M$  вероятности излучения  $B=b_j(v_k)$  представляет вероятность наблюдения в момент времени  $t$  символа  $v_k$  при заданном состоянии  $E_j$ .

- Последовательность наблюдений  $V=v_1, v_2, \dots, v_M$  - это последовательность из  $M$  наблюдений, где  $v_j \in \{\text{Низкий; Средний; Высокий}\}$ .

Символы наблюдений относятся к физическим характеристикам системы. В работе они состоят из оставшейся наблюдаемой мощности узлов для разных состояний и называются низким (L), средним (M) и высоким (H).

- Для простоты в остальной части текста мы используем обозначение  $i$  для  $E_i$ ,  $k$  для  $v_k$ .

В главе 3 приводится более подробная информация, относящаяся к предлагаемому методу. Описывается архитектура системы и марковская реализация.

### **1.3. Подходы к разработке математического и программного обеспечения управления данными в специальных задачах**

#### ***1.3.1. Проблема обнаружения аномальных данных в системе автоматизации измерения мощности***

Система автоматизации учета электроэнергии относится к системе, которая может собирать, отслеживать и анализировать электрические дан-

ные на стороне производства электроэнергии, на стороне электроснабжения, на стороне распределения электроэнергии и на стороне продажи электроэнергии электростанциями, подстанциями, трансформаторами общего пользования, специальными трансформаторами и потребителями низкого напряжения, включая главную станцию автоматизации учета, канал связи и терминал автоматизации учета. Система автоматизации учета обеспечивает дистанционное автоматическое считывание показаний счетчика в режиме реального времени, аварийную сигнализацию о ненормальном потреблении электроэнергии, мониторинг напряжения и качества электроэнергии, анализ обрыва линии и предоплату путем сбора, мониторинга, анализа и обработки такой информации, как данные, напряжение, ток, нагрузка и другая информация каждого терминала мониторинга, например управление сборами, проверка потребления электроэнергии, управление нагрузкой и контроль за ней, а также статистика перебоев в подаче электроэнергии обеспечивают информационную поддержку для работы сети и управления ею [4.1]. Но прежде чем приступить к реализации этих передовых прикладных исследований, самой основной предпосылкой является обеспечение своевременности, целостности и надежности данных, собираемых системой.

В настоящее время данные, передаваемые в систему автоматизации измерений, в основном включают различные показатели терминала, такие как скорость онлайн-подключения, скорость автоматического считывания показаний счетчика и т.д., собранные данные, такие как табличные коды, электричество и коэффициент мощности, и эти базовые данные подвергаются вторичным вычислениям, таким как усовершенствованное приложение для деления напряжения такие данные, как потеря данных, обрыв линии, обрыв линии на подстанциях и т.д. Столкнувшись с таким большим объемом данных, мы не смогли эффективно гарантировать надежность качества данных, полагаясь на ручные или традиционные программные

средства для работы с базами данных для ежедневной проверки качества данных. Традиционным методом обнаружения посторонних данных является использование классификатора RBF для обнаружения посторонних данных, и временные затраты на этот метод обнаружения для обнаружения данных большой системы слишком высоки, а погрешность результатов обнаружения также велика [4.2]. Таким образом, основываясь на приведенном выше анализе, рассматривается метод обнаружения аномальных данных в системе автоматизации измерения мощности.

### ***1.3.2. Система динамического мониторинга потерь больших объемов данных в мобильной сети на базе Интернета вещей***

С целью решения проблемы большой ошибки синхронизации времени сбора данных, вызванной резким увеличением объема данных в традиционных системах мониторинга, разработана система динамического мониторинга потерь больших объемов данных в мобильной сети на базе Интернета вещей. В системе организована беспроводная сенсорная сеть, спроектирован многоканальный узел базовой станции, а датчики расположены в разных каналах для достижения цели передачи данных. Микросхема Er3c16q240 выбрана в качестве основной микросхемы управления многоканальным узлом базовой станции, исходя из вышеуказанной аппаратной схемы, функции мониторинга кластера, мониторинга производительности узла и мониторинга выполнения заданий предназначены для загрузки информации о состоянии больших данных мобильной сети для пошагового мониторинга выполнения заданий в соответствии с потребностями динамического мониторинга потерь данных. На данный момент общий дизайн системы завершен. Результаты экспериментов показывают, что: по сравнению с традиционной системой мониторинга, разработанная система мониторинга, основанная на Интернете вещей, имеет меньшую ошибку синхронизации по времени сбора данных, лучшую производительность

синхронизации сбора данных и повышает точность динамического мониторинга потерь больших объемов данных в мобильной сети.

### *Динамический мониторинг потерь больших объемов данных в мобильной сети Интернета вещей*

Являясь важной вехой в современной истории развития человечества, Интернет является символом человеческих инноваций и мудрости, а также важным символом стремительного развития науки и техники [4.9]. С момента своего появления и до настоящего времени Интернет внес большие изменения в индустрию и жизнь всего мира. Он не только реализует предпринимательскую мечту многих людей, но и делает жизнь людей неразрывной и оказывает взаимное влияние. Он также вносит большие изменения в организационную структуру общества [4.10]. Интернет оказывает широкое и далеко идущее влияние, жизнь людей стала более удобной, а способы распространения информации - разнообразными. Благодаря быстрому развитию науки и техники развитие Интернета всегда шло относительно быстрыми темпами. С быстрым развитием масштабов сети услуги, предоставляемые сетью, также диверсифицируются, что значительно облегчает жизнь людей [4.11]. Развитие Интернета стало более масштабным, особенно с появлением мобильных клиентов, которые коренным образом изменили жизнь интернет-пользователей. С появлением все большего числа мобильных приложений, таких как мобильный банкинг, заказ билетов на поезд и еды навынос, Интернет начал в полной мере удовлетворять потребности пользователей, делая жизнь людей все более «сетевой» [4.12].

В первые десятилетия после появления компьютерных сетей они использовались в основном университетскими исследователями для отправки электронной почты и сотрудниками компаний для совместного использования принтеров. В этих случаях безопасность не будет замечена [4.13]. В настоящее время миллионы обычных людей пользуются Интернетом.

То, что мы хотим делать в реальном мире, должно осуществляться в Интернете; личные звонки, сохранение личных документов, подписание писем и контрактов, онлайн-голосование, электронные публикации, банковские операции и покупки требуют защиты [4.14]. Сетевая безопасность является основным условием существования Интернета, что превращает компьютерную сеть из важного бизнес-инструмента в научный инструмент. Ограничения безопасности также стали ограничениями Интернета [4.15]. Уязвимости в системе безопасности обнаруживались одна за другой, и сетевая безопасность стала актуальной темой, которой люди уделяют пристальное внимание. В настоящее время, сталкиваясь с проблемой потерь больших объемов данных в мобильных сетях, особенно важно внедрить механизм мониторинга в систему управления сетевой безопасностью и создать мощную систему динамического мониторинга потерь больших объемов данных в мобильных сетях [4.16]. Используя этот механизм мониторинга, можно отслеживать безопасность данных в мобильной сети в режиме реального времени и предупреждать о таких нештатных ситуациях, как потеря данных. В случае сбоя необходимо своевременно информировать персонал управления после возникновения проблемы, чтобы обеспечить безопасность данных в мобильной сети.

В предыдущих исследованиях говорилось, что существует множество развитых технологий мониторинга и систем мониторинга с открытым исходным кодом. Система мониторинга может предоставлять информацию о состоянии сети и системы, а также предоставлять функции уведомления о сбоях в работе. Мониторинг может осуществляться как на локальных, так и на удаленных серверах. Необходимо лишь изменить конфигурационный файл [4.17, 4.18]. В настоящее время, столкнувшись с проблемой потерь данных в мобильной сети, многие начинают осваивать облачные вычисления и изучать различные решения для мониторинга больших данных. Очевидными результатами являются система мониторинга на базе ZigBee

и веб-система мониторинга. Однако, учитывая текущее состояние растущего объема данных, вышеупомянутые две системы с трудом обеспечивают синхронизацию сбора данных, и существует проблема больших ошибок синхронизации во времени сбора. Учитывая это явление, в работе предлагается и проектируется система динамического мониторинга потерь больших объемов данных в мобильной сети на основе Интернета вещей. Разработано аппаратное обеспечение системы мониторинга. Микросхема Er3c16q240 выбрана в качестве основной управляющей микросхемы многоканального узла базовой станции. На основе проектирования аппаратного обеспечения разрабатывается системное программное обеспечение, включая функцию мониторинга кластера, функцию мониторинга производительности узлов и функцию мониторинга работы. С помощью проектирования аппаратного обеспечения системы и программного обеспечения, завершается проектирование системы динамического мониторинга потерь больших объемов данных в мобильной сети на основе Интернета вещей, что сокращает время синхронизации сбора данных, повышает точность динамического мониторинга потерь больших объемов данных в мобильной сети и повышает производительность синхронизации сбора данных.

#### **1.4. Постановка задач работы**

Таким образом, проведено исследование проблем управления средствами разработки специального энергоэффективного программного обеспечения перемещения данных в беспроводных информационных системах на основе агентного подхода и фреймворка динамической реконфигурации для технологии слияния. Отмечено, что повысить эффективность такого управления можно путем применения модификации стратегии управления данными, отличающейся применением агентного подхода и фреймворка динамической реконфигурации для технологии слияния на основе дерева, позволяющей повысить энергоэффективность маршрутизации в крупно-

масштабных беспроводных сенсорных сетях, создания алгоритма улучшения показателей энергоэффективной связи объектов беспроводных информационных систем, отличающегося автоматическим выбором гетерогенных факторов ранжирования узла слияния – кандидата с наибольшей реализуемостью, и обеспечивающего сокращение общего времени, необходимого для завершения процесса кластеризации, применения гибридного итерационного алгоритма поиска аномальных данных, сочетающего применение матрицы сходства аномальных признаков данных и кластерный анализ с использованием обучения iForest, разработки структуры программной системы динамического мониторинга потерь больших объемов данных с многоуровневым сбором информации (уровни кластера, производительности узлов и работы заданий), обеспечивающей максимальную точность мониторинга. Результат анализа потребовал формализации данных задач, а также алгоритмизации их решения с учетом особенностей, отраженных на рис. 1.1. Сформулирована цель и задачи исследования.

**Целью работы** является создание моделей анализа и алгоритмов управления процессами разработки специального энергоэффективного программного обеспечения перемещения потоков данных в беспроводных информационных системах.

**Задачи исследования.** Для достижения поставленной цели необходимо решить следующие задачи:

1. Проанализировать проблематику моделирования и алгоритмизации процессов разработки специального энергоэффективного программного обеспечения перемещения потоков данных в беспроводных информационных системах.

2. Разработать модель снижения энергопотребления за счет использования вероятностных методов и кластеризации.

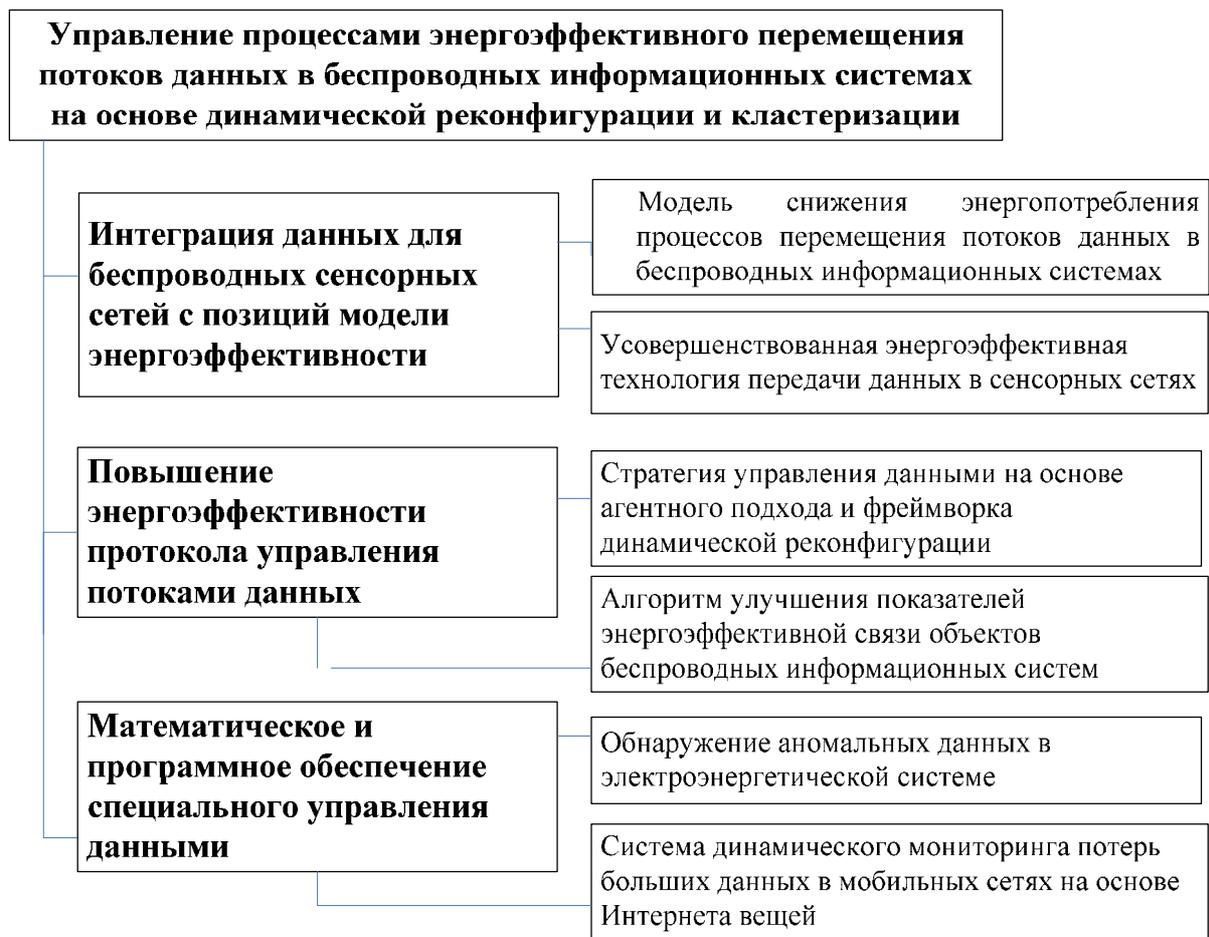


Рис. 1.1. Дизайн исследования

3. Предложить модификацию стратегии управления данными, позволяющую повысить энергоэффективность маршрутизации в крупномасштабных беспроводных сенсорных сетях на основе агентного подхода и фреймворка динамической реконфигурации для технологии слияния на основе дерева.

4. Разработать алгоритм улучшения показателей энергоэффективной связи за счет сокращения общего времени, необходимого для завершения процесса.

5. Разработать структуру программной системы динамического мониторинга потерь больших объемов данных с многоуровневым сбором информации (уровни кластера, производительности узлов и работы заданий), повышающую максимальную точность мониторинга.

## **2. Интеграция данных для беспроводных сенсорных сетей с позиций модели энергоэффективности**

### **2.1. Проблема проектирования энергоэффективных маршрутов в беспроводных сенсорных сетях**

Беспроводные сенсорные сети внесли значительный вклад в развитие систем беспроводной сенсорной связи, основанных на ограниченных ресурсах и вычислительных возможностях датчиков. За последнее десятилетие было предпринято несколько целенаправленных исследовательских работ, направленных на изучение и поиск решений проблем, связанных с объединением данных об энергоэффективности в беспроводных сенсорных сетях. Однако проблема проектирования энергоэффективных маршрутов так и не была решена. Гарантировать продление срока службы датчика на более длительный период довольно сложно из-за ограниченных вычислительных возможностей автономных датчиков, которые часто сопровождаются ограничениями по энергопотреблению. Результаты этой работы представляют собой усовершенствованную энергоэффективную технологию передачи данных в сенсорных сетях, которая состоит из трех различных инновационных систем.

### **2.2. Усовершенствованная энергоэффективная технология передачи данных в сенсорных сетях**

Предложенная система - объединение данных с потенциальной энергоэффективностью (DFWPEE), отвечает за оптимизацию энергопотребления. Предлагаемая работа позволяет снизить энергопотребление за счет использования вероятностных методов и кластеризации. В процессе объединения данных архитектура Multiple Zone Data Fusion (MZDF) использует глобулярную топологию, которая помогает распределить нагрузку. Стратегия представляет собой подход к маршрутизации, который используется

для повышения энергоэффективности маршрутизации в крупномасштабных беспроводных сенсорных сетях. Внедряя идею агентов маршрутизации, фреймворк для технологии слияния на основе дерева (ТВФТ), предлагает метод динамической реконфигурации. Этот план позволяет системе определить, какой датчик имеет более высокую скорость рассеивания энергии, а затем немедленно передать задачу по объединению данных узлу, который является более энергоэффективным. Этот метод, основанный на пороговых значениях, позволяет датчику выполнять как роль головного узла кластера, так и функцию узла-члена. Узел ведет себя как головной узел кластера, пока не достигнет порогового значения остаточной энергии, и функционирует как узел-участник после того, как он преодолет пороговое значение остаточной энергии. Обе эти роли могут выполняться одновременно. Математическое моделирование проводилось с использованием традиционной модели использования энергии радиоканалов, что повысило надежность полученных результатов. Предлагаемая система обеспечивает улучшенные показатели энергоэффективной связи при сравнении с существующими стандартами для энергоэффективных схем. Усовершенствованный метод использует почти вдвое меньше энергии, чем LEACH, при этом основное внимание уделяется сокращению общего времени, необходимого для завершения процесса, что приводит к повышению производительности.

Основная цель предлагаемой модели - оптимизировать многие доступные в настоящее время источники энергии. В представленном исследовании представлена структура оптимизации, известная как Potential Energy Efficient Data Fusion (DFWPEE). Эта структура гарантирует, что центр кластера в области беспроводных сенсорных сетей может быть выбран наиболее эффективным из возможных способов, тем самым повышая общую производительность сети. Аппарат теории вероятностей применен для проведения расчетов вероятности выбора главы кластера на основе резуль-

татов имитационного исследования. Каждый из сенсорных элементов произвольно размещен в пространстве моделирования, на этом основан предложенный метод. Метод стартует при получении датчиком уникальных сведений о конкретном событии. Работа продолжается до тех пор, пока полученные данные не будут ретранслированы на базовую станцию.

$$P_{Dx} = L(P_{elec} + P_{fx}xd^2), d < d_0 \quad (2.1)$$

$$P_{tx} = L(P_{elec} + P_{mp}Xd^4), d \geq d_0 \quad (2.2)$$

Особенностью предлагаемой системы является разнотипность сенсорных узлов, похожих друг на друга. Особенностью инфраструктуры также является множество факторов, например, расстояние между узлами датчиков и центром кластера, а также количество начальной энергии, которое изначально было присуще каждому датчику. Ожидается, что базовая станция будет расположена на большом расстоянии от кластера, чтобы получать информацию о стабильности и долговечности кластеров на протяжении всего срока их службы. Стартовый этап процесса проектирования системы основан на ориентированной на энергоэффективность сетевой модели.

$$d_0 = \sqrt{\frac{P_{fo}}{P_{up}}} \quad (2.3)$$

Практическая часть исследования посвящена верификации модели. При этом считаем, что для обработки предоставлено  $L$  бит данных, расстояние  $d$ , причем количество энергии для обработки фиксировано. Величина  $P_{elec}$  в исследовании есть энергозатраты на аппаратную обработку одного бита информации как источником, так и получателем. Из (3) вычисляется дополнительная переменная расстояния  $d_0$  для представления доли неэффективно расходуемой энергии из-за неоптимального метода передачи. В результате приведенных выше рассуждений энергия, необходимая для передачи информации длиной  $L$  бит между парой сенсорных узлов,

может быть представлено в виде

$$P_{Dx} = L \cdot P_{elec} \quad (2.4)$$

Более подробный анализ уравнения (2.4) показывает, что уравнение корректно тогда и только тогда, когда расстояние передачи  $d$  полностью определяет мощность передачи  $P_{Dx}$ . Этот вывод был сделан после более детального анализа уравнения. По сравнению с (2.2), в котором используется четвертая степень дальности передачи, рассеиваемая мощность при передаче будет уменьшена, если расстояние передачи будет возведено в квадрат, а не в четвертую степень. Это произойдет потому, что квадрат дальности передачи будет больше, чем расстояние передачи. Как следствие, совершенно очевидно, что в данном сценарии используется стратегия эксплуатации. В частности, система должна соответствовать эмпирическому уравнению, которое включает квадрат расстояния, чтобы поддерживать более низкий уровень рассеиваемой мощности при пропускании. Это необходимо для того, чтобы система функционировала должным образом. Экспериментальное выражение количества ненужной энергии, которое расходуется на получение  $L$  пакетов данных в битах, лучше отражает количество энергии, которое тратится впустую.

$$P_{CL} = \frac{\alpha n}{c_k} - 1 \frac{\ddot{o}}{\emptyset} L g_{dik} P + \frac{n}{k} L g_{du} P + L g_{slax} E + L s_{fs} g_{gs}^2 \quad (2.5)$$

Когда дело доходит до надежности выполнения уравнений в приложениях любого размера в беспроводных сенсорных сетях, это довольно легко обосновать без каких-либо дополнительных сравнений. В целях дальнейшей стандартизации работы метода предлагаемая система учитывает тот факт, что выбор оптимального кластера осуществляется следующим образом, если измерение расстояния между сенсорными узлами превышает допустимое. Если это так, то стандартизация работы метода будет улучшена:

$$P_{SCL} = L P_{diec} + L s_{fs} d_a^2 \quad (2.6)$$

Предлагаемая система, в отличие от ранее существовавших методов оптимизации, таких как рекурсивная оптимизация, предусматривает шаги, которые являются одновременно простыми и итеративно менее сложными, чтобы вычислить оптимальное количество сенсорных узлов, которые должны участвовать в процессе объединения данных. Эти шаги предназначены для определения оптимального количества сенсорных узлов, которые должны участвовать в процессе объединения данных. Это сделано для того, чтобы предлагаемая система отличалась от других подходов к оптимизации, которые уже существуют. Уравнение описывает эмпирический подход к расчету общего числа сенсорных узлов elite в беспроводной сенсорной сети, которые способны осуществлять слияние данных. Этот метод может быть использован для определения общего числа сенсорных узлов elite в сети (2.6). При таком подходе учитывается очень низкое энергопотребление беспроводных сенсорных сетей. В отличие от традиционного метода, при котором выбор центра кластера осуществляется исключительно на основе остаточной энергии и итеративного процесса, предложенная система учитывает широкий спектр характеристик, прежде чем прийти к оценке требуемой мощности. Это отличается от традиционного метода, при котором выбор центра кластера производится исключительно на основе остаточной энергии и итеративного процесса. Используется принцип оптимизации, и эмпирически определяется повышенная вероятность того, что обычный узел будет выбран в качестве вероятного центра кластера, как наилучшее значение  $n$  к  $p$ . Эта формула отражает повышенную вероятность того, что обычный узел будет выбран в качестве вероятного центра кластера. Из-за этого вероятность того, что будет выбран обычный узел, повышается.

$$d_{CL}^2 = \iint_{0 \ 0}^{x_1 \ x} (x^2 + y^2) \mathbf{g}(x, y) dx dy = \frac{M^2}{2pk} \quad (2.7)$$

Приведенное уравнение подчеркивает вероятность того, что обыч-

ный датчик сможет выполнить объединение данных с наивысшим уровнем достоверности (2.7). Принимая это во внимание, представляется, что предлагаемая система помогает извлекать информацию, необходимую для выбора датчика, который является наиболее подходящим при проведении теста. Этот процесс позволяет избавиться от значительного числа повторяющихся процессов, что, в свою очередь, сокращает время, необходимое для реагирования на событие, и перекладывает бремя выполнения задачи со старого центра кластера на более современного. Поскольку неправильный выбор центра кластера не обеспечит эффективную передачу данных от кластеров к базовой станции, что приведет к перерасходу энергии и ненужному потреблению ресурсов, основной целью предлагаемого исследования DFWPEE является выбор эффективного центра кластера. Это связано с тем, что неправильный выбор центра кластера не приведет к распространению данных. Это связано с тем, что исследование, которое было проведено с целью подчеркнуть, что общее количество потребляемой энергии во всей сенсорной сети растет с каждой итерацией, привело к выводу, что это действительно так.

$$P_t = Lg(2nP_{die} + nP_{Dd} + (kd_{ks}^2 + ndd_{CL}^2)) \quad (2.8)$$

$$d_{As}^2 = \underset{A}{\overset{\circ}{O}}(x^2 + y^2) \underset{A}{g} \frac{1}{2} = 0.765g \frac{M}{2} \quad (2.9)$$

Распространенной практикой является распространение информации о количестве энергии, которое все еще доступно в сообщающихся сенсорных узлах, при разработке мероприятий, включающих обновление данных. Это делается с целью увеличения периода времени, в течение которого сеть способна поддерживать работу сенсорных узлов. Этот процесс выполняется таким же образом даже на более позднем этапе выбора центра кластера.

Можно продолжить процедуру при условии, что накопленная энер-

гия сенсорных узлов не будет полностью израсходована, но это зависит от выполнения данного требования. Предлагаемый метод не предусматривает выбора потенциального центра кластера, который уже выполнял обязанности центра кластера в прошлом. Это сделано для того, чтобы процесс выбора центра кластера был настолько простым и однозначным, насколько это в человеческих силах. В следующем разделе обсудим метод исследования, который был использован в данном исследовании для достижения целей, поставленных перед данным исследованием.

### **2.3. Метод исследования**

В предлагаемом исследовании метод DFWPEE включает в себя как аналитический, так и эмпирический подход. Использование вероятностного подхода было ключом к достижению основной цели метода исследования, которая заключалась в разработке системы, которая позволила бы сенсорным узлам участвовать в выборе центра кластера наиболее эффективным образом. Это было достигнуто с помощью метода исследования. Схематическую иллюстрацию предложенного дизайна DFWPEE можно увидеть в выделенной области рис. 2.1. Процедура начинается с построения модели сети. Затем исследуется передача посредством центра кластера объединенных данных между датчиками и базовой станцией. Процедуру завершает этап анализа данных. Кластеризация выполняется иерархическим алгоритмом, что обусловлено учетом остаточной мощности сенсорных узлов и характеристиками энергоэффективности самого алгоритма. Метод иерархической кластеризации был выбран из-за этих особенностей. Рекомендованный подход будет использовать меньшее число для начальной энергии, и ему будет предоставлен шанс выбрать центр кластера для каждого раунда роста, который равен 5% от общей вероятности. Эти изменения внедряются в систему с целью приведения ее в более полное соответствие с текущим положением дел.

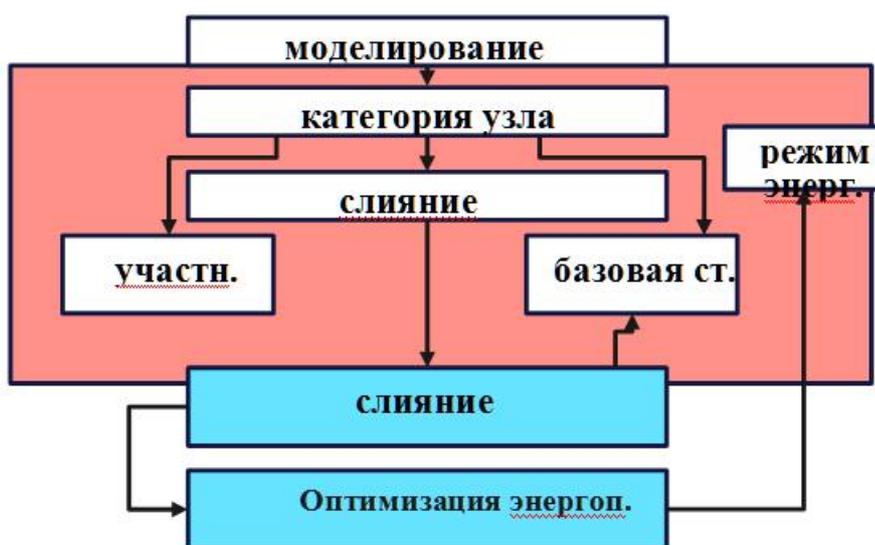


Рис. 2.1. Предлагаемая архитектура

Разработка схемы вероятностного выбора, которая может эффективно выбирать оптимальный центр кластера и управлять другими узлами, чтобы они становились центрами кластера в последующих раундах имитационного моделирования, является наиболее важным аспектом принципа проектирования. Это связано с тем, что схема должна быть способна эффективно выбирать оптимальный центр кластера. Каждый из узлов-кандидатов обладает достаточным количеством энергии, что позволяет им самостоятельно участвовать в следующем раунде конкурса на звание центра кластера, который проводится в рамках имитационного исследования. Метод учитывает не только узлы, которые являются претендентами на пост центра кластера, но и узлы, которые не являются кандидатами. Это узлы, которые имеют гораздо меньшее общее количество энергии и меньшую вероятность занять эту позицию в сети. С другой стороны, на следующем этапе моделирования существует вероятность того, что узлы, не являющиеся кандидатами, будут переведены в положение узлов-кандидатов. Это возможно. Напомним, что методика сбора определена в основном на уровне накопленной энергии.

## 2.4. Результаты экспериментов и их анализ

Исследование проводится на 32-разрядной операционной системе Windows с двухъядерным процессором, работающим на частоте 1,84 ГГц, в качестве интегрированной среды разработки используется MATLAB. Вот несколько различных сценариев, которые были приняты во внимание, каждый из которых имеет от 60 до 100 узлов и случайное распределение. Процедура инициализации состоит из 150 циклов, а начальная энергия узлов установлена на уровне 0,1 Джоулей. Для завершения процесса также требуется выполнить 150 раундов. Метод LEACH, который используется большим количеством исследователей, сравнивается со стратегией, которая была представлена выше. Ниже приведен список параметров, которые учитывались при выполнении сетевого моделирования.

Единственная цель этой работы - обеспечить, чтобы реализация схемы максимально увеличивала бесперебойность объединения данных и одновременно сводила к минимуму любые нежелательные трудности с точки зрения энергопотребления. Компоненты правила были разработаны таким образом, что гетерогенный тип ранжирования является улучшенной версией детерминированной схемы, которая используется в однородном типе ранжирования.

На рис. 2.2 показаны результаты моделирования предложенной системы DFWPEE. В контексте этого моделирования можно продемонстрировать локальную оптимальность информации о ранжировании узлов. В рамках гетерогенного аспекта ранжирования, в дополнение к мощности, будет использована информация, касающаяся оценок толерантности и надежности данных (0,1J). Эксперимент проводится путем корректировки размеров соседних узлов. Ясно видно, что количество соседних узлов, которые находятся ближе друг к другу, напрямую связано с практической надежностью данных на узле объединения данных.

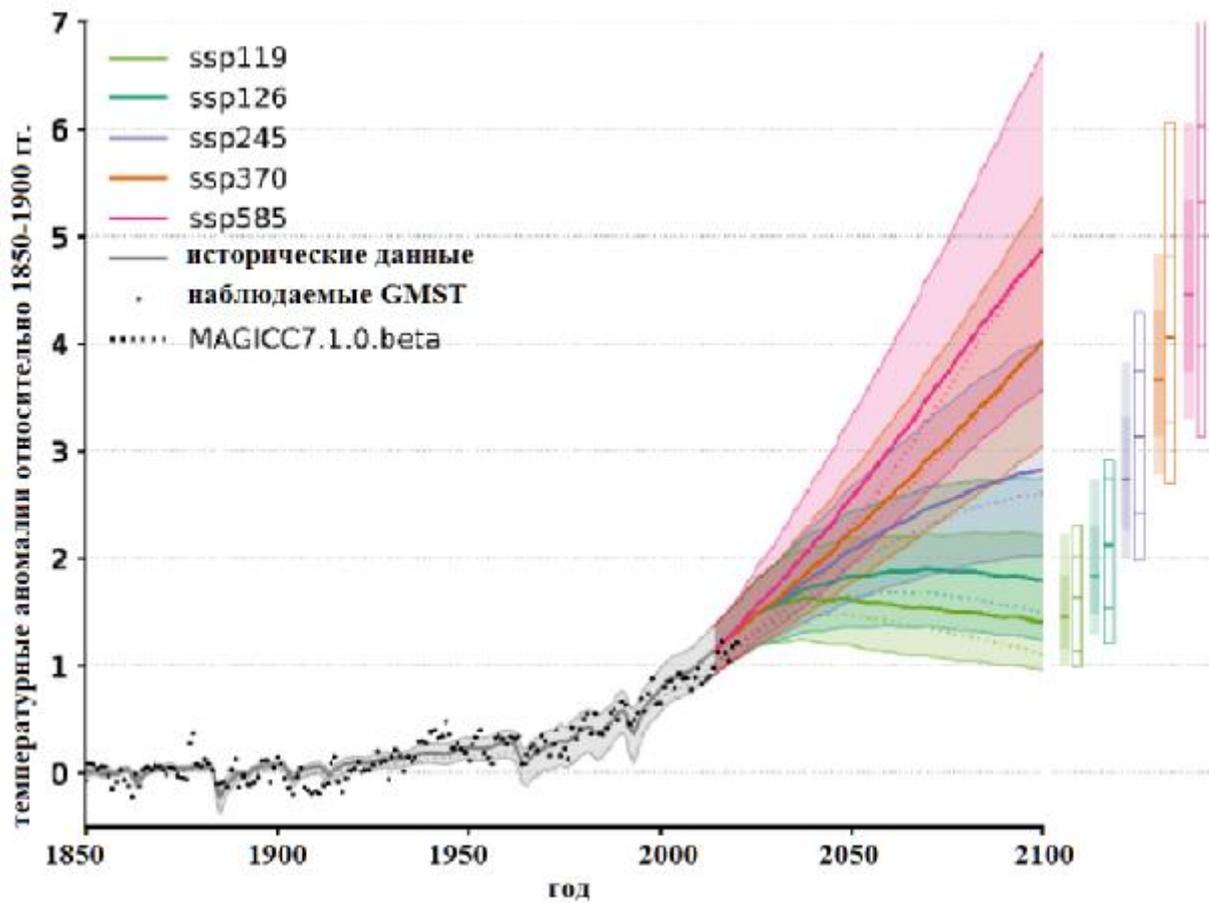
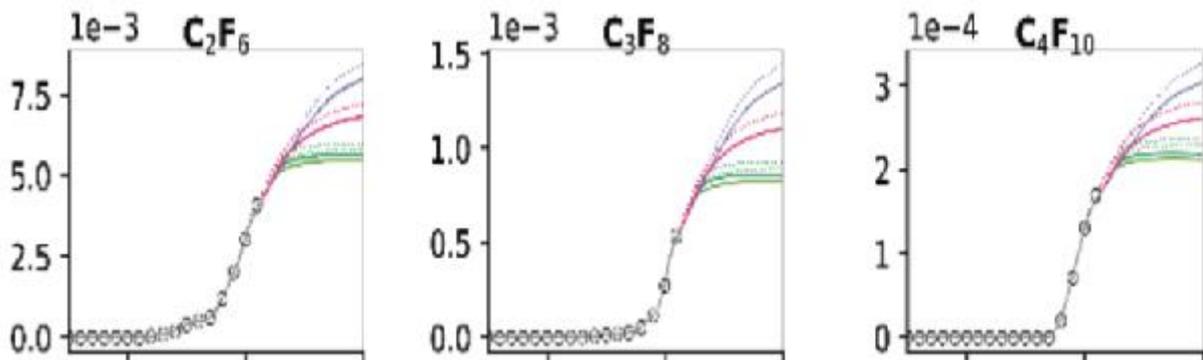
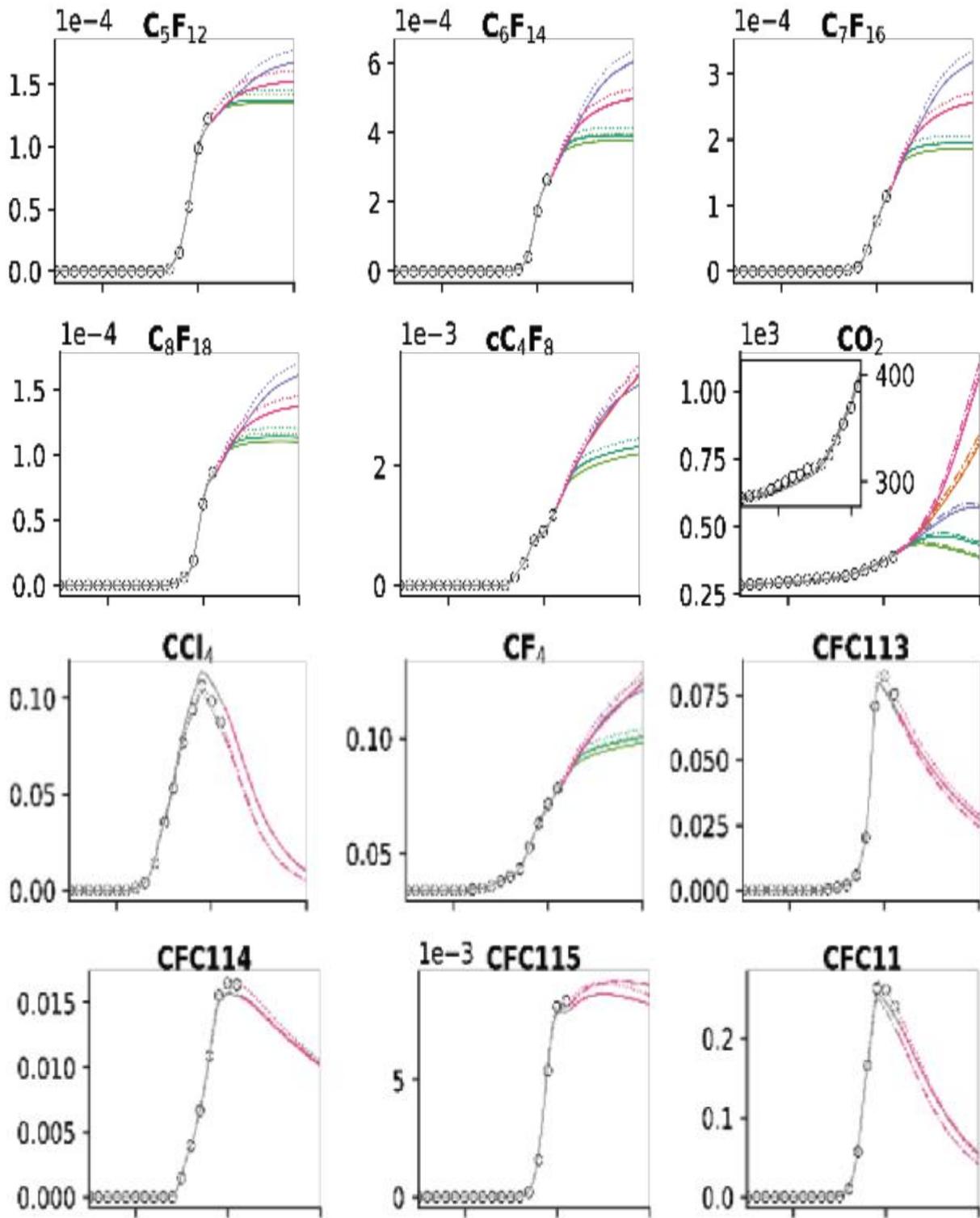
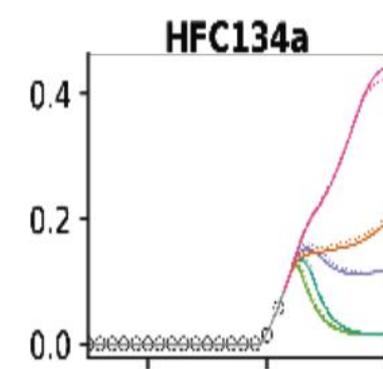
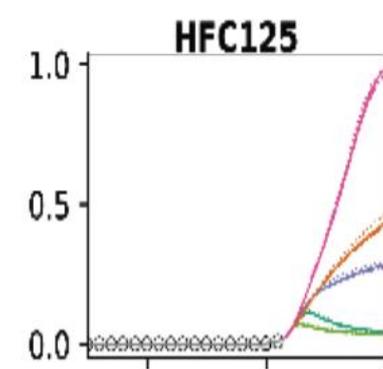
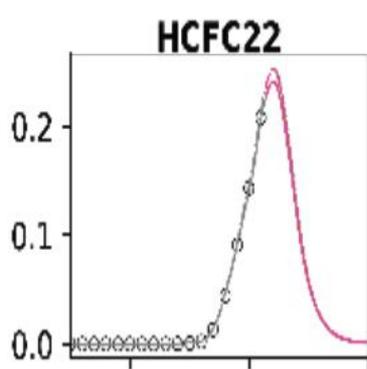
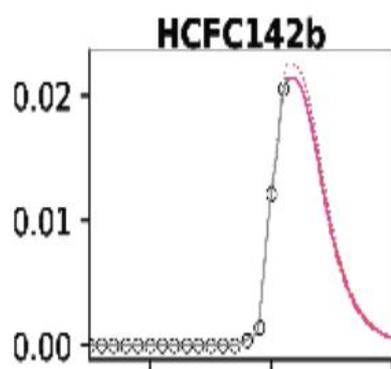
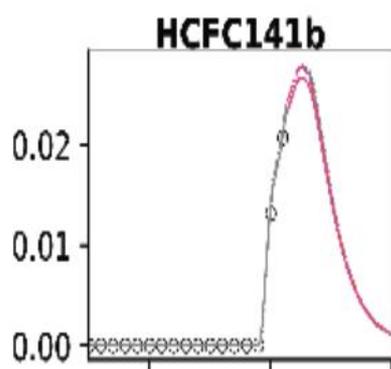
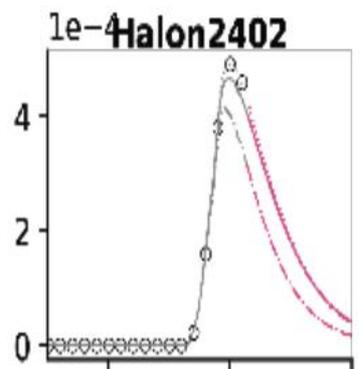
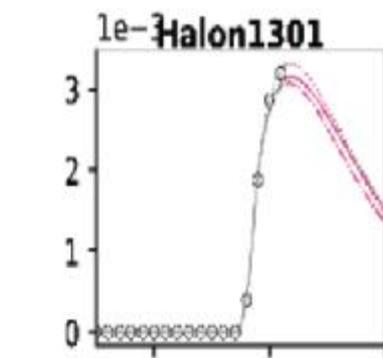
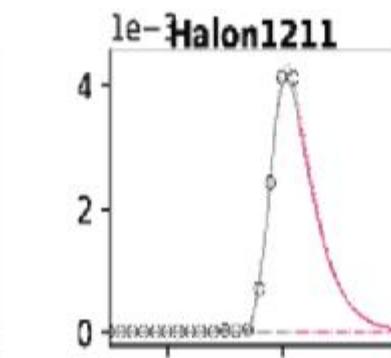
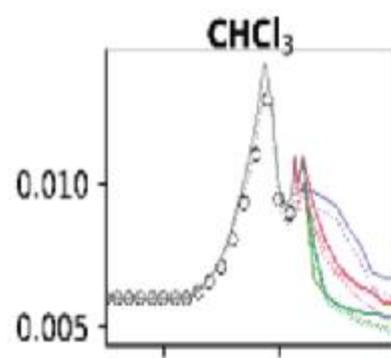
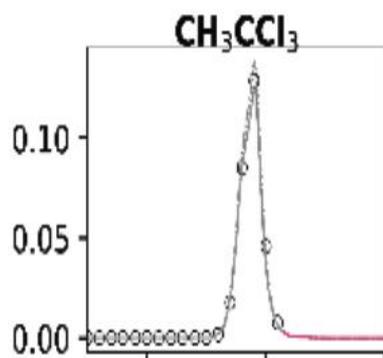
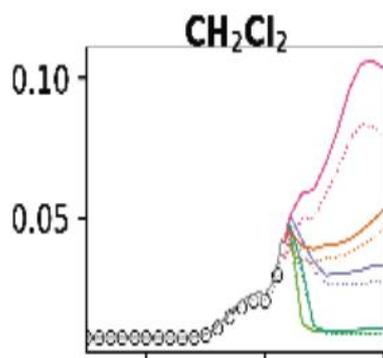
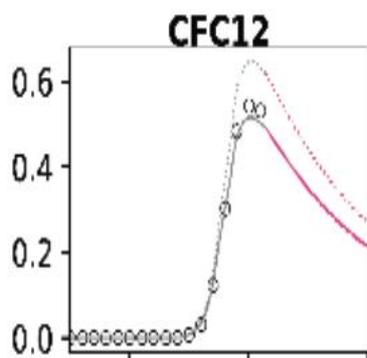


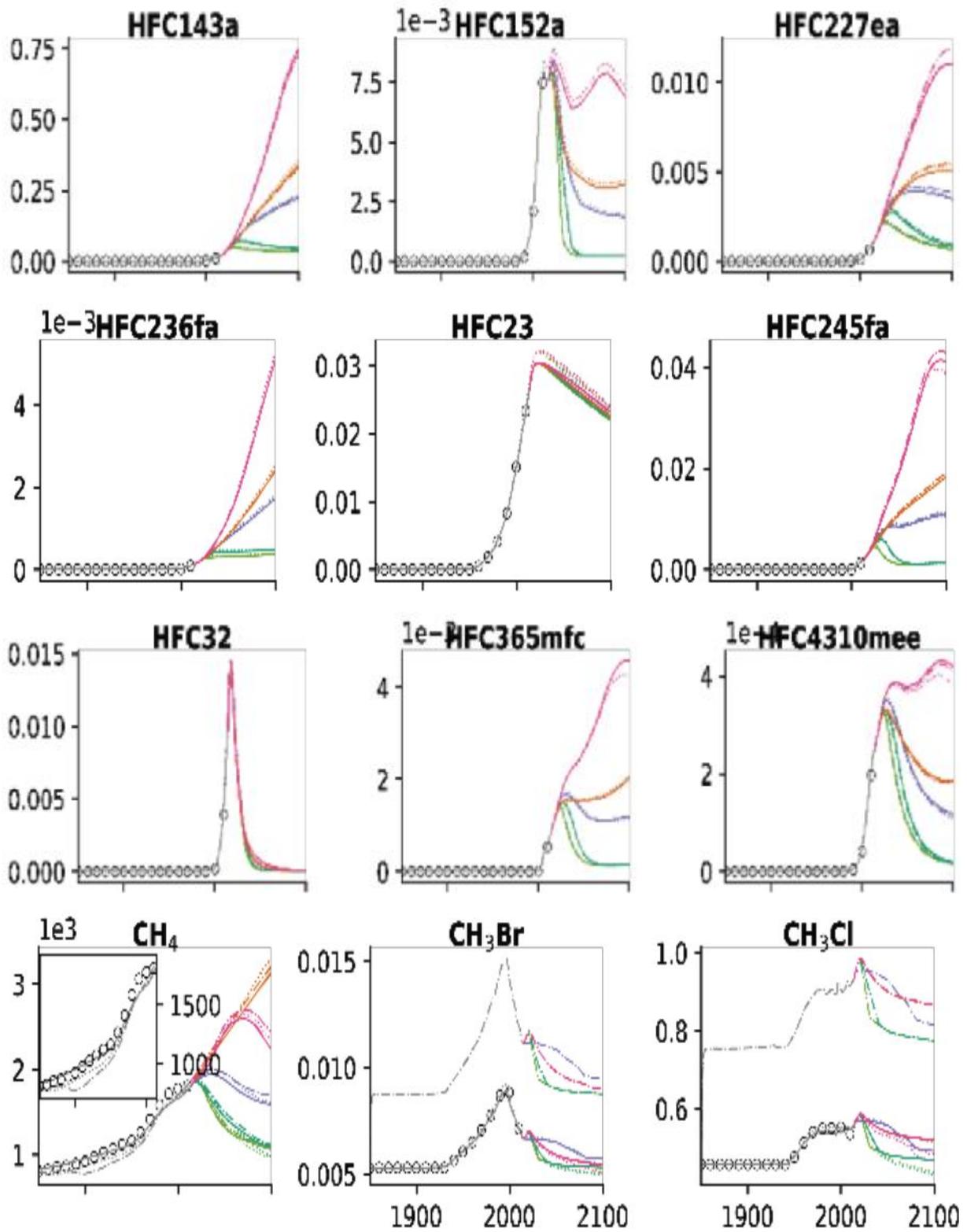
Рис. 2.2. Моделирование схемы DFWPEE

На рис. 2.3 представлена информация о показателе надежности слияния данных, а также о допустимости неоднородного коэффициента ранжирования.









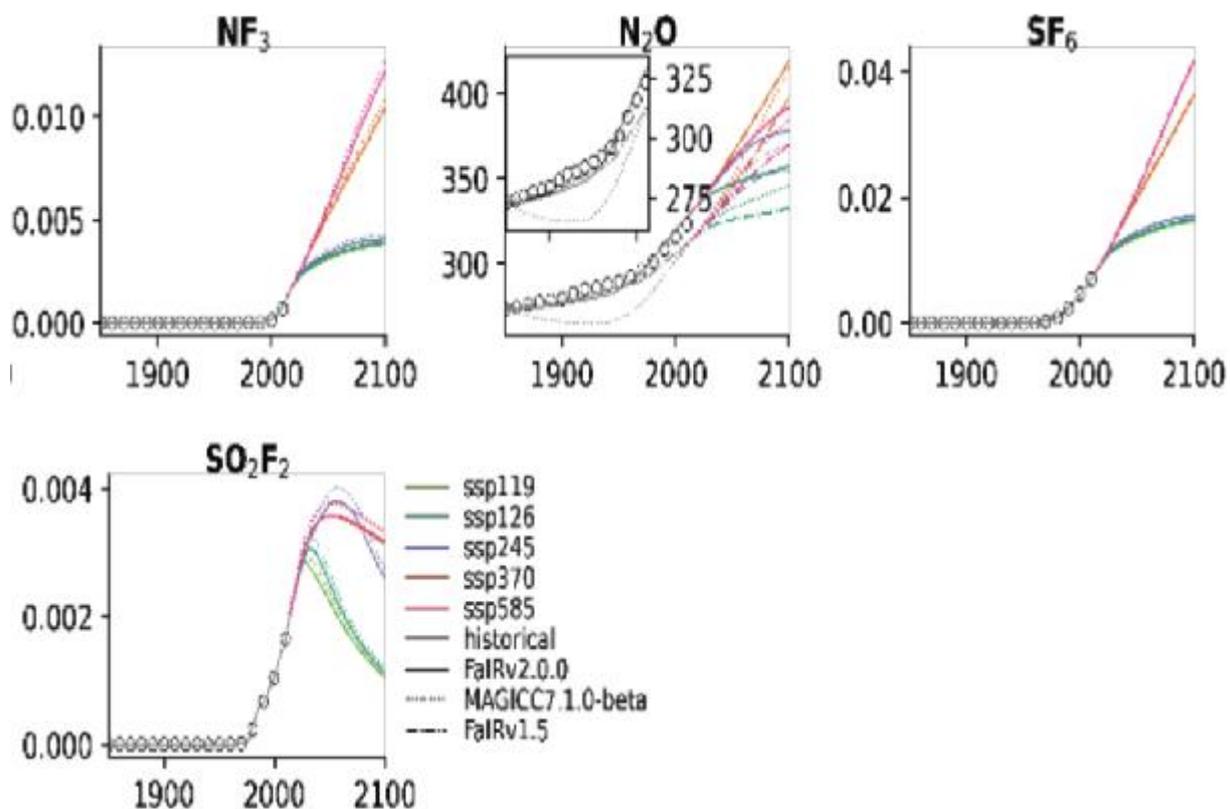


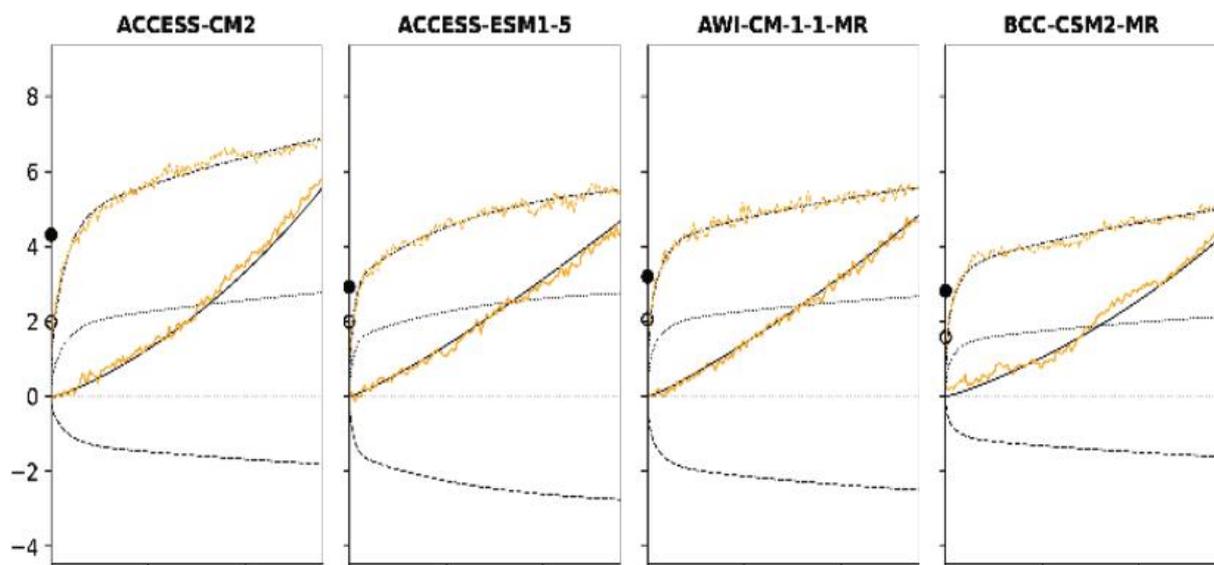
Рис. 2.3. Влияние DFWPEE на CPF (с использованием коэффициента неоднородности ранжирования)

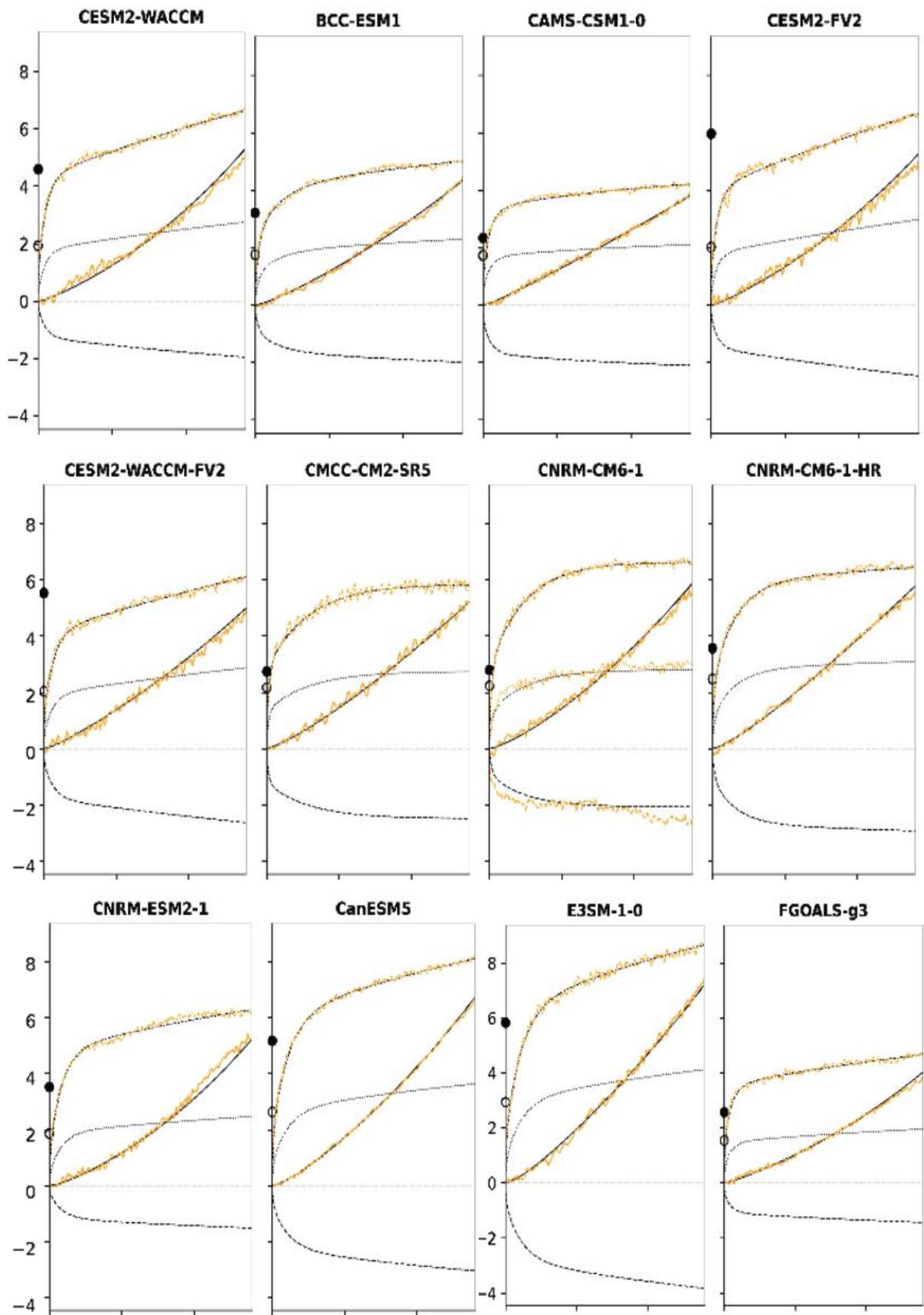
Этот материал относится к нескольким аспектам энергетики. Из рисунка видно, что эффективность узла слияния возрастает по мере увеличения количества узлов, а также их индивидуальных уровней энергии. Это повышает вероятность получения уникально проиндексированного пакета данных, который представлен на рис. 2.3. Другим важным результатом является то, что гетерогенный фактор ранжирования сам выбирает узел слияния-кандидата, который обладает наибольшей реализуемостью, принимая во внимание энергию и общую плотность окрестностей. Это важный результат, поскольку он демонстрирует, как работает HRF. Как прямое следствие этого, успешная реализация схемы DFWPEE в сети WSN является однозначной.

На рис. 2.3 также показана зависимость от однородного коэффициента ранжирования и информации о показателе надежности данных. В

этом случае легко заметить, что увеличение значения коэффициента однородного ранжирования (CPF) в зависимости от расстояния, положения и общей вовлеченности узлов fusion повышает возможность уникальной индексации принимаемых пакетов данных. На рис. 2.3 также показано, как переменная DFWPEE повлияла на CPF (коэффициент однородного ранжирования).

На рис. 2.4 приведена иллюстрация перехода к однородному коэффициенту ранжирования, а также информация о показателе надежности данных в зависимости от расстояния, отделяющего узел слияния от приемника. На рис. 2.4 показано, что, несмотря на увеличение значения коэффициента независимости (IF), коэффициент однородности остается лучшим для кривых с CPF=3000 и TxD=30m по сравнению с кривыми с CPF=6000 и TxD=60m. Это связано с тем, что кривые с CPF=3000 и TxD=30m имеют более низкое значение коэффициента непосредственности (IF), чем кривые с CPF=6000 и T. Это остается актуальным, несмотря на то, что значения IF становятся все выше. Появление этого события свидетельствует о том, что DFWPEE использует более эффективный метод оптимизации.





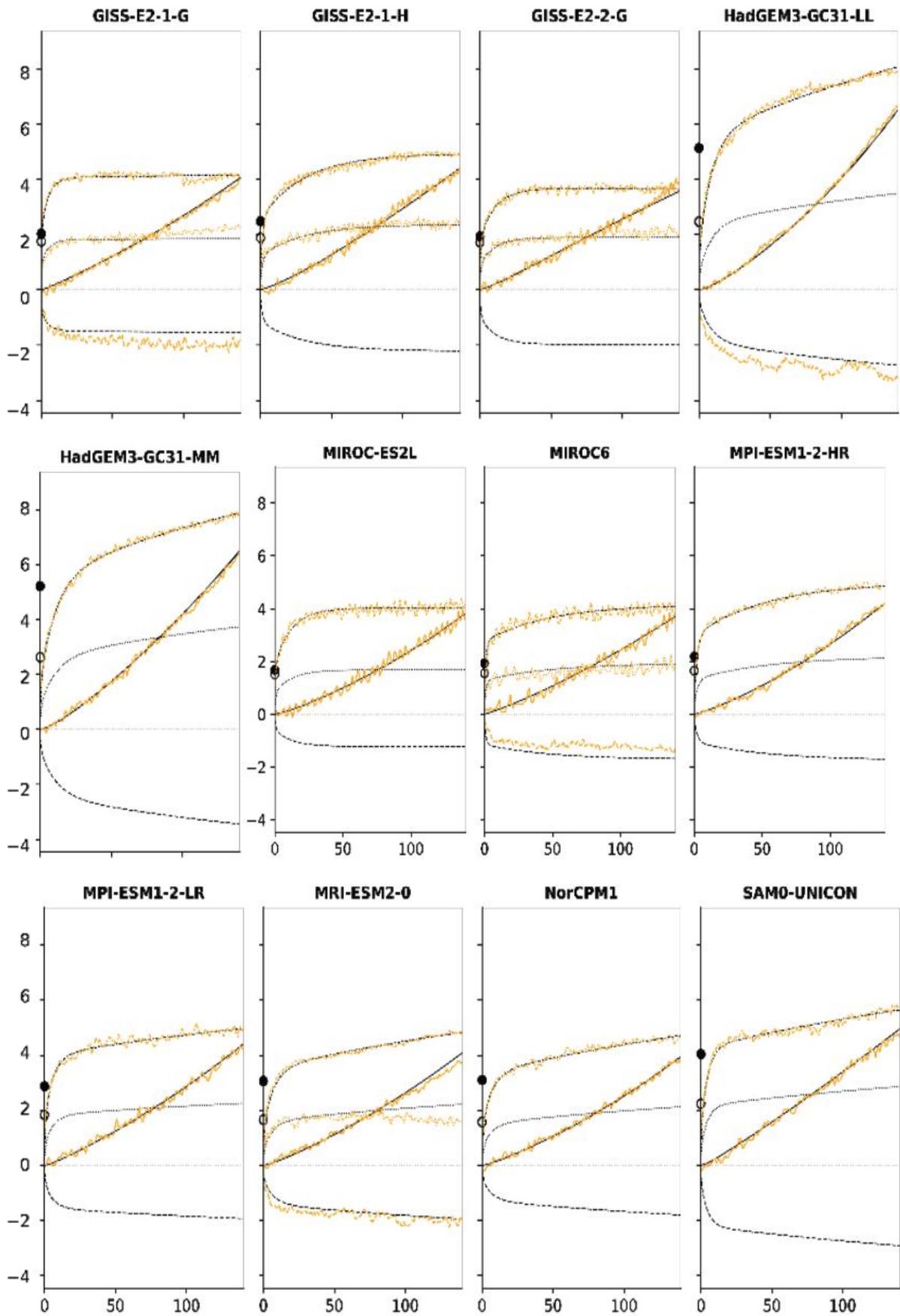


Рис. 2.4. Влияние DFWPEE на энергетику

Когда число узлов слияния увеличивается с 30 до 40, а затем до 50, можно увидеть, что коэффициент неоднородности ранжирования достигает своего максимального значения одновременно с увеличением общего числа узлов. Следовательно, разумно предположить, что вероятность индексации уникальных входящих данных возрастает пропорционально количеству присутствующего неоднородного компонента ранжирования. Это может быть выражено как "вероятность индексации уникальных входящих данных".

На рис. 2.5 показано влияние DFWPEE на CPF и TxD (коэффициент однородного ранжирования).

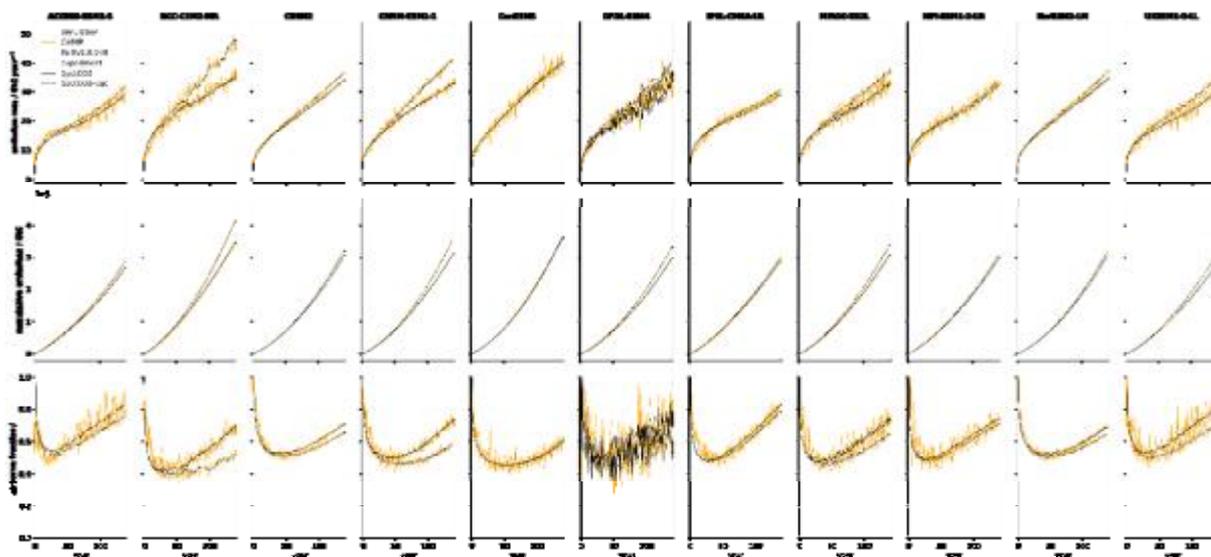


Рис. 2.5. Влияние DFWPEE на CPF и IF (коэффициент однородного ранжирования)

Информация о показателе перехода к однородному коэффициенту ранжирования в зависимости от расстояния до приемника еще раз выделена, и ее можно увидеть на рис. 2.6. В результате становится ясно, что предложенная система значительно оптимизировала коэффициент однородного ранжирования путем реализации недетерминированной логики. Это было сделано для того, чтобы сохранить максимальное количество энергии, ко-

торое можно было сохранить, а также для того, чтобы сделать лучший и более эффективный выбор надежного узла слияния при минимальных затратах на сеть. Это было сделано для того, чтобы максимально сократить объем работ, необходимых для поддержания работоспособности сети. Как следствие, описанная система обладает высокой эффективностью в плане обработки данных и даже может использоваться в среде, работающей в режиме реального времени.

Широко известный протокол LEACH используется наряду с протоколом DFWPEE, и был проведен сравнительный анализ обоих протоколов. Это сравнение зависит от параметров моделирования, которые были приняты во внимание ранее. На рис. 2.6 представлено сравнение количества активных узлов, обнаруженных с помощью предложенного метода и протокола LEACH. Эти результаты ясно указывают на превосходство предлагаемого метода в производительности по сравнению с протоколом LEACH. В течение первых нескольких сотен итераций протокола LEACH общее количество активных узлов начинает постепенно уменьшаться, но в конечном итоге эта тенденция прекращается. С другой стороны, разработанный протокол DFWPEE предусматривает постоянное и плавное уменьшение количества узлов, которые все еще активны. Согласно полученным результатам, DFWPEE может обеспечить максимальное увеличение общего срока службы беспроводной сенсорной сети.

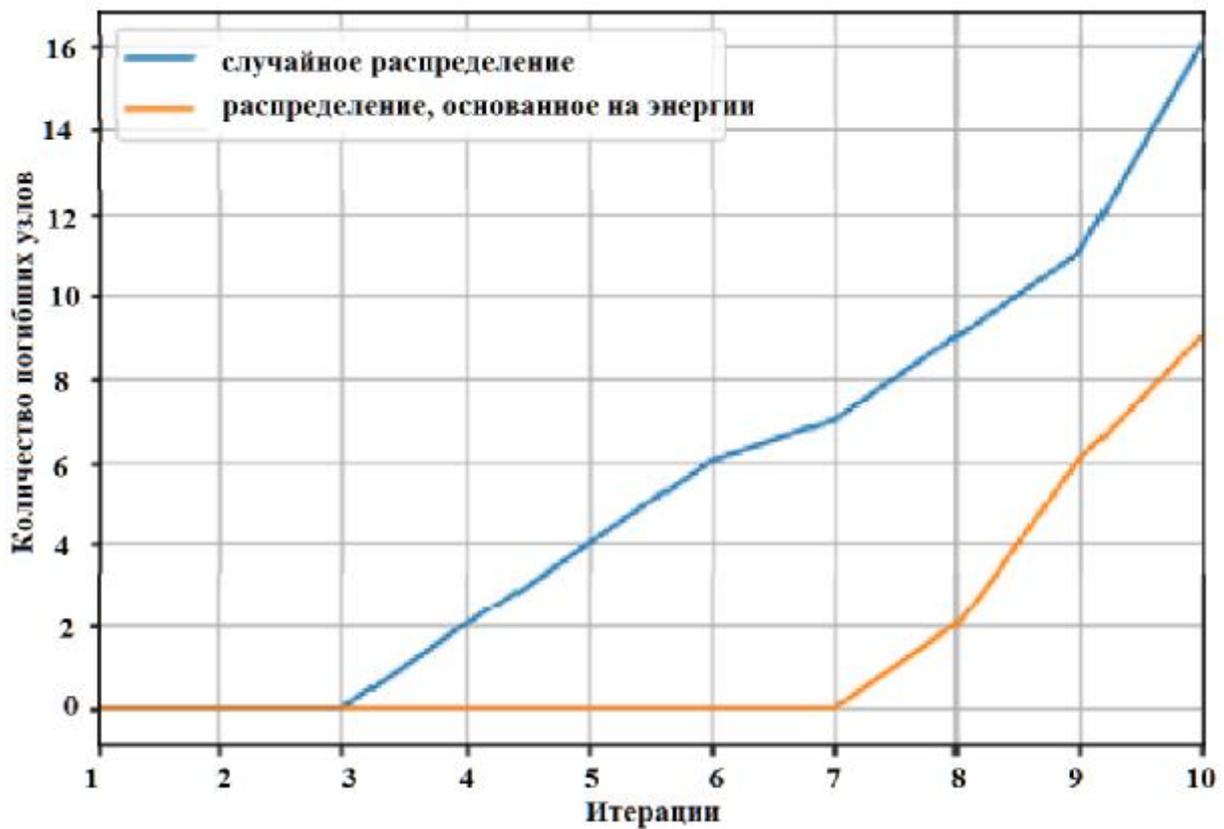


Рис. 2.6. Сравнение с LEACH (живые узлы и итерации)

Одно из наиболее существенных различий между методологиями внедрения LEACH и DFWPEE показано на рис. 2.7. Можно видеть, что наблюдается резкое уменьшение наклона кривой, которая используется для представления остаточной энергии при LEACH, при количестве итераций ниже 300 и около него. Важно, что предложенный ENLPL отличается плавным затуханием вплоть до полного через около 700 итераций. В качестве очевидного вывода отметим, что PEEDF потенциально ведет к улучшению энергосбережения. Дело в том, что быстрый разряд датчика гарантируется при использовании протокола LEACH элемента, а предложенный протокол, напротив, максимизирует остаточную энергию.

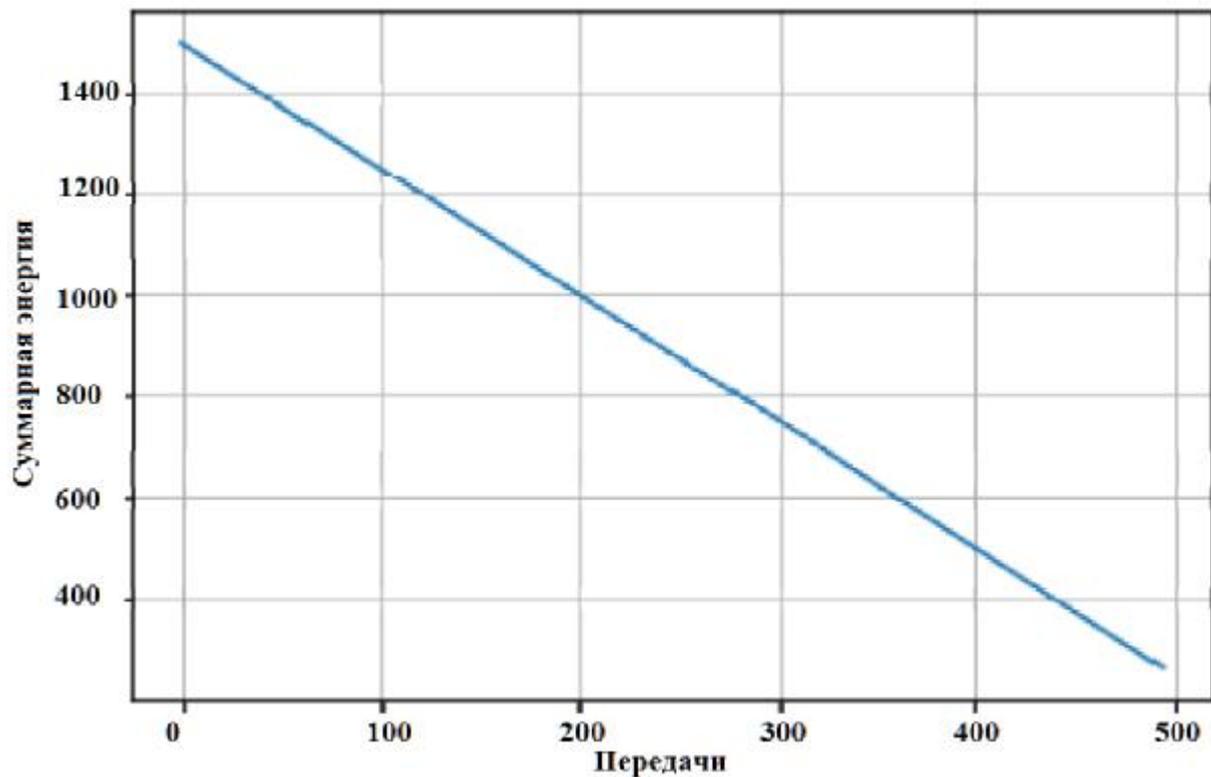


Рис. 2.7. Остаточная потенциальная энергия узлов

Следовательно, одним из наиболее важных аспектов сравнительного анализа должен быть сбор данных о параметре дисперсии мощности. Было замечено, что предложенный протокол имеет среднее значение FND (гибель первого узла), равное примерно 656 из 700 циклов наблюдения, в то время как протокол LEACH имеет только 378 из 700 циклов наблюдения. Зависимость мощности от общего числа 700 циклов наблюдения показана на рис. 2.8.

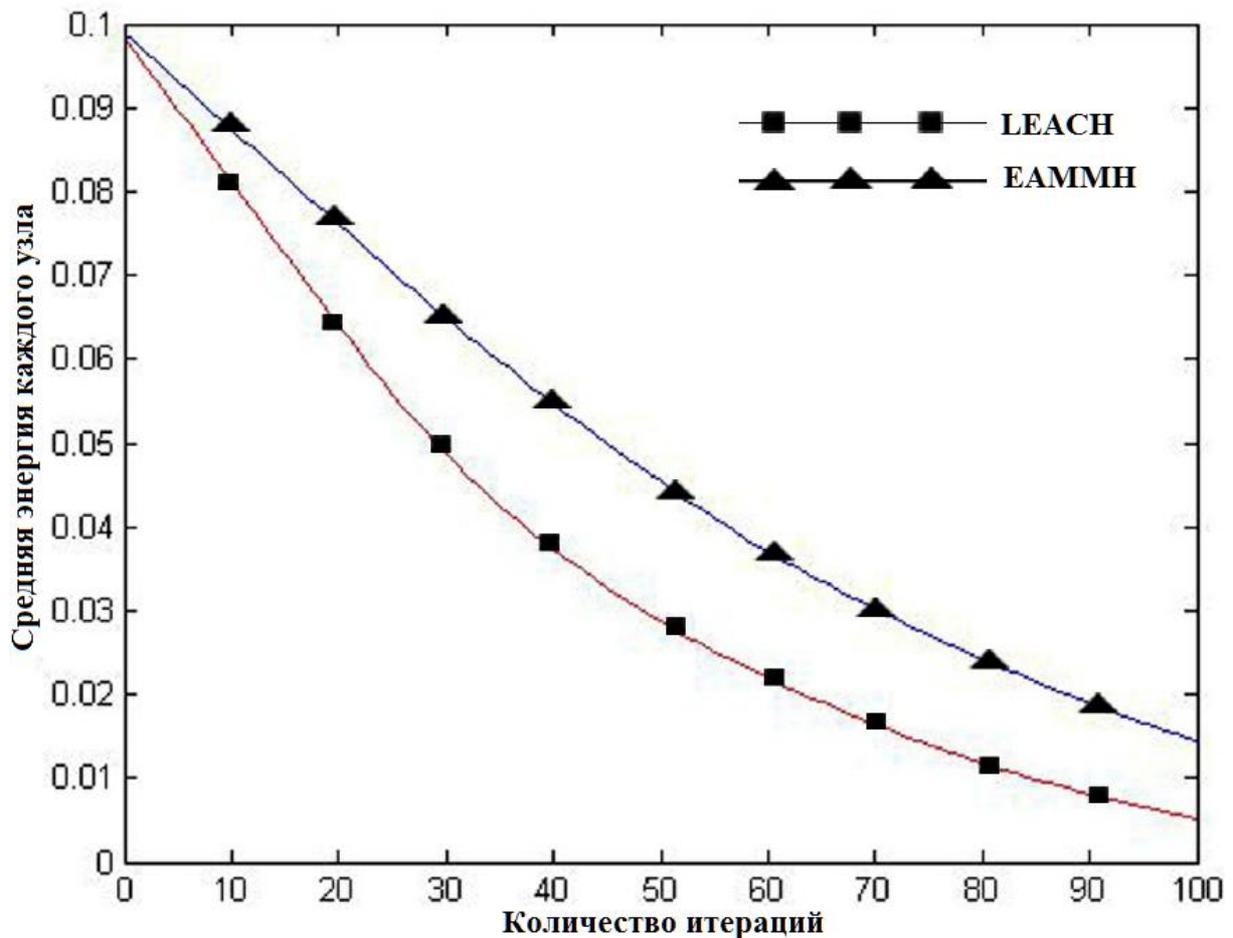


Рис. 2.8. Иллюстрирует разницу в мощности в зависимости от количества итераций

Для визуализации и сравнения эффективности предложенного алгоритма при анализе энергоэффективности системы с использованием протокола LEACH будем учитывать параметр FND (отказ первого узла). Решаемая задача при этом - максимизация совокупного срока службы сети.

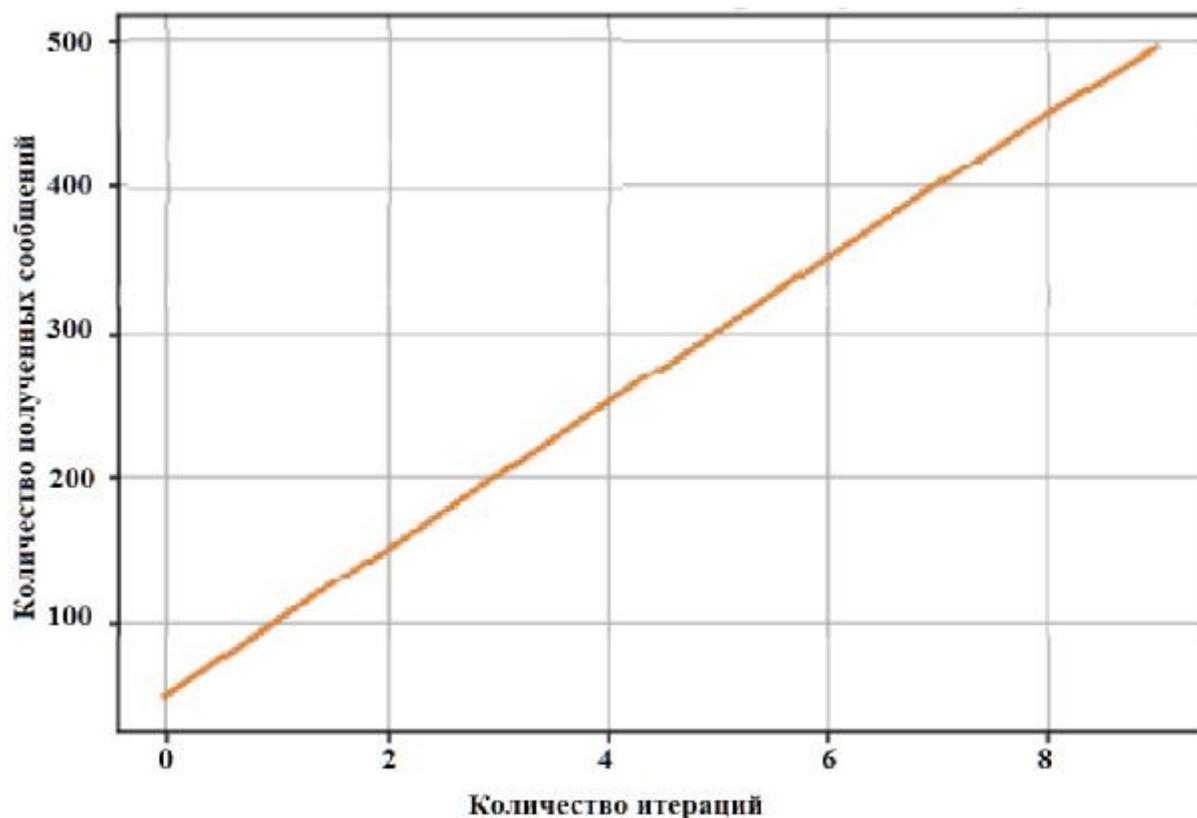


Рис. 2.9. Данные для первого события отказа узла как для предложенного метода, так и для метода LEACH

Из рис. 2.9 становится ясно, что предложенный метод действительно помогает датчикам стабильно использовать электроэнергию.

Также можно сказать, что мощность, выделяемая датчикам в обычном протоколе LEACH, часто отклоняется для отдельных датчиков, которые имеют чрезвычайно высокую мощность, в то время как другие датчики теряют энергию до тех пор, пока у них не останется 0 джоулей. Однако очевидно, что проблемы с протоколом LEACH, которые обсуждались ранее, были преодолены с помощью предлагаемой системы. Согласно протоколам LEACH, датчики разрядятся примерно через 370 циклов, в то время как в предлагаемой системе заряд не иссякнет примерно через 690 циклов.

## 2.5. Выводы

Для задач маршрутизации в WSN энергопотребление есть наиболее

существенный фактор. Именно поэтому изучается механизм кластеризации на основе LEACH. Метод LEACH обеспечивает поиск центра кластера на основе вероятностного подхода. С другой стороны, LEACH просто основан на вероятностной структуре, которая утверждает, что некоторые кластеры могут быть топологически близки друг к другу и находиться на границе WSN. Понятно, что использование таких кластеров неоптимально для улучшения энергоэффективности.

Протокол ENLPL устраняет сложности при коммуникации на основе вероятностной логики и при решении задачи выбора основных факторов при поиске центра кластера, таких, как уровень энергии, норма трансмиссии и топологическая плотность узлов.

Наше внимание было привлечено к тому факту, что генерация кластеров в протоколе LEACH полностью зависит от универсальных знаний. Без сомнения, для того, чтобы получить доступ к этой информации, требуются определенные стратегии маршрутизации в полном объеме. С этой точки зрения, протокол LEACH можно рассматривать как полураспределенный протокол для беспроводных сенсорных сетей. Другой существенной проблемой протокола LEACH является его произвольный метод выбора лидеров кластеров. Такой подход не гарантирует, что будет выбрано необходимое количество лидеров кластера, и не гарантирует, что выбранные лидеры кластера будут равномерно распределены. Ни то, ни другое нельзя гарантировать.

## **Литература к главе 2**

2.1. Afzaal A.A., Kamal A. An intelligent neural-WSN based schema for energy resources forecast// International Journal of Advanced Science and Technology, 2011, vol. 33, pp. 121130.

2.2. Shiny A.V.A., Nagarajan V. Energy Efficient Routing Protocol for Mobile WSN// International Journal of Computer Applications, 2012, vol. 43, no. 2, pp. 1-5.

2.3. Kumar A.S., Kanthan L.N., Priya S.N., Amala G, Revathi A., Revathi R. A Survey on Flow Balanced Routing in WSN// International Journal Of Technology Enhancements And Emerging Engineering Research, 2014, vol. 2, no. 9.

2.4. Balamurugan A. An energy efficient fitness based routing protocol in WSN, ICTACT Journal on Communication Technology, 2014, vol. 5, no. 1, pp. 894-899.

2.5. Kirankumar B.B., Ajayakumar C.K., Poornima M.C., Gururaj K. Energy Efficient Multipath Routing Protocol with Guaranteed Data Delivery for WSN// International Journal of Information and Education Technology, 2014, vol. 4, no.5, pp. 430-435.

2.6. Ben-Othman J., Bashir Y. Energy efficient and QoS based routing protocol for WSN// Journal of Parallel and Distributed Computing, 2010, vol. 70, no. 8, pp. 849-857.

2.7. Bokare M., Ralegaonkar A. WSN: A Promising Approach for Distributed Sensing Task// Excel Journal of Engineering Technology and Management Science, 2012, vol. 1, no. 1, pp. 1-9.

2.8. Boonsawat V., Jurarat E., Kulwadee B., Somsak K. 'XBee WSN for temperature monitoring// The second conference on application research and development, Chon Buri, Thailand. 2010, 117.

2.9. Chakraborty A., Swarup K.M., Mrinal K.N. A Genetic algorithm inspired routing protocol for WSN// International Journal of Computational Intelligence Theory and Practice, 2011, vol. 6, no. 1, pp. 1-8.

2.10. Chaudhary S., Singh N., Pathak A., Vatsa A.K. Energy Efficient Techniques for Data aggregation and collection in WSN// Int. J. Comp. Sci. Eng. Appl., 2012, vol. 2, no. 4, pp.37-40.

2.11. Chen R.-C., Haung Y.-F., Hsieh C.-F. 'Ranger intrusion detection system for WSN with sybil attack based on ontology// New Aspects of Applied Informatics, Biomedical Electronics and Informatics and Communications, WSEAS International Conference on Applied Informatics and Communications, 2010, pp. 176-180.

2.12. Chitra S.M., Vinoba V. Padmavathy T.Y. Link Reliability Routing Protocol in WSN using Game Theory Approach// Proceedings of the International Conference on Applied Mathematics and Theoretical Computer Science, 2013, pp. 235-238.

2.13. Cobb M. Powering High Speed Analog-to-Digital Converters with Switching Power Supplies// Power Management Design Line, Analog Devices, 2009, May 27.

2.14. Elrahim A.G.A., Hussein E.A., Salwa R.E., Magdy I.M. An energy aware WSN geographic routing protocol// Universal Journal of Computer Science and Engineering Technology, 2010, vol. 1, no. 2, pp. 105-111.

2.15. Enami N., Reza A.M., Kouros D., Hoseini M. Neural network based energy efficiency in WSN: A survey// International Journal of Computer Science & Engineering Survey, 2010, vol. 1, no. 1, pp. 39-53.

2.16. Fan G.J., Jin S.Y. Coverage problem in WSN: A survey// Journal of networks, 2010, vol. 5, no. 9, pp. 1033-1040.

2.17. Felemban E. Advanced border intrusion detection and surveillance using WSN technology', Int. J. Communications, Network and System Sciences, 2013, no. 6, pp. 251-259.

2.18. Flouri K., Beferull-Lozano B., Tsakalides P. Energy-Efficient Distributed Support Vector Machines for Wireless Sensor Networks// Proc. 2006 European Workshop on Wireless Sensor Networks (EWSN '06), Zurich, Switzerland, 2006, February 13-15.

### **3. Повышение энергоэффективности протокола управления потоками данных со скрытой марковской моделью в беспроводных сенсорных сетях**

#### **3.1. Предлагаемая архитектура**

Предлагаемая архитектура в основном состоит из трех модулей, обозначенных пунктирными линиями на рис. 3.1. Это проектирование сети, инициализация случайных параметров и поиск оптимального скрытого пути. Все эти шаги позволяют системе прогнозировать наилучшие параметры управления потоками данных.

##### ***3.1.1. Общая схема системы***

Подход к минимизации энергопотребления в предлагаемом стохастическом процессе управления потоками данных проиллюстрирован на рис. 3.1.

Первый модуль в блок-схеме позволяет наблюдать за поведением существующей сети и извлекать уроки из него. В этом блоке, функции управления потоками данных собираются на основе традиционных алгоритмов и извлекают соответствующую информацию. Второй модуль состоит из получения функций, подготовленных в первом модуле, для инициализации параметров скрытой адаптированной марковской модели. Он вычисляет их с помощью эмпирических формул. Третий модуль - это блок принятия решений, который позволяет выбрать оптимальный путь из числа кратчайших, повторно оценить новые параметры в системе и найти более эффективный процесс управления потоками данных.

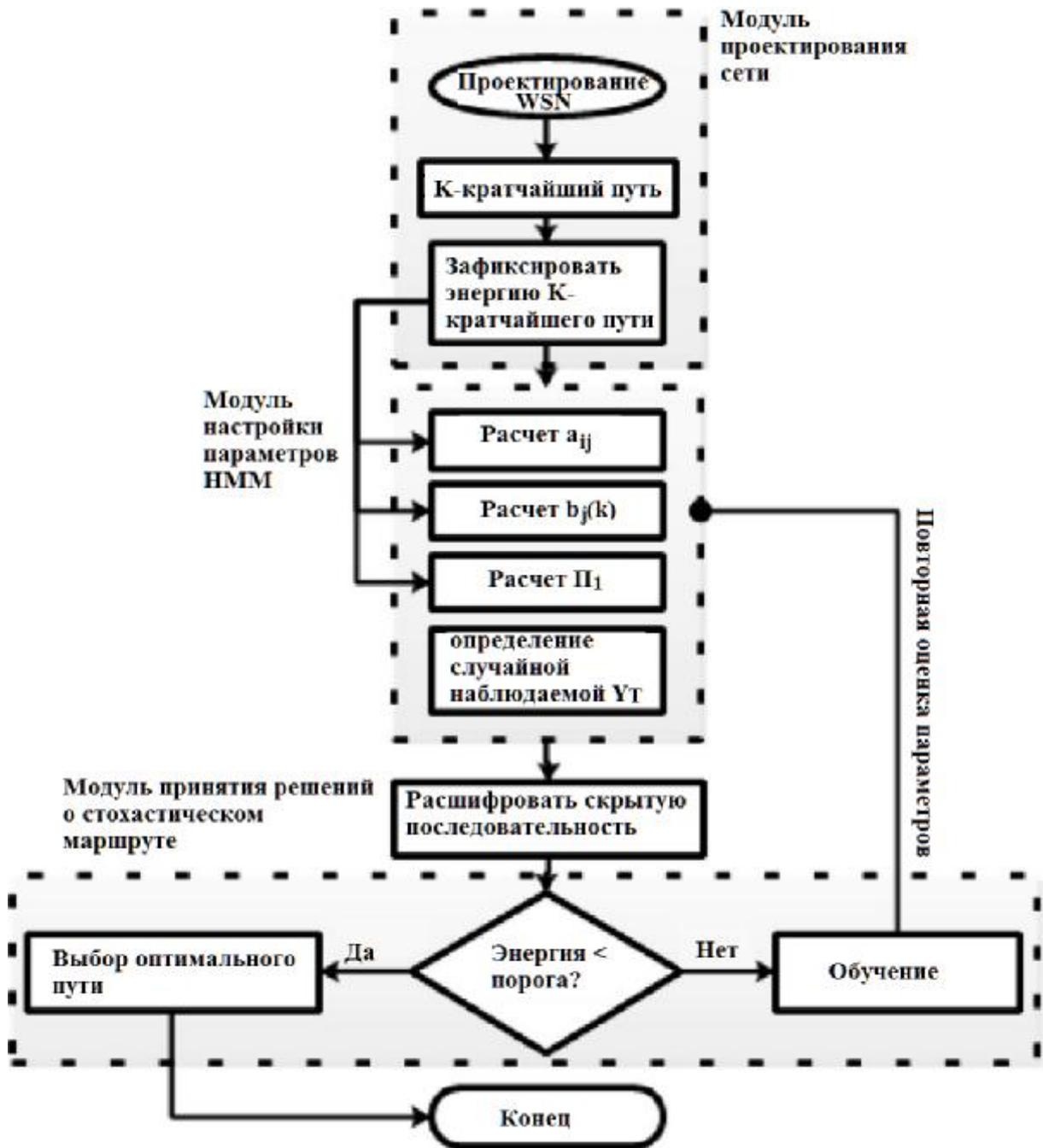


Рис. 3.1. Структура основных модулей общей системы стохастической управления потоками данных

### Алгоритм 3.1. Стохастический поиск кратчайшего пути

Требуется:  $G=(V, E)$

Задача: Найти K-Кратчайший путь

1: Общая энергия пути:  $E_k = \sum_{i=1}^n e_{k_i}$

- 2: Общий вес пути:  $Cost_k = \sum_{i=1}^n \cos t_{k_i}$
- 3: Вычисляем начальные параметры (А, В, П)
- 4: Фиксируем символы  $Y_t$  для наблюдения за временными рядами
- 5: Применяем алгоритмы "Вперед-назад" и "Витерби"
- 6: **If** Достигнуто Пороговое значение,
- 7: **Then** Выбираем оптимальный маршрут
- 8: **Else**
- 9: Обучаем модель
- 10: Возврат (3)
- 11: **EndIF**

В алгоритме 3.1 К кратчайших путей были вычислены с использованием процесса, предоставленного алгоритмом Йена [3.41], на основе стоимости ребер. В строке 1 общее потребление энергии на каждом кратчайшем пути составляет  $E_k = \sum_{i=1}^n e_{k_i}$ . В строке 2 стохастическая модель инициализируется эмпирическим путем. В строке 3 используются уравнения (3.1)-(3.3) соответственно. В строке 4 используется уравнение (3.5), строка 5 соответствует уравнениям (3.6)-(3.13), а строка 7 соответствует уравнению (3.14).

Образец магистральной сетевой топологии, использованный для демонстрации модели, представлен в следующем разделе.

### ***3.1.2. Примерная топология сети***

Рассмотрим  $N$  сенсорных узлов, развернутых в прямоугольной области  $M \times N$ , обозначенной как типовой граф  $G=(V, E)$  [3.42]. Здесь  $V=\{v_1, v_2, \dots, v_n\}$  представляет вершин количеством  $|V|$ , а  $E=\{e_1, e_2, \dots, e_m\}$  представляет ребра между узлами датчиков количеством  $|E|$ . Путь или маршрут

от узла-источника  $n_i \in V$  к узлу-получателю  $n_j \in V$  представляет собой последовательность узлов, выбранных протоколом управления потоками данных для пересылки данных.  $Link(i,j) \in E$  включается в маршрут и оценивается с точки зрения энергозатрат, потребляемых содержащими узлами.

Рис. 3.2 используется в качестве демонстрационной модели для иллюстрации предлагаемой методологии. Узлы обозначены символами  $N_k$ , где  $k$  - ссылка на узел в топологии.

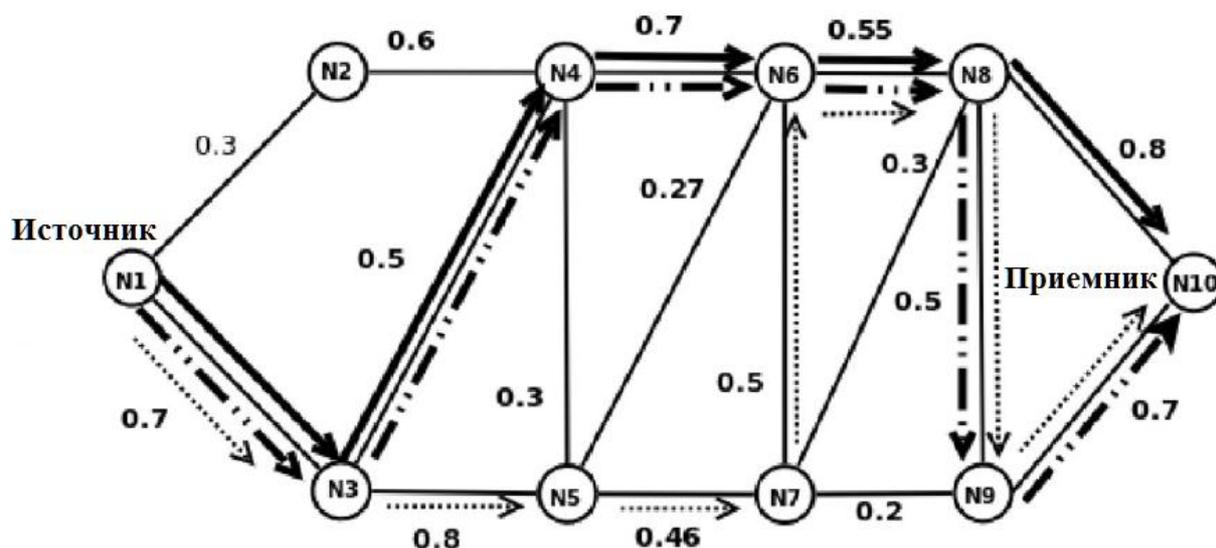


Рис. 3.2. Пример сетевой топологии, используемой для демонстрации протокола управления потоками данных по стохастическим кратчайшим путям: — ребро;  $\Rightarrow$   $E_1$  (маршрут 1);  $\cdots \Rightarrow$   $E_2$  (маршрут 2);  $\cdots \cdots \Rightarrow$   $E_3$  (маршрут 3)

Простые линии обозначают общие ребра связи, жирные черточки, пунктирные стрелки и пунктирные стрелки - точки, обозначающие рассматриваемые пути. Значения энергии всех связей (ребер), которые регистрируются с помощью работающей сети датчиков, находятся в диапазоне 0,2-0,8. Эти значения используются для инициализации параметров стохастической модели управления потоками данных НММ.

Пороговое значение энергии фиксируется на уровне 0,4 [3.43] для

выбора ребер, участвующих в процессе управления потоками данных.

### 3.1.3. Эмпирическая оценка исходных параметров НММ

Входные данные нашей системы используют численные значения энергий, показанные на предыдущем графе в качестве примера, для вычисления соответствующих коэффициентов на основе предложенного эмпирического подхода.

Формулы, используемые для получения начальных параметров системы, используют минимальное значение энергопотребления  $E^{\min}$  в каждом канале. Каждый передаваемый пакет содержит минимальное количество энергии, необходимое для передачи информации по сети.

Минимальная энергия пути - это минимальное потребление всеми узлами на выбранном пути [3.40].

Ниже приведены формулы для инициализации модели.

$$\rho_i = \frac{E_i^{\min}}{\sum_{j=1}^3 E_j^{\min}} \quad (3.1)$$

$$a_{ij} = \frac{\left| \frac{E_i^{\min} - E_j^{\min}}{\sum_{i,j=1}^3 (E_i^{\min} - E_i^{\min})} \right|, i \neq j, E_i^{\min} \neq E_j^{\min}}{\left| \frac{E_i^{\min} - E_j^{\min}}{\sum_{i,j=1}^3 (E_i^{\min} - E_i^{\min})} \right|, i \neq j, E_i^{\min} \neq E_j^{\min}} \quad (3.2)$$

Уравнения (3.1) и (3.2) показывают стратегию, основанную на эмпирической энергии системы, для получения всех соответствующих коэффициентов.

Распределение в уравнении (3.3) получено в соответствии с традици-

онной стоимостью каждого пути.

$$b_{E_j}(v_k) = \frac{\left| \cos t_i - \cos t_j \right|}{\left| \cos t_i - \overset{\circ}{a} \prod_{j=1}^3 \cos t_j \right|}, i \neq j, \cos t_i \neq \cos t_j \quad (3.3)$$

$$1 - \overset{\circ}{a} \prod_{k=1}^3 b_j(k), j = k, E_j^{-1} v_k$$

Уравнения (3.1), (3.2) и (3.3) должны учитывать предположения о вероятностях Маркова, выраженные в уравнениях (1.3), (1.5) и (1.7).

### 3.2. Стохастическое решение о маршруте

В этом разделе описывается, как обычные параметры управления потоками данных AODV (протокол динамической маршрутизации для мобильных ad-hoc сетей (MANET) и других беспроводных сетей) используются для стохастического поиска кратчайших путей [3.44]. На этом этапе система использует численное значение эмпирического расчета для определения наилучшей скрытой последовательности в соответствии с исходной моделью.

Декодирование производится с использованием алгоритма Витерби, и предлагаемая система оценивает, требуется ли шаг обучения с использованием метода Баума-Уэлча для оптимизации модели.

#### 3.2.1. Предложенная полностью связанная модель

Чтобы определить нашу стохастическую модель, мы запускаем последовательность  $Y$  в момент времени  $T=10$  для наблюдаемых символов. Каждый символ  $Y_t$  представляет значение энергии, которое, вероятно, наблюдается на пути в это время.

Предположим, что мы наблюдаем следующую короткую последовательность (3.4) уровней (3.5):

$$Y = v_1, v_2, v_3, v_3, v_2, v_1, v_1, v_2, v_3, v_1 \quad (3.4)$$

$$Y = L, M, H, H, M, L, L, M, H, L \quad (3.5)$$

Последовательность  $Y$  представляет собой выбранные случайным образом символы наблюдения, где каждый символ определяет уровень энергий в системе [3.45]. Вероятность наблюдения последовательности, заданной моделью  $I$ , оценивается в Прямом алгоритме. Каждый символ  $v_k$  иллюстрирует энергетический уровень, на котором может находиться каждое состояние:

$$V = v_1 = \text{Low}(L); v_2 = \text{Medium}(M); v_3 = \text{High}(H).$$

На рис. 3.3 коэффициенты перехода между состояниями не равны нулю; именно поэтому предложенная НММ является взаимосвязанной эргодической моделью. Коэффициенты  $a_{ij}$  представляют собой переходы между состояниями. Выбор скрытых состояний не является фиксированным и может быть скорректирован в соответствии с потребностями системы и соответствовать количеству начальных кратчайших путей.

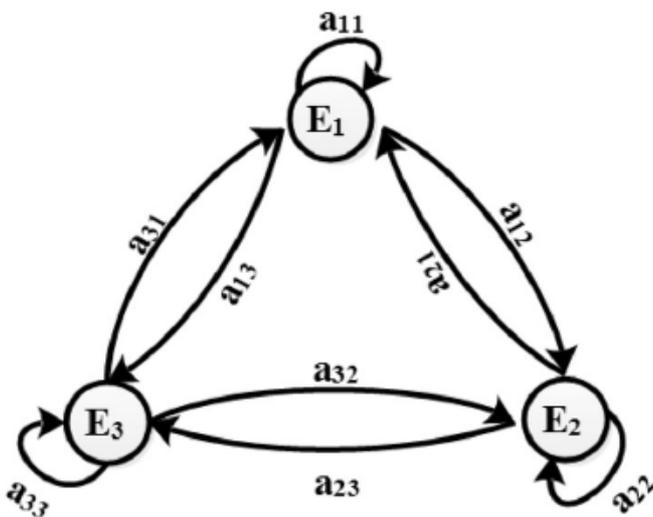


Рис. 3.3. Эргодический полностью связанный граф НММ, используемый в протоколе стохастического управления потоками данных

Наглядная и подробная иллюстрация полностью подключенной модели приведена на рис. 3.8.

Алгоритм Витерби используется для поиска оптимальной скрытой последовательности. Учитывая последовательность наблюдений и модель НММ, алгоритм возвращает путь к состоянию через НММ, который присваивает последовательности наблюдений максимальное правдоподобие. Мы используем следующее обозначение:  $P(X, Y/l)$ .

Определим величину  $d_t(j)$  для  $j=1\dots N$ , для  $t=1\dots T$ ,  $d_t(j)$  - это максимальная вероятность, учитывая количество  $k$  параметров, того, чтобы пройти через последовательность состояний  $X_{1\otimes t}$ , которая заканчивается в  $E_j$  в момент времени  $t$ , и наблюдать последовательность  $Y_{1\otimes t}$ :

$$d_t(j) = \text{Max}_{1\otimes t\otimes 1} P(X_{1\otimes t\otimes 1}, X_t=E_j, Y_{1\otimes t} / l) \quad (3.6)$$

Из состояния  $E_j$  в момент времени  $t$  и состояния  $E_i$  в момент времени  $t_1$  при  $t=2\dots T$ ,  $d_t(j)$  - максимальная вероятность того, что пути заканчиваются в состоянии  $E_j$ .

Методом индукции получаем:

$$d_{t+1}(j) = \max_{E_i} (a_{ij} d_t(i)) g_{E_j}(Y_{t+1}) \quad (3.7)$$

Таким образом, определение максимального аргумента  $\arg \max_{E_j} d_t(j)$

похоже на получение

$$\arg \max_{E_j} a_{ij} d_{t-1}(j) \quad (3.8)$$

Таким образом, принимая аргумент для максимума  $d_t(j)$  для всех  $t=1,2,\dots,10$  и  $j=1\dots N$ , мы получаем оптимальную последовательность состояний:

$$y_t(E_i) = \arg \max_{E_j} d_t(j) \quad (3.9)$$

Для любых  $t \in [1, T]$  и  $j \in [1, N]$  переменные  $d_t$  и  $\psi_t$  используются для оценки наилучшего скрытого пути с использованием подхода Витерби.

Пороговая вероятность для процесса обнаружения была установлена равной 0,4 (рис. 3.1).

### ***3.2.2. Обучение стохастической модели управления потоками данных***

Обучение осуществляется после первой оценки декодирования Витерби. Оно основано на исходных параметрах стохастической управления потоками данных, установленных на основе реальных характеристик сети и эмпирических формул. Алгоритм Баума-Уэлча используется для обучения исходной модели в соответствии с последовательностью наблюдаемых изменений энергий во время работы сети. От этапа обучения к этапу повышается вероятность модели и точность решения об управлении потоками данных.

Следующее предположение должно быть проверено после каждого этап обучения.

$$P(Y_{i+1}) > P(Y_i) \quad (3.10)$$

В уравнении (3.17)  $P(Y_i)$  - это вероятность исходной последовательности, а  $P(Y_{i+1})$  - это вероятность после обучающего набора, заданного нашей исходной стохастической моделью управления потоками данных. После определенного количества итераций стохастическая модель управления потоками данных должна достичь стабильности.

### ***3.2.3. Процесс принятия решения об управлении потоками данных***

Для создания процесса принятия решения необходим набор вычислений. В этих вычислениях использовались алгоритмы прямого и обратного хода, как показано в [3.46], для получения вероятностей всех возможных скрытых путей управления потоками данных, которые могли бы генерировать последовательность наблюдений  $Y$ .

Учитывая наблюдаемую последовательность  $Y_{1@t}=Y_1, \dots, Y_t$ , получим вероятности того, что первый  $Y_{1@t}$  завершится состоянием  $E_i$ , для  $i=1 \dots T-1$ . Рассмотрим прямые переменные  $\alpha_i$ , определенные по формулам:

$$a_t(i) = \prod_{E_i} g_{E_i}(Y_t) \quad (3.11)$$

$$a_{t+1}(j) = b_j(Y_{t+1}) \prod_{i=1}^N a_t(i) g_{ij} \quad (3.12)$$

Каждый элемент  $\alpha_i$  в уравнениях (3.11)–(3.12) представляет собой вероятность использования энергии пути после просмотра первого наблюдения  $t$ .

Например, мы использовали процесс завершения в прямом/обратном алгоритме для оценки вероятности соблюдения наилучшей последовательности с помощью уравнения (3.13):

$$P(Y/l) = \prod_{i=1}^{N=10} a_T(i) \quad (3.13)$$

В каждый момент времени  $t$  мы наблюдаем уровень энергопотребления, обозначенный символом  $v_k$ . Это означает, что уровень энергопотребления системы может зависеть от вероятности. Результаты этого расчета приведены в табл. 3.9.

Процесс завершения (см. (3.13)) используется для определения решения об управлении потоками данных, как показано в следующей таблице. В табл. 3.2 решение об управлении потоками данных основано на уравнениях (3.13) и (3.14). От тренировки к тренировке вероятность этого повышается, а вместе с ней и эффективность решения о маршруте. В данном случае решение о маршруте достигает порогового значения после четырех тренировок. В другом случае количество тренировок может отличаться.

Предварительная оценка помогает оценить вероятность управления потоками данных с использованием наилучшего кратчайшего пути с минимальными затратами энергии. Коэффициент принятия решения об

управлении потоками данных вычисляется на основе следующей формулы:

$$P(Y_{i+1})' e \quad (3.14)$$

где  $e$  - постоянный параметр, эквивалентный начальной энергии в каждом узле в начале моделирования. Он умножается на вероятность перехода от одной тренировки к другой для оптимизации параметров НММ. Результаты моделирования для этой работы представлены в следующем разделе.

Таблица 3.2

Решение о декодировании оптимальных значений скрытой последовательности

Обучение	Вероятность	Решение о маршруте
1	1.32e-09	1.32e-6
2	3.92e-05	0.039
3	2.24e-04	0.22
4	4.59e-04	0.5

### 3.3. Анализ результатов моделирования

#### 3.3.1. Настройка параметров

Сетевой симулятор версии 2 (NS2) используется для реализации системного сетевого графика для нашего стохастического алгоритма управления потоками данных [3.47]. NS2 - это хорошо известный объектно-ориентированный сетевой симулятор с высокой степенью параметризации, используемый научным сообществом, который позволяет экспериментировать со многими вариациями параметров сети и наблюдать относительные характеристики исследуемой системы. Его версия ns-2.35 была реализована в среде Ubuntu 14.04 LTS. Возможности моделирования приведены в табл. 3.3.

Таблица 3.3

Параметры моделирования

Параметр	Значение
Протоколы управления потоками данных	AODV/предлагаемые
Модель мобильности	равномерная
Размер сети	100 на 200 м
Начальная энергия	100 Дж
Мощность передачи $T_x$	0,05 Вт
Мощность приема $R_x$	0,024 Вт
Скорость	1 Мб/с
Количество узлов	10
Время	100 с

### ***3.3.2. Оценка вероятности управления потоками данных***

Во время реализации сети узлы используют определенный уровень энергии для обмена данными и участвуют в процедуре управления потоками данных. Сводная информация о потреблении типовых узлов приведена в табл. 3.4.

В табл. 3.4 представлено иллюстративное представление значений энергии для пяти узлов. В начале моделирования узлы имеют одинаковый начальный уровень энергии для запуска моделирования.

Входные значения модели НММ рассчитываются в соответствии с этими значениями энергий, чтобы иметь реалистичный прогноз при принятии окончательного решения о выборе кратчайшего пути.

В табл. 3.5 приведены данные об общем потреблении энергии на каждом  $k$  кратчайших путях от узла  $N_1$ , который является источником, и узла  $N_{10}$ , который является пунктом назначения.

В табл. 3.6 указаны такие характеристики управления потоками данных, как минимальное потребление энергии путями, затраты и узлы, вхо-

дящие в каждый путь. Мы выбираем ограниченное количество путей и их потребление для расшифровки скрытой последовательности и, при необходимости, обучаем систему для оптимизации выходных данных.

Таблица 3.4

Потребление энергии узлами-образцами

Идентификатор узла	2
4	6
8	10
Энергия (Дж)	29,98
29,82	28,71
30,45	27,91

Таблица 3.5

Потребление энергии между исходным и конечным узлами

Параметр управления потоками данных	Значение
Источник	$N_1$
Пункт назначения	$N_{10}$
Путь 1	1,2,4,5,7,9,10
$E^{\min 1}$	196,895 Дж
Путь 2	1,2,4,5,7,8,10
$E^{\min 2}$	197,339J
Путь 3	1,2,4,5,6,8,10
$E^{\min 3}$	198,246 Дж

Моделирование сгенерировало вычисленные скрытые состояния, как показано ниже, для случайной последовательности шагов наблюдения за общее время,  $T=10$ . Таким образом, мы выводим из этого вычисления наилучшую последовательность скрытых состояний

$D=E1, E2, E3, E3, E1, E1, E1, E2, E3, E1$

$$P(Y_{III})=3.92e^{-05}$$

которая определяется начальными параметрами. Алгоритм Витерби обеспечивает для нашей вероятностной системы уровень энергии  $E_1$ , соответствующий низкому уровню в момент времени  $t$  при минимальной энергии. Мы можем использовать соответствующий путь для увеличения срока службы сети. В качестве альтернативы, согласно подходу Витерби, нашим вероятным резервным маршрутом может быть  $E_3$ , чтобы избежать нехватки энергии и проблем с управлением потоками данных.

Таблица 3.6

Особенности управления потоками данных путей

Path <sub>k</sub>	Узлы	$E^{min_k}$	Стоимость <sub>k</sub>
1	1,2,4,5,7,9,10	196,895 Дж	2,56
2	1,2,4,5,7,8,10	197,339 Дж	2,76
3	1,2,4,5,6,8,10	198,246 Дж	2,82

Условные обозначения:  $E^{min_k}$  = минимальная энергия Path<sub>k</sub>

Повторная оценка параметров с помощью алгоритма Баума-Уэлча, использованного в предыдущем процессе обратного отслеживания Витерби, стабилизирует систему на оптимальном скрытом пути:

$D=E1, E2, E3, E3, E3, E1, E1, E1, E2, E3, E1$

$$P(Y_{III})=2.24e^{-045}$$

То же самое наблюдение в скрытой последовательности. Только вероятность получения начальной последовательности, заданной моделью, значительно повышается. Начальный маршрут -  $E1$  с пятью повторениями, а резервный -  $E3$  с тремя повторениями. Оценивается наилучшая общая вероятность управления потоками данных (BRP) использования энергоресурса в течение определенного времени. На рис. 3.4 показано вероятное

поведение системы при использовании энергии в начале управления потоками данных. Вероятность высока, учитывая временной ряд  $Y_T$  и модель  $k$ .

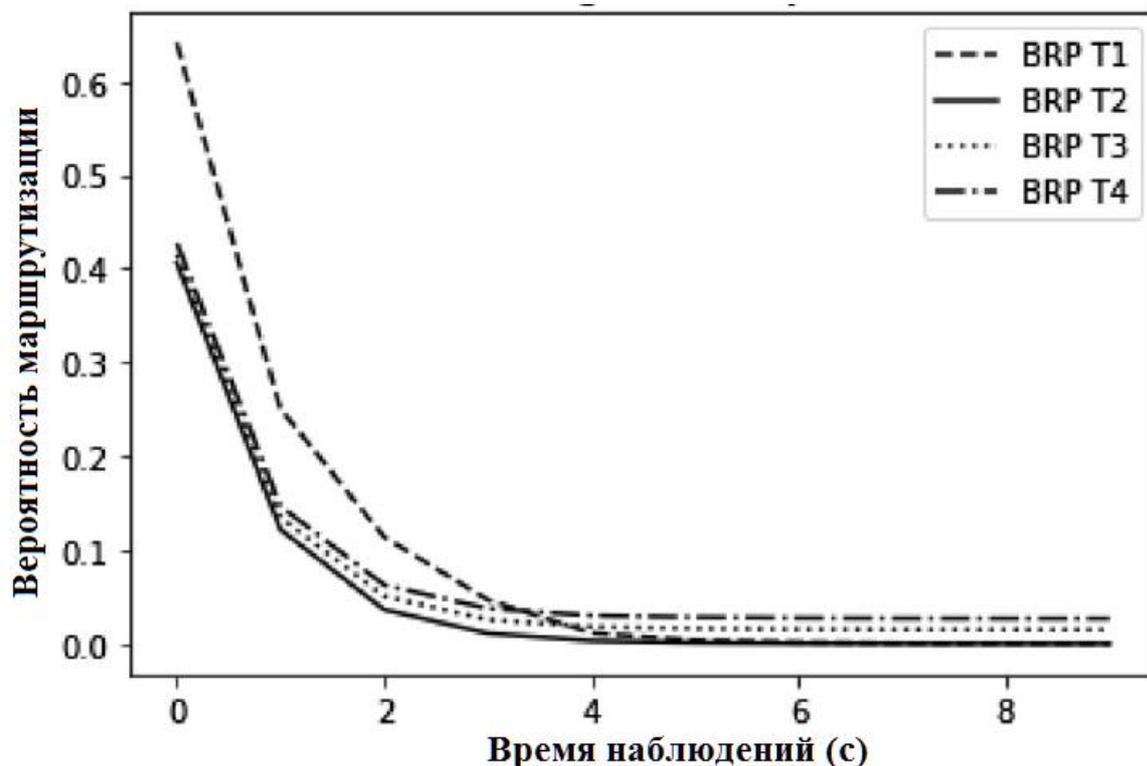


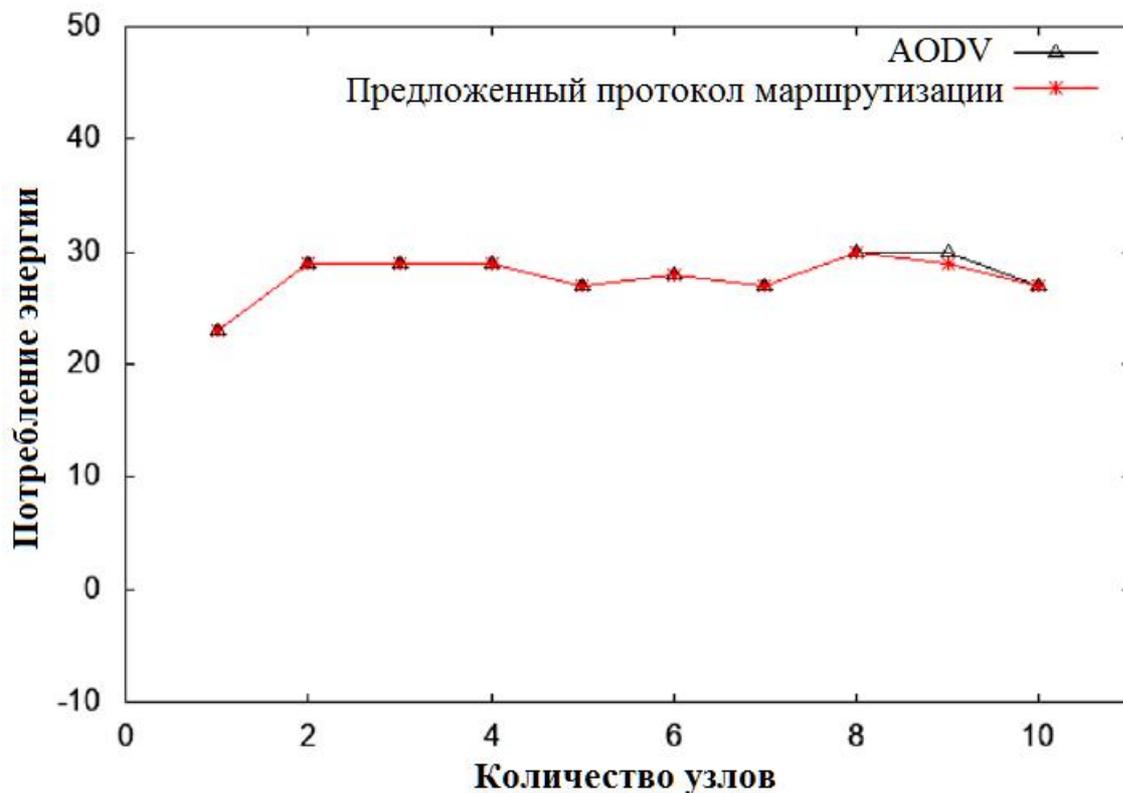
Рис. 3.4. Сравнение тенденций наилучшей вероятности управления потоками данных (BRP) между четырьмя этапами обучения

На рис. 3.4 показана тенденция изменения маршрута на четырех этапах обучения. Начиная с этапа 1, обозначенного пунктирной линией, высокая вероятность составляет около 0,6%. Вероятность на этапе 2, обозначенном пунктирной линией, значительно снижается. А с этапа 3 по этап 4 общее снижение энергопотребления составляет почти 0,4%. Динамика между наблюдениями при  $t=0$  и  $t=1$  показывает, что вероятность управления потоками данных в начале процесса высока. По истечении первой единицы времени это значение уменьшается до определенного системного порога. Кривая продолжает сходиться. В момент времени  $t=2$  траектория соответствует траектории 1, выбранной модулем принятия решений, и ее вес

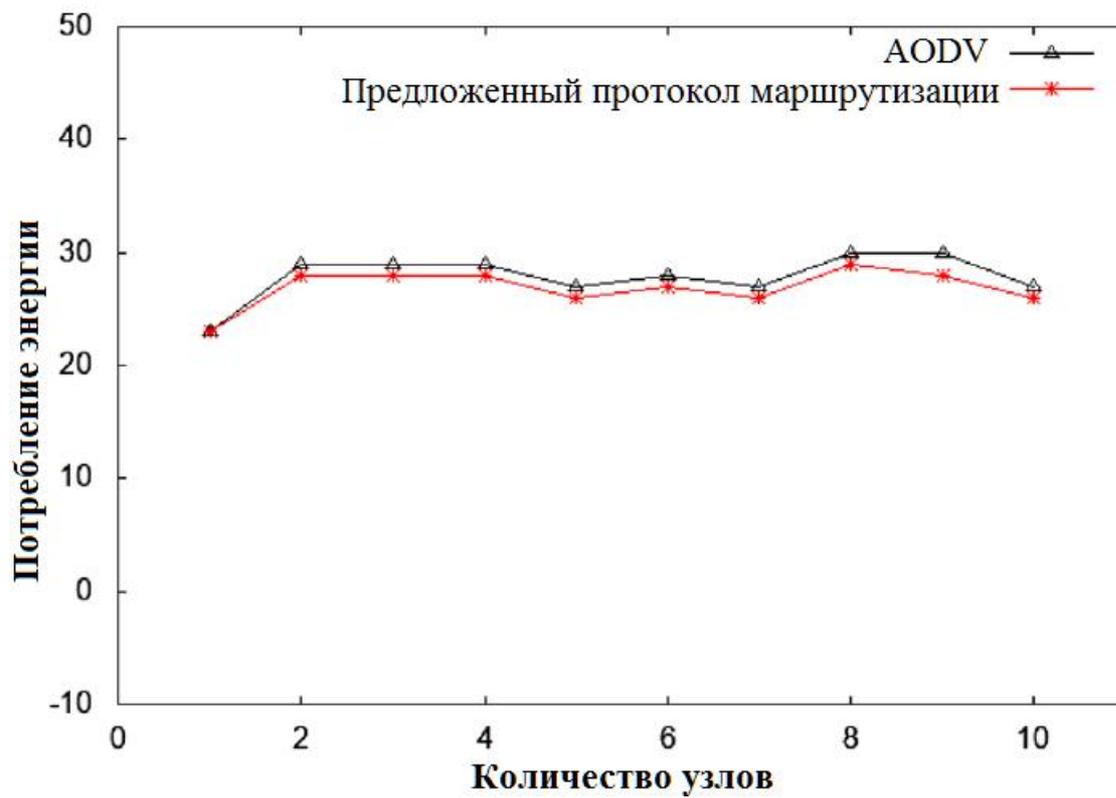
составляет  $cost=7$  хопов при низкой энергии около 196,895 Дж. На данном этапе это наименьшее значение по сравнению с другими траекториями.

### 3.4. Сравнение производительности

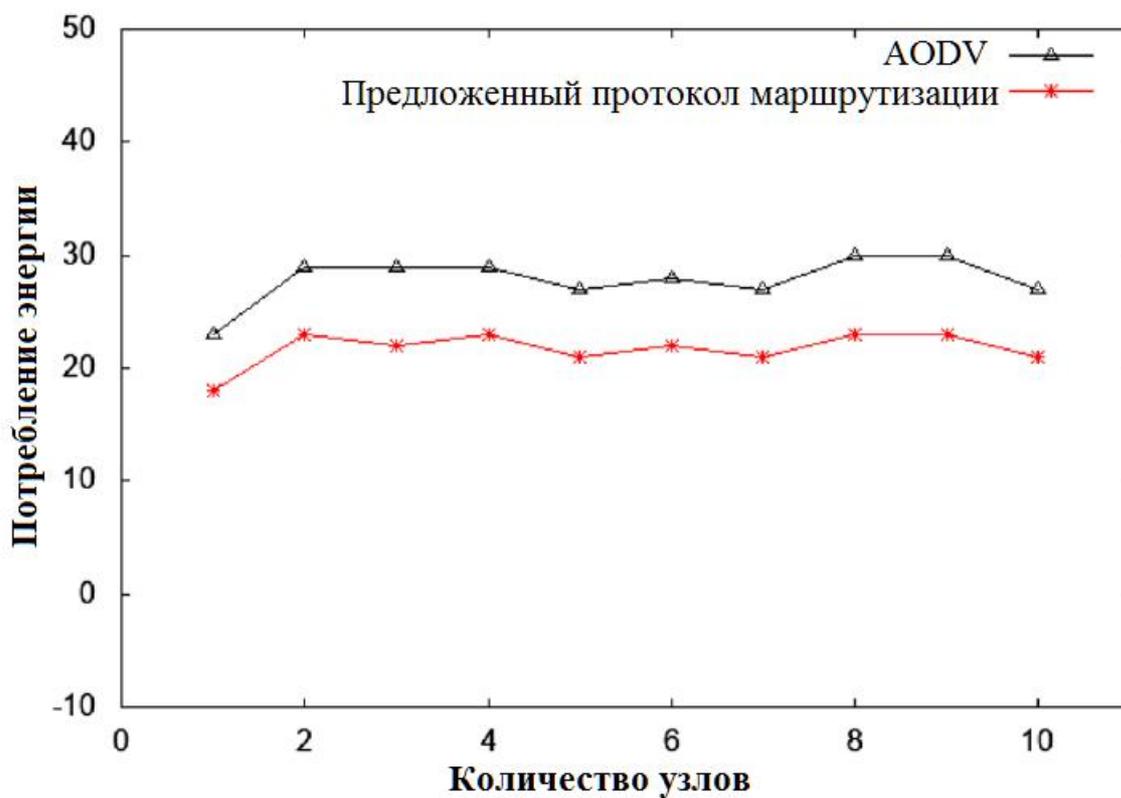
Проводится сравнение с традиционными протоколами управления потоками данных, такими как ad hoc on-demand distance vector routing protocol (AODV), для оценки эффективности предлагаемого протокола управления потоками данных на базе НММВ. Представлены значения энергозатрат после каждой серии тренировок и проведено среднее сравнение между этим подходом и обычным.



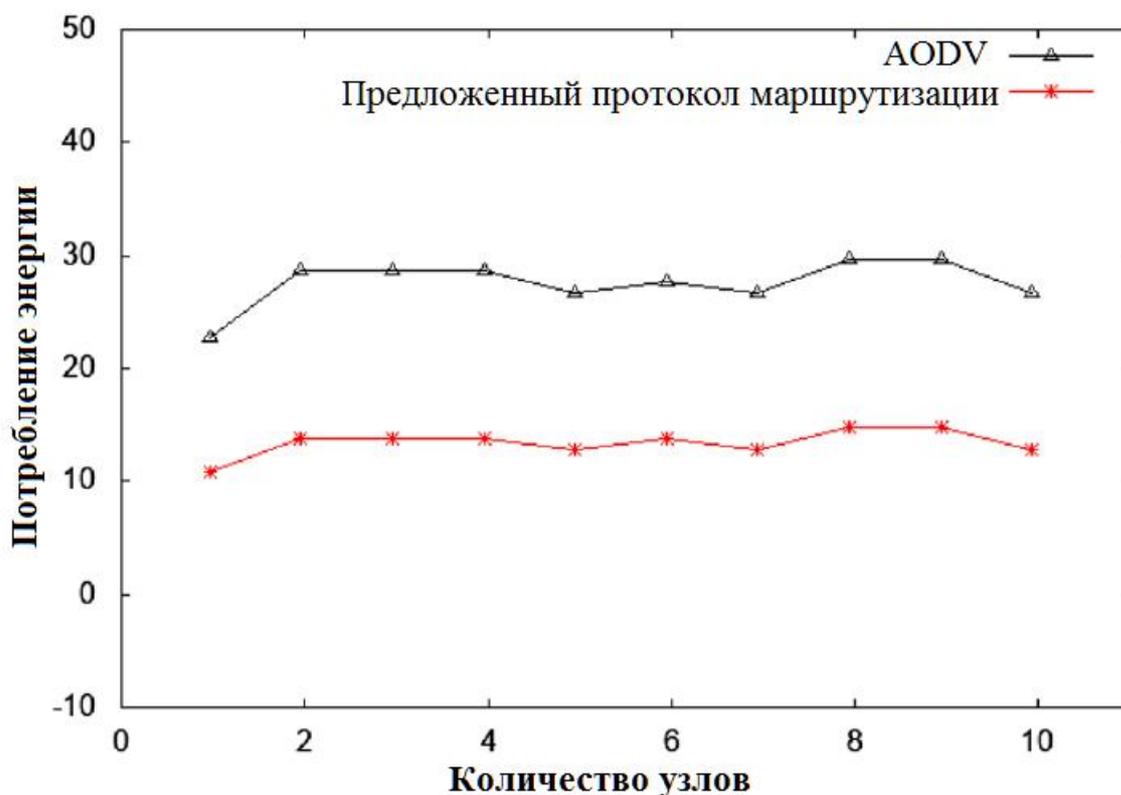
а) Потребление энергии на первом шаге



б) Потребление энергии на втором шаге



в) Потребление энергии на третьем шаге



г) Потребление энергии на четвертом шаге

Рис. 3.5. Сравнение поведения системы после нескольких циклов оптимизации: а) сравнивается начальное потребление энергии в системе со значениями после первого цикла и решения об управления потоками данных. Потребление близко к обычной вероятности управления потоками данных; б) система сравнивает потребление энергии после второго тренировочного цикла; в) сравнение с первоначальным потреблением энергии в системе и значениями после тридцатого тренировочного цикла. Снижение хорошо заметно; г) сравнивается потребление энергии после стабилизационного цикла

От тренировки к тренировке (табл. 3.5, рис. 3.5) потребление энергии в системе меняется. После этапа стабилизации достигается пороговое значение, и потребление энергии в системе значительно снижается.

Результаты ясно показывают, что предлагаемый алгоритм, основанный на стохастическом методе, снижает общее потребление энергии.

Кроме того, на рис. 3.6 показано среднее значение энергопотребления системы после всего обучения. Конечное энергопотребление в предлагаемой системе ниже, чем в исходном протоколе управления потоками данных.

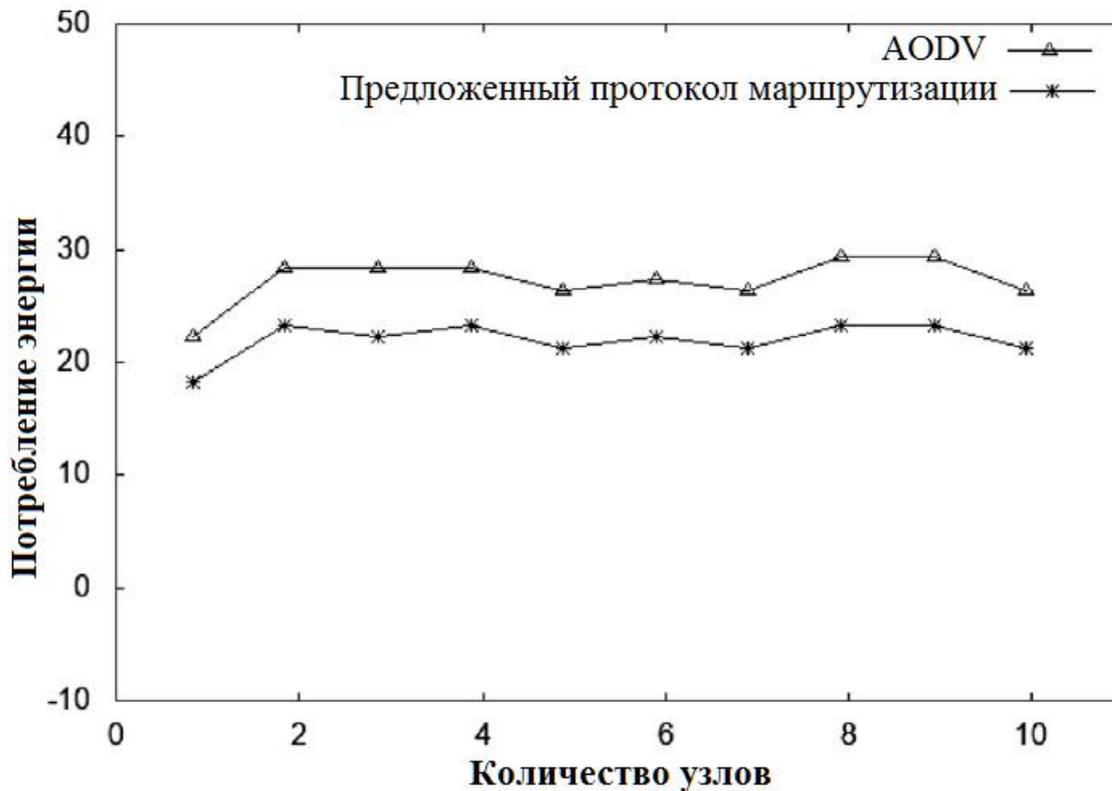


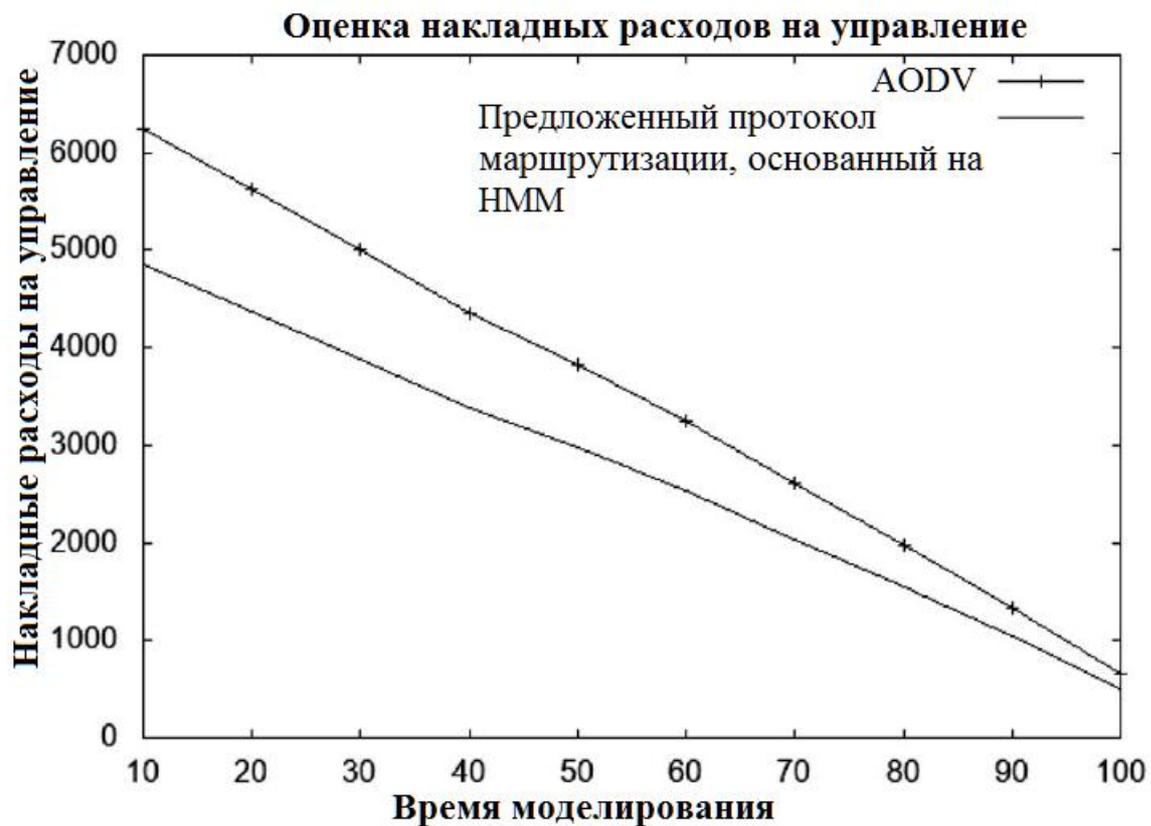
Рис. 3.6. Среднее потребление энергии системой

На рис. 3.6 мы ясно видим, что предлагаемая система значительно снижает энергопотребление. Среднее потребление энергии AODV составляет 285,46 Дж. В противном случае, после четырех тренировок в нашей системе среднее потребление составляет 231,07 Дж. Это значение соответствует снижению энергопотребления в системе с предложенной стохастической моделью на 19,05%.

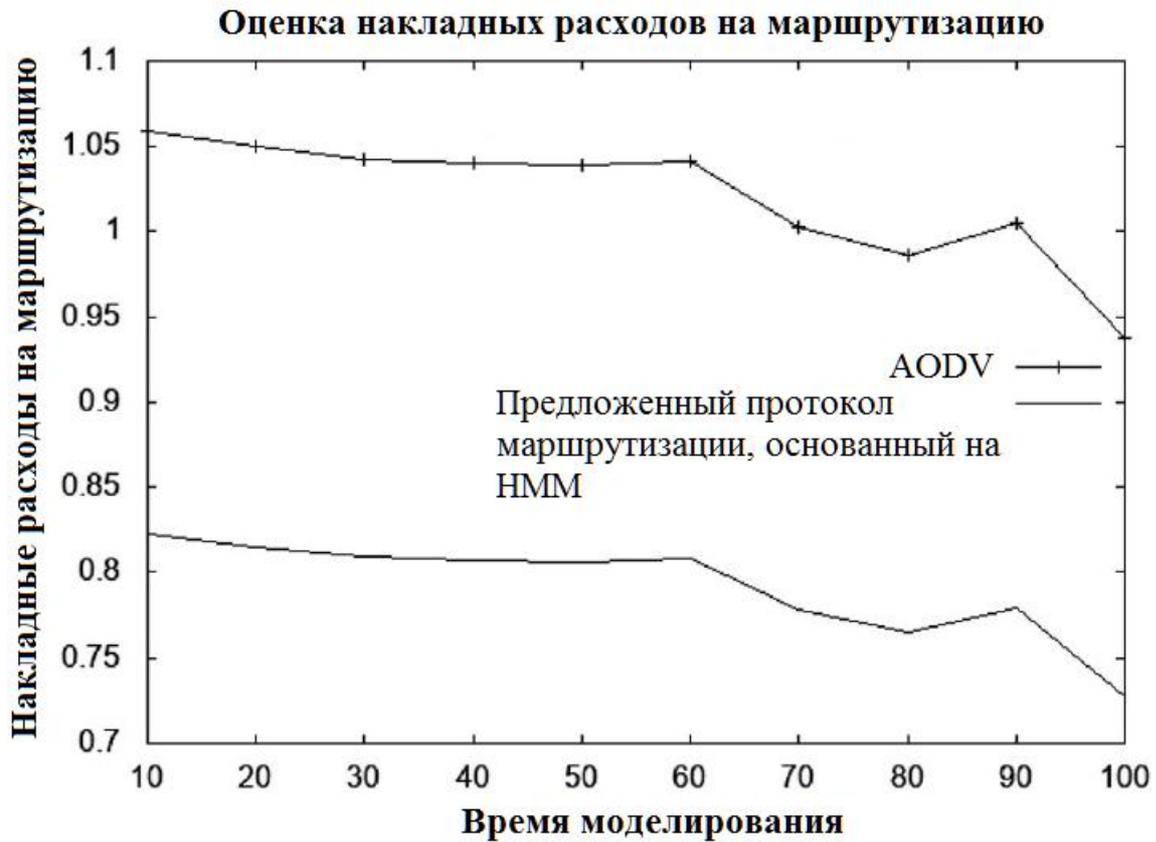
Кроме того, мы также оцениваем нагрузку на сеть. Это относится к контролю информации, отправляемой на тот адрес, который фактически получен каждым узлом сети. Он также определяет время обработки, необходимое для передачи данных узлом, и включает другие параметры, такие

как поддержка канала, обнаружение сети и узла, задержка переадресации и время ожидания. Издержки управления потоками данных позволяют получить отношение общего количества отправленных пакетов к общему количеству доставленных пакетов.

На рис. 3.7 сравниваются затраты на управление потоками данных между AODV и предлагаемым протоколом. Результаты ясно показывают, что протокол управления потоками данных на основе НММ обладает лучшими характеристиками.



а) накладные расходы на управление



б) накладные расходы на управление потоками данных

Рис. 3.7. Сравнение параметра управления потоками данных с накладными расходами в сети: а) накладные расходы на управление сокращаются примерно в то же время, что и для двух протоколов управления потоками данных. Значение параметра управления потоками данных, основанного на НММ, соответствует исходному протоколу управления потоками данных; б) сравнение с нормализованным протоколом управления потоками данных.

Это связано с тем, что предлагаемая модель определяет процесс управления потоками данных как стохастическую переменную и оценивает вероятность выбора маршрута с оптимальными параметрами.

Исходный протокол управления потоками данных приобретает все большее значение по сравнению с протоколом управления потоками данных на основе НММ.

### 3.5. Особенности результатов

Предлагаемый способ направлен на повышение энергопотребления в сети WSN. Например, если в период инициализации достигается пороговое значение, мы вычисляем и декодируем соответствующую скрытую последовательность, используя метод Витерби, который вычисляет вероятность кратчайшего пути с учетом модели  $I$  и принимает решение о маршруте. Напротив, если система не может достичь порогового значения по истечении времени инициализации, выполняется этап обучения с использованием метода Баума-Уэлча, который является частным случаем алгоритма максимизации математического ожидания (EM) для улучшения решения о маршруте в нашей исходной стохастической модели управления потоками данных  $I$ .

Максимальные вероятности перехода в скрытые низкоэнергетические состояния в каждый момент времени  $t$ , а также общее энергопотребление, оцениваемое как вероятность принятия наилучших решений об управлении потоками данных, зависят от производительности системы в этот момент времени. Конкретно, в начале процесса у нас есть высокая вероятность реализации управления потоками данных с более высоким значением энергопотребления в системе. Поскольку эта вероятность значительно превышает пороговое значение системы, нам необходимо установить аварийный сигнал для управления и мониторинга каждого объекта управления потоками данных, участвующего в процессе. По прошествии времени вероятность становится меньше начального порогового значения. Можно настроить генерацию аварийных сигналов, чтобы избежать неконтролируемой активности узлов.

С помощью нашей системы мы можем прогнозировать и предотвращать чрезмерное потребление энергии. Вполне вероятно, что если на пути у нас больше узлов, чем на других, то этот узел будет потреблять больше энергии в процессе управления потоками данных. Начальное время про-

цесса управления потоками данных показывает высокую вероятность высокого энергопотребления. Но в течение этого времени наша система регулирует использование энергии и оптимизирует срок службы сети. Мы понимаем, что при инициализации предложенной схемы стохастической управления потоками данных мы наблюдаем высокую загрузку ресурсов, которая со временем снижается.

В табл. 3.7 представлена максимальная вероятность получения наиболее оптимальной последовательности скрытых состояний с учетом случайной серии наблюдений ( $Y_t$ ) в момент времени  $t=10$  после четырех этапов обучения. Чтобы сгенерировать последовательность символов  $O(10)=Y_t$ , более вероятно начать сначала с состояния  $t_1=E_1$ , затем  $t_2=E_2$  и закончить в момент времени  $T=10$  в состоянии  $t_{10}=E_1$ . Максимальная вероятность может показаться очень малой, но это нормально. Существует множество возможных последовательностей скрытых состояний, которые могут быть сгенерированы в количестве порядка  $M^T$  [48], эквивалентном  $3^{10}$  в нашем конкретном случае.

Таблица 3.7

Максимальная оценка управления потоками данных

Скрытая последовательность	Вероятность
E1,E2,E3,E3,E3,E1,E1,E1,E2,E3,E3	$4.59e^{-04}$

### 3.6. Расчет параметров стохастического маршрута (пример)

Чтобы проиллюстрировать методологию, использованы числовые значения, приведенные в примере графа на рис. 3.2, рассчитанные с помощью уравнений (1.2), (3.2) и (3.3), которые обозначают начальные вероятности, вероятности перехода и вероятности выбросов соответственно.

$$\Pi = \begin{pmatrix} 0.332 \\ 0.333 \\ 0.335 \end{pmatrix};$$

$$A = \begin{pmatrix} 0.668 & 0.082 & 0.25 \\ 0.015 & 0.817 & 0.168 \\ 0.046 & 0.031 & 0.923 \end{pmatrix}$$

$$B = \begin{pmatrix} 0.918 & 0.036 & 0.047 \\ 0.007 & 0.982 & 0.011 \\ 0.009 & 0.002 & 0.989 \end{pmatrix}$$

Вектор  $\Pi$  определяет, в каком состоянии система должна находиться в начале, матрица  $A$  отражает связь между каждым состоянием, и, наконец,  $B$  описывает вероятности того, что система заметит символы  $L, M, H$ , находящиеся в данном состоянии. Мы запускаем последовательность шагов наблюдения  $Y$  случайным образом в течение времени  $T=10$ . Каждый объект  $Y_t$  представляет случайное решение о траектории, принятое в этот момент времени  $t$ , и выглядит следующим образом:

$$Y = \{Y_1=v_1, Y_2=v_2, Y_3=v_3, Y_4=v_3, Y_5=v_2, Y_6=v_1, Y_7=v_1, Y_8=v_2, Y_9=v_3, Y_{10}=v_1\}$$

Мы предполагаем, что наблюдаем короткую последовательность действий, приведенную ниже:

$$Y = \{v_1, v_2, v_3, v_3, v_2, v_1, v_1, v_2, v_3, v_1\} \quad (3.15)$$

Последовательность  $Y$  представляет собой выбранные случайным образом символы наблюдения, где каждый символ определяет уровень энергий в системе. Вероятность наблюдения последовательности, заданной моделью  $k$ , оценивается с помощью прямого алгоритма. Символы  $L, M$  и  $H$  соответствуют низкой, средней и высокой энергиям соответственно. Каждый символ  $V$  иллюстрирует энергетический уровень, при котором каждое состояние может быть:  $V=v_1$ =Низкий ( $L$ );  $v_2$ =Средний ( $M$ );  $v_3$ =Высокий( $H$ ).

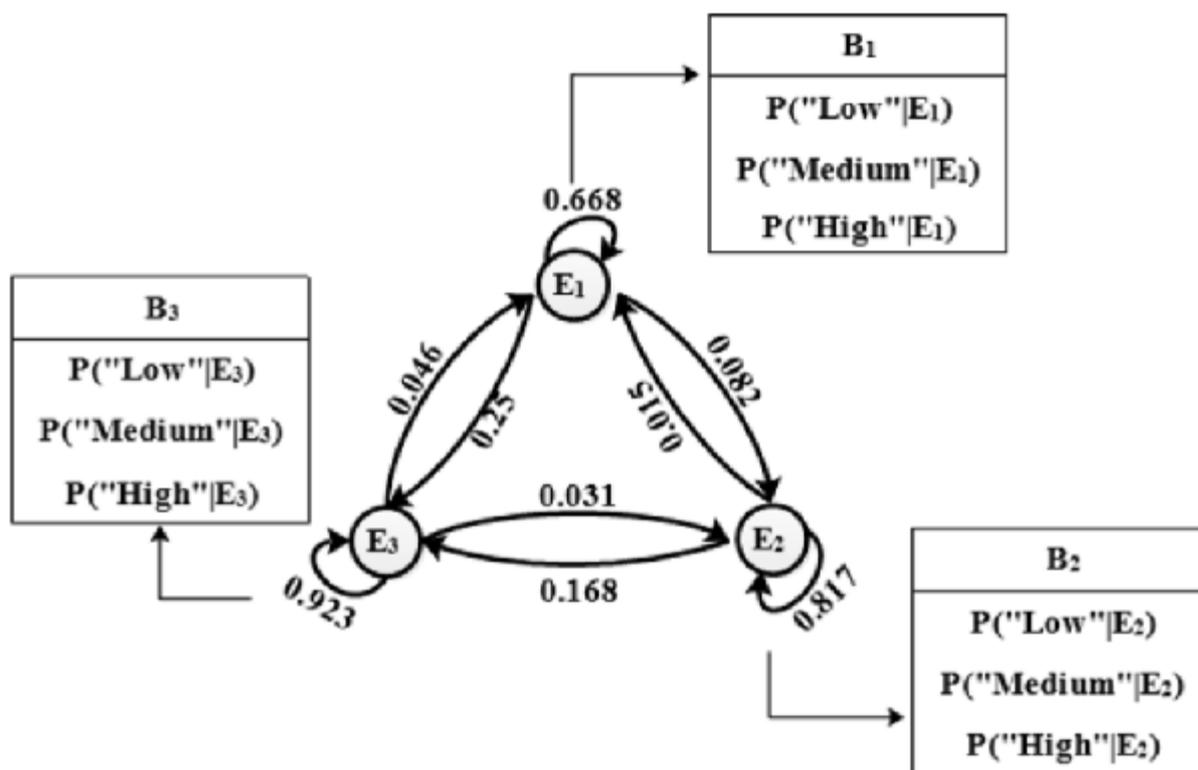


Рис. 3.8. Полносвязный граф НММ

На рис. 3.8 показаны коэффициенты переходов и начальный вектор для предложенной модели. Граф взаимосвязан тем, что все коэффициенты перехода между скрытыми состояниями не равны нулю. Коэффициенты перехода между состояниями приведены в табл. 3.8.

Таблица 3.8

Коэффициенты перехода между состояниями

К	Символ	Значение	$i=1, E_1$	$i=2, E_2$	$i=3, E_3$
1	Low	Низкая энергия	0,668	0,082	0,25
2	Medium	Средняя энергия	0,015	0,817	0,168
3	High	Высокая энергия	0,046	0,031	0,923

В табл. 3.8 представлен переход между скрытыми состояниями. Скрытые состояния обозначают энергию, потребляемую узлами, соответствующими кратчайшему пути.  $E_1$  - энергия в кратчайшем пути 1 или со-

стоянии 1, E2 - энергия, потребляемая узлами в кратчайшем пути 2, и E3 - энергия, потребляемая узлами в состоянии 3. Эта часть является завершением раздела, в котором описана полносвязная скрытая марковская модель. Затем эти коэффициенты вычисляются на основе уравнения, предложенного для оценки исходных параметров. K=3 кратчайших пути между исходным и конечным узлами с наилучшей минимальной энергией в соответствии с сетевым графом (рис. 3.2, 3.3).

### 3.9. Прямой расчет вероятностей

Основные этапы прямого вычисления вероятностей состоят из инициализации, ввода в действие и завершения следующим образом:

#### *Инициализация*

$$\alpha_1(E_1) = \Pi_{E_1} \cdot b_{E_1}(Y_1 = v_1(L)) = 0.332 \times 0.668 = 0.304776$$

$$\alpha_1(E_2) = \Pi_{E_2} \cdot b_{E_2}(Y_1 = v_1(L)) = 0.333 \times 0.082 = 0.027306$$

$$\alpha_1(E_3) = \Pi_{E_3} \cdot b_{E_3}(Y_1 = v_1(L)) = 0.335 \times 0.250 = 0.08375$$

#### *Шаг индукции*

$$\alpha_2(j) = b_j(Y_2) \cdot \sum_{i=1}^{N=3} \alpha_1(i) \cdot a_{ij}$$

$$\alpha_2(E_1) = b_{E_1}(Y_2 = v_2(M)) \cdot [\alpha_1(E_1) \cdot a_{11} + \alpha_1(E_2) \cdot a_{21} + \alpha_1(E_3) \cdot a_{31}]$$

$$\alpha_2(E_2) = b_{E_1}(Y_2 = v_2(M)) \cdot [\alpha_1(E_1) \cdot a_{12} + \alpha_1(E_2) \cdot a_{22} + \alpha_1(E_3) \cdot a_{32}]$$

$$\alpha_2(E_3) = b_{E_1}(Y_2 = v_2(M)) \cdot [\alpha_1(E_1) \cdot a_{13} + \alpha_1(E_2) \cdot a_{23} + \alpha_1(E_3) \cdot a_{33}]$$

...

$$\alpha_{10}(j) = b_j(Y_{10}) \cdot \sum_{i=10}^{10} \alpha_9(i) \cdot a_{ij}$$

#### *Завершение*

Процесс завершения вычисляет вероятность получения наилучшей последовательности наблюдений. Используется уравнение (3.16).

$$P(Y | I) = \prod_{i=1}^{10} a_T(i) \quad (3.16)$$

В каждый момент времени (t) мы наблюдаем символ. Это означает, что система может иметь уровень энергии, зависящий от вероятности. В момент времени T=10 производится суммирование, чтобы получить вероятность наблюдения нашей случайной последовательности с учетом исходной модели I.

В таблице 3.9 мы представляем результаты прямого анализа за определенный промежуток времени. В таблице результаты прямого алгоритма, обозначенные как  $\alpha$  (рис. 3.9), представляют собой вероятности обнаружения символа  $v_k$  в каждом скрытом состоянии на основе начальной последовательности случайных наблюдений. Важно показать это, поскольку эти числовые значения используются для определения наилучшей вероятности прохождения маршрута (BRP) в стохастической модели.

Таблица 3.9

Вероятность прямого вычисления ( $\alpha$ )

Время	Символ	E <sub>1</sub>	E <sub>2</sub>	E <sub>3</sub>
t <sub>1</sub>	L	0.305	0.002	0.003
t <sub>2</sub>	M	0.007	0.027	0.0002
t <sub>3</sub>	H	0.0002	0.0002	0.006
t <sub>4</sub>	H	2.156e <sup>-005</sup>	4.695e <sup>-006</sup>	0.006
t <sub>5</sub>	M	1.03139e <sup>-005</sup>	0.000185662	1.1568e <sup>-005</sup>
t <sub>6</sub>	L	9.401e <sup>-006</sup>	1.056e <sup>-006</sup>	4.080e <sup>-007</sup>
t <sub>7</sub>	L	5.793e <sup>-006</sup>	1.139e <sup>-008</sup>	2.668e <sup>-008</sup>
t <sub>8</sub>	M	1.387e <sup>-007</sup>	4.773e <sup>-007</sup>	3.126e <sup>-009</sup>
t <sub>9</sub>	H	4.661e <sup>-009</sup>	4.477e <sup>-009</sup>	1.163e <sup>-007</sup>
t <sub>10</sub>	L	7.858e <sup>-009</sup>	5.289e <sup>-011</sup>	1.004e <sup>-009</sup>

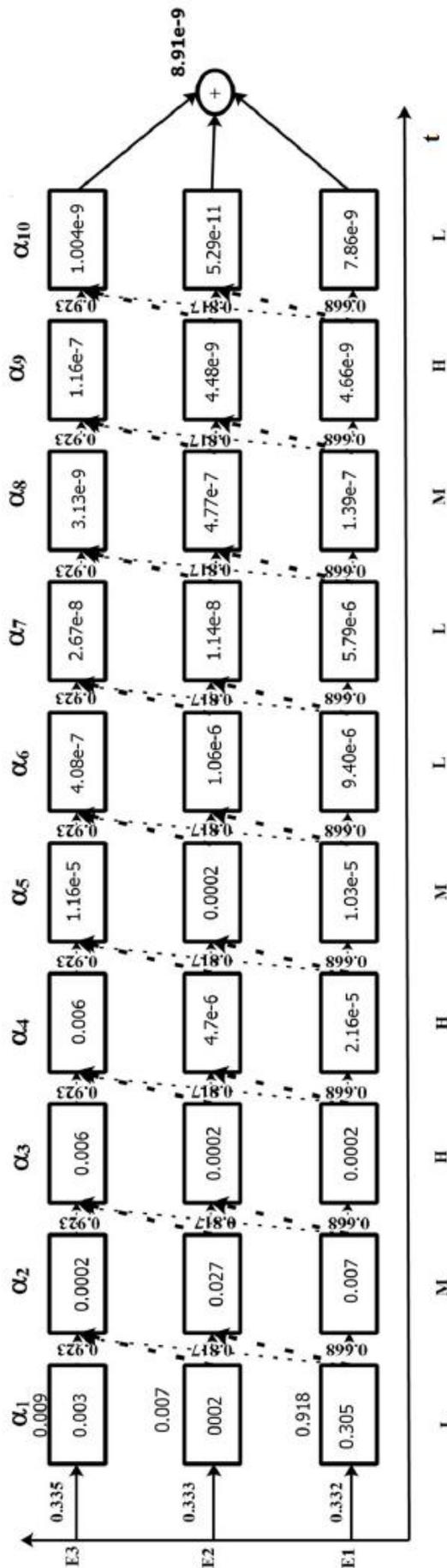


Рис. 3.9. Процесс вычисления  $\alpha$  для последовательности наблюдений. Эта оценка основана на последовательности символов, наблюдаемых в последовательности  $Y_t$

На рис. 3.9 мы вычисляем максимальную оценку, чтобы получить наилучший путь, соответствующий последовательности наблюдений  $Y$ . Это означает, что нужно найти наилучшую последовательность в модели  $I$ , которая максимизирует  $P(X, Y/I)$  (см. (3.13)). Основываясь на этом процессе и других промежуточных результатах, мы прогнозируем, что наилучшим скрытым состоянием, в результате которого исходная последовательность  $Y$  получила высокий балл, является то, которое показано в табл. 3.7.

### **3.7. Выводы к главе 3**

Экономия потребляемой энергии является важнейшим аспектом, который может повлиять на производительность беспроводной сенсорной сети. В этой работе мы предлагаем новый метод управления потоками данных, основанный на использовании скрытых марковских стохастических методов, который использует данные из существующего протокола управления потоками данных, чтобы изучить поведение сети и предложить новую схему управления потоками данных с оптимальными параметрами для выбора пути, который потребляет наименьшее количество энергии. Благодаря этому новому подходу увеличивается остаточный расход энергии, что позволяет продлить срок службы сети как минимум на 19,05% по сравнению с предыдущим подходом.

### **Литература к главе 3**

- 3.1. Sohraby K, Minoli D, Znati T (2007) *Wireless sensor networks: technology, protocols, and applications*. John Wiley and sons, New Jersey.
- 3.2. Kalkha H, Satori H, Satori K (2016) Performance evaluation of AODV and leach routing protocol. *Adv Inf Technol Theory Appl* 1(1):112–118.
- 3.3. Medina Carlos, Segura Jose´ C, de la Torre Angel (2013) Accurate time synchronization of ultrasonic TOF measurements in IEEE 802.15.4 based wireless sensor networks. *Ad Hoc Netw* 11(1):442–452.
- 3.4. Almesaeed R, Jedidi A (2021) Dynamic directional routing for mobile wireless sensor networks. *Ad Hoc Netw* 110:102301.
- 3.5. Kalkha H, Satori H, Satori K (2017) A dynamic clustering approach for maximizing scalability in wireless sensor network. *Trans Mach Learn Artif Intell*. <https://doi.org/10.14738/tmlai.54.3328>.
- 3.6. Akyildiz IF, Weilian S, Sankarasubramaniam Y, Cayirci E (2002) *Wireless sensor networks: a survey*. *Comput Netw* 38(4):393–422.
- 3.7. Zurita G, Shukla AK, Pino JA, Merigo´ JM, Lobos-Ossando´n V, Muhuri PK (2020) A bibliometric overview of the journal of network and computer applications between 1997 and 2019. *J Netw Comput Appl* 165:102695.
- 3.8. Russel A, Moundounga A, Satori H, Satori K (2020). An overview of routing techniques in wsns. In: 2020 IEEE fourth international conference on intelligent computing in data sciences (ICDS), pp 1–7.
- 3.9. Basheer A, Sha K (2017) Cluster-based quality-aware adaptive data compression for streaming data. *J Data Inf Qual (JDIQ)* 9(1):1–33.
- 3.10. Razzaque MA, Bleakley C, Dobson S (2013) Compression in wireless sensor networks: a survey and comparative evaluation. *ACM Trans Sens Netw (TOSN)* 10(1):1–44.
- 3.11. Wang C-F, Shih J-D, Pan B-H, Tin-Yu W (2014) A network lifetime enhancement method for sink relocation and its analysis in wireless sensor networks. *IEEE Sens J* 14(6):1932–1943.

3.12. Capo-Chichi EP, Guyennet H, Friedt JM (2009) K-rle: a new data compression algorithm for wireless sensor network. In: 2009 third IEEE international conference on sensor technologies and applications, pp 502–507.

3.13. Ortega AP, Ramchurn SD, Tran-Thanh L, Merrett GV (2021) Partner selection in self-organised wireless sensor networks for opportunistic energy negotiation: a multi-armed bandit based approach. *Ad Hoc Netw* 112:102354.

3.14. Rawat P, Singh KD, Chaouchi H, Bonnin JM (2014) Wireless sensor networks: a survey on recent developments and potential synergies. *J Supercomput* 68(1):1–48.

3.15. Swain RR, Dash T, Khilar PM (2019) A complete diagnosis of faulty sensor modules in a wireless sensor network. *Ad Hoc Netw* 93:101924.

3.16. Park GY, Kim H, Jeong HW, Youn HY (2013) A novel cluster head selection method based on k-means algorithm for energy efficient wireless sensor network. In: 2013 27th IEEE international conference on advanced information networking and applications workshops, pp 910–915.

3.17. Eshaftri M, Al-Dubai AY, Romdhani I, Yassien MB (2015) A new energy efficient cluster based protocol for wireless sensor networks. In: 2015 IEEE federated conference on computer science and information systems (FedCSIS), pp 1209–1214.

3.18. Hamzah A, Shurman M, Al-Jarrah O, Taqieddin E (2019) Energy-efficient fuzzy-logic-based clustering technique for hierarchical routing protocols in wireless sensor networks. *Sensors* 19(3):561.

3.19. Li J, Liu D (2015) Dpso-based clustering routing algorithm for energy harvesting wireless sensor networks. In: 2015 IEEE international conference on wireless communications and signal processing (WCSP), pp 1–5.

3.20. Singh DP, Bhateja V, Soni SK (2014) Prolonging the lifetime of wireless sensor networks using prediction based data reduction scheme. In: 2014 IEEE international conference on signal processing and integrated networks (SPIN), pp 420–425.

3.21. Boudries A, Aliouat M, Siarry P (2014) Detection and replacement of a failing node in the wireless sensors networks. *Comput Electr Eng* 40(2):421–432.

3.22. Obado V, Djouani K, Hamam Y (2012) Hidden markov model for shortest paths testing to detect a wormhole attack in a localized wireless sensor network. *Proc Comput Sci* 10:1010–1017.

3.23. Kalkha H, Satori H, Satori K (2019) Preventing black hole attack in wireless sensor network using hmm. *Proc Comput Sci* 148:552–561

3.24. Saihi M, Boussaid B, Zouinkhi A, Abdelkrim N (2015) Distributed fault detection based on hmm for wireless sensor networks. In: 2015 IEEE 4th international conference on systems and control (ICSC), pp 189–193.

3.25. Xiaofei X, Zhang Z, Chen Y, Li L (2020) Hmm-based predictive model for enhancing data quality in WSN. *Int J Comput Appl* 42(4):351–359.

3.26. Kumar S, Tiwari SN, Hegde RM (2015) Sensor node tracking using semi-supervised hidden markov models. *Ad Hoc Netw* 33:55–70.

3.27. Anand S, Rafeeqe KM (2022) Enhancing the security in wireless sensor network using hidden markov model. *Soft computing for security applications*. Springer, Berlin, pp 409–423.

3.28. Tabatabaei S (2020) A novel fault tolerance energy-aware clustering method via social spider optimization (sso) and fuzzy logic and mobile sink in wireless sensor networks (wsns). *Comput Syst Sci Eng* 35(6):477–494.

3.29. Fu X, Pace P, Aloï G, Li W, Fortino G (2021) Toward robust and energy-efficient clustering wireless sensor networks: A doublestage scale-free topology evolution model. *Comput Netw* 200:108521.

3.30. Song F, Zhu M, Zhou Y, You I, Zhang H (2019) Smart collaborative tracking for ubiquitous power IoT in edge-cloud interplay domain. *IEEE Int Things J* 7(7):6046–6055.

3.31. Bahbahani MS, Alsusa E (2017) A cooperative clustering protocol with duty cycling for energy harvesting enabled wireless sensor networks. *IEEE Trans Wirel Commun* 17(1):101–111.

3.32. Budianu C, Ben-David S, Tong L (2006) Estimation of the number of operating sensors in large-scale sensor networks with mobile access. *IEEE Trans Signal Process* 54(5):1703–1715.

3.33. Gupta P, Kumar PR (1999) Critical power for asymptotic connectivity in wireless networks. *Stochastic analysis, control, optimization and applications*. Springer, Berlin, pp 547–566.

3.34. Mozaffari M, Safarinejadian B, Shasadeghi M (2020) A novel mobile agent-based distributed evidential expectation maximization algorithm for uncertain sensor networks. *Trans Inst Meas Control* 43(7):1609–1619.

3.35. Wang Q, Hassanein H, Takahara G (2004) Stochastic modeling of distributed, dynamic, randomized clustering protocols for wireless sensor networks. In: *Workshops on mobile and wireless networking/high performance scientific, engineering computing/network design and architecture/optical networks control and management*. *Ad Hoc and Sensor Networks/Compil*, pp 456–463.

3.36. Mini RAF, Loureiro AAF, Nath B (2004) The distinctive design characteristic of a wireless sensor network: the energy map. *Comput Commun* 27(10):935–945.

3.37. Huang X, Acero A, Hon H-W, Reddy R (2001) *A guide to theory, algorithm, and system development, spoken language processing*. Prentice Hall PTR, New Jersey.

3.38. Hu P, Zhou Z, Liu Q, Li F (2007) The hmm-based modeling for the energy level prediction in wireless sensor networks. In: *2007 2nd IEEE conference on industrial electronics and applications*, pp 2253–2258.

3.39. Nazli Tekin and Vehbi Cagri Gungor (2020) Analysis of compressive sensing and energy harvesting for wireless multimedia sensor networks. *Ad Hoc Netw* 103:102164.

- 3.40. Yang S, Ma L, Jia S, Qin D (2019) A novel markov model-based low-power and secure multihop routing mechanism. *J Sens* 2019(2019):1–11.
- 3.41. Yen JY (1971) Finding the k shortest loopless paths in a network. *Manag Sci* 17(11):712–716.
- 3.42. Stahlbuhk T, Shrader B, Modiano E (2019) Learning algorithms for scheduling in wireless networks with unknown channel statistics. *Ad Hoc Netw* 85:131–144.
- 3.43. Malek A-G, Chunlin L, Zhiyong Y, Hasan AHN, Xiaoqing Z (2012) Improved the energy of ad hoc on-demand distance vector routing protocol. *IERI Proc* 2:355–361.
- 3.44. Liu S, Srivastava R, Koksal CE, Sinha P (2009) Pushback: a hidden markov model based scheme for energy efficient data transmission in sensor networks. *Ad Hoc Netw* 7(5):973–986.
- 3.45. Rohit Kumar and Joy Chandra Mukherjee (2021) On-demand vehicle-assisted charging in wireless rechargeable sensor networks. *Ad Hoc Netw* 112:102389.
- 3.46. Rabiner LR (1989) A tutorial on hidden markov models and selected applications in speech recognition. *Proc IEEE* 77(2):257–286.
- 3.47. Issariyakul T, Hossain E (2009) Introduction to network simulator 2 (ns2). *Introduction to network simulator NS2*. Springer, Berlin, pp 1–18.
- 3.48. Vouma Lekoundji J-B (2014) *Mode`les de Markov cache's*, PhD thesis, Universite´ du Que´bec a` Montre´al.

## **4. Математическое и программное обеспечение специального управления данными в прикладных задачах**

### **4.1. Обнаружение аномальных данных в электроэнергетической системе**

В разделе рассматриваются методы обнаружения аномальных данных в системе автоматизации измерения мощности, которые не требуют больших затрат времени и не дают низкой точности традиционных методов обнаружения аномальных данных в системе автоматизации измерения мощности. Разработана таблица структуры хранения данных в базе данных системы автоматизации учета электроэнергии, исправлены недостающие данные и устранены помехи в таблице данных. Выполнены предварительные вычисления на основе данных, чтобы получить последовательность элементов данных. После того, как алгоритм Р-кластеризации осуществит кластеризацию данных, модель леса используется для обнаружения аномальных данных для завершения исследования метода. Результаты экспериментов показывают, что предложенный метод обнаружения обладает преимуществами короткого времени обнаружения и высокой точностью 91,26–95,67%.

#### ***4.1.1. Метод обнаружения аномальных данных в системе автоматизации измерения электроэнергии***

Данные, используемые для обнаружения аномальных данных в системе автоматизации учета электроэнергии, содержатся в таких таблицах, как таблица потока сообщений системного терминала и таблица записей эксплуатации и технического обслуживания в полевых условиях. Следовательно, таблица структуры системных данных должна быть создана в базе данных терминала.

Таблица потока сообщений используется для записи состояния связи между каждым терминалом и главной станцией, в основном, включая поток сообщений, количество повторных подключений и время работы в режиме онлайн и т.д. Конкретное определение приведено в табл. 4.1.

Таблица 4.1

Таблица потока сообщений

Объект	Имя поля БД	Данные
Код терминала	rtuid	
Дата передачи данных	datetime	
Байт отправки (нисходящий поток)	sendbytes	
Байт получения (восходящий поток)	recvbytes	
Время повторного подключения	logintimes	
Поток данных	databytes	
Поток сигналов тревоги	alarmbytes	
Время работы потока тактов	linkbytes	
Время	onlinetimes	

Данные в этой таблице собираются и подсчитываются главной станцией с нулевым значением каждый день, что отражает передачу данных за предыдущий день. Если терминал выйдет из строя, главная станция не сможет собирать текущие данные, поэтому в базе данных нет соответствующей записи. Отправляющий и принимающий байты в таблице относятся к главной станции. Отправляющие байты представляют собой количество байтов, отправленных с главной станции на терминал, а полученные байты представляют собой количество байтов, полученных главной станцией [4.3]. Поток данных в таблице указывает на поток, используемый для передачи данных об электроэнергии среди всех потоков. В дополнение к вышеуказанному потоку, в таблицу также включена другая управляющая информация терминала, включая количество повторных подключений, ко-

личество аварийных сигналов и количество мерцаний. Время работы в режиме онлайн показывает количество сигналов, полученных главной станцией, и которые терминал передает один раз в минуту. Если значение равно 1440, это означает, что терминал подключен к сети в течение 24 часов (табл. 4.2).

Таблица 4.2

Таблица потерь на линии

№ п/п	Объект	Имя поля БД
1	Номер строки	LINEID
2	Название строки	DISC
3	Дата передачи данных	DATETIME
4	Входная мощность	ENERGY_IN
5	Выходная мощность	ENERGY_OUT
6	Частота потери данных на линии	LINELOSS_RATE
7	Дата отправки данных	SENDDATE

В таблице эксплуатации и технического обслуживания на месте регистрируются результаты технического обслуживания и инспекций, проведенных персоналом по эксплуатации и техническому обслуживанию в случае возникновения неисправности, включая информацию о терминале и информацию о неисправностях. Таблица генерируется системой для получения информации, относящейся к терминалу, включая всевозможные номера, время передачи данных и местоположение с главной станции, а также информацию о неисправностях, которая заполняется персоналом по эксплуатации и техническому обслуживанию. Тип неисправности и описание неисправности заполняются вручную персоналом по эксплуатации и техническому обслуживанию, а затем вводятся в систему. Таким образом, для одной и той же неисправности имена, указанные разными специалистами по эксплуатации и техническому обслуживанию, могут отличаться,

что требует дополнительной обработки.

Обрыв линии электропередачи относится к потере электрической энергии при передаче и преобразовании электроэнергии. Таблица потерь на линии используется для записи входной и выходной мощности и частоты потерь на линии каждой линии и содержит основную информацию, относящуюся к ежедневным потерям на линии для каждой линии, как показано в табл. 4.2.

После того, как вышеуказанная таблица данных будет введена в базу данных системы автоматизации учета электроэнергии, данные, собранные и сохраненные системой в базе данных, будут обработаны.

#### ***4.1.2. Обработка данных системы автоматизации учета электроэнергии***

По разным причинам данные будут неполными и противоречивыми. Эти данные называются данными об ошибках и оказывают большое влияние на последующее обнаружение аномалий. Поэтому фильтрация данных очень важна для обнаружения аномалий.

При фильтрации данных сначала необходимо удалить избыточные данные в наборе данных. Избыточные данные - это единственная характеристика, которая уничтожает все записи в наборе данных. При появлении нескольких идентичных записей избыточные данные необходимо удалить. У каждого пользователя должны быть данные о потреблении электроэнергии за каждый час дня, и серьезное отсутствие определяется следующим образом:

- 1) на кривой не хватает 20% точек измерения;
- 2) на кривой постоянно отсутствуют более 2 последовательных точек считывания. Если данные отсутствуют в значительной степени, пользователь исключается из области исследования, и для восстановления пропущенного значения используется метод многоуровневой лагранжевой ин-

терполяции. Формула восстановления пропущенного значения выглядит следующим образом [4.4]:

$$P_t = \frac{\overset{m_1}{\underset{k=1}{\overset{\circ}{a}}} P_{t-k} + \overset{m_2}{\underset{i=1}{\overset{\circ}{a}}} P_{t+i}}{m_1 + m_2} \quad (4.1)$$

где  $m_1$  - количество периодов прямой передачи,

$m_2$  - количество периодов обратной передачи,

$t$  - время, когда системные данные отсутствуют,

$P_t$  - значение, отсутствующее после ремонта,

$P_{t-k}$  - системные данные на момент времени  $k$  до  $t$ ,

$k$  - время  $i$  по истечении времени  $t$ .

После внесения исправлений в данные устраняются шумы путем сглаживания кривой зависимости времени работы системы.

Установим общее время работы  $T$  системы в течение периода времени, в течение  $t'$  которого система находилась в ненормальном состоянии, регистрируемом оборудованием наблюдения, начальная точка ненормального состояния записывается как  $P_{start}$ , конечная точка ненормального состояния записывается как  $P_{end}$ . Предположим, что набор данных  $P_{i\phi}$ , собранных системой автоматизации учета электроэнергии в  $t'$ , выражается следующим образом:

$$P_{i\phi} = \{P_{start}, P_2, \dots, P_{n-1}, P_{end}\}$$

и точка данных  $P_{i\phi}$  в  $t'$  после обработки записывается как [4.5]:

$$P_{i\phi} = \frac{P_{i\phi} \pm (P_{start} - P_{end})}{time_{P_{end}} - time_{P_{start}}} \quad (4.2)$$

После обработки точки данных на кривой соотношения временных рядов можно разделить только на два типа: нормальные данные и аномальные данные. После обработки данных автоматизированной системы измерения электроэнергии извлекаются признаки аномальных данных.

### 4.1.3. Обнаружение признаков аномальных данных

В процессе обнаружения аномальных данных в системе автоматизации измерения электроэнергии сначала получают временной ряд для каждого из данных в сетевой базе данных, в качестве элемента которого используется среднее значение временного ряда. Конкретные шаги заключаются в следующем.

Предположим, что  $Q = \{q_1, q_2, \dots, q_m\}$  и  $C = \{c_1, c_2, \dots, c_n\}$  представляют два временных ряда данных, а  $w_q$  и  $w_c$  представляют временные ряды двух рядов данных. Используем следующую формулу для вычисления среднего значения всех элементов данных в каждом временном ряду [4.6]:

$$q_i = \frac{1}{w} (m) \frac{Q \{q_1, q_2, \dots, q_m\} \xi^{w_q, w_c} \xi}{C \{c_1, c_2, \dots, c_n\} \xi} \quad (4.3)$$

В формуле (4.3)  $w$  представляет собственное значение для формирования новой последовательности данных,  $m$  представляет кусочно-агрегированную аппроксимацию, а  $w_q$  и  $w_c$  представляют среднее значение элементов данных. Предполагая, что расстояние  $d(i, j)$  представляет собой динамический временной сдвиг,  $\xi$  представляет собственное значение для формирования нового ряда данных, а  $N$  представляет форму изменения данных во временном ряду данных, затем используем уравнение (4.4), чтобы получить характеристический ряд данных со средним значением временного ряда как элемент [4.7].

$$r(i, j) = \frac{L_{DTW} \cdot D_{w_q, w_c} \mathbf{g}^d(i, j)}{[\mathbf{x} \mathbf{g}^N] \mathbf{g}_i} [\mathbf{q}] \quad (4.4)$$

В формуле (4.4)  $L_{DTW}$  представляет процесс, в ходе которого два временных ряда сначала преобразуются в последовательности объектов,  $D_{w_q, w_c}$  представляет исходные данные временных рядов, а  $[\mathbf{q}]$  представляет матрицу накопления расстояний. После извлечения аномальных данных из системы автоматизации учета электроэнергии обнаруживаются аномаль-

ные данные.



Рис. 4.1. Выявление аномальных данных

#### 4.1.4. Реализация обнаружения аномальных данных

При кластерном анализе алгоритма AP используется матрица сходства аномальных признаков данных, и сходство между точками данных выражается квадратом отрицательного евклидова расстояния. Если имеется  $n$  данных, то эти точки данных образуют матрицу подобия  $S$   $n \times n$ ,  $S(i,j)$  представляет собой сходство между точками данных  $i$  и  $j$ , формула расчета выглядит следующим образом:

$$S(i, j) = - \|x_i - x_j\|^2, S(i, j) \in (-\infty, 0] \quad (4.5)$$

Значения элементов  $S(i,j)$  на диагонали матрицы подобия используются для определения центра кластера.  $S(i,j)$  называется параметром предпочтения, который указывает на пригодность точки данных  $i$  в качестве центра кластера класса. Если его значение больше, это означает, что точка больше подходит в качестве центра кластера. Установим среднее значение матрицы подобия в качестве предпочтительного:

$$\text{preference} = \frac{\sum_{i,j=1}^n S(i,j)}{n \cdot (n - 1)} \quad (4.6)$$

Алгоритм AP постоянно обновляет матрицу притяжения  $R$  и матрицу атрибутов  $A$  в процессе итерации. Информация о притяжении  $r(i,k)$  - это информация, отправляемая точкой  $i$  выборки в возможный центр кластеризации  $k$ , которая указывает степень притяжения точки  $i$  выборки к  $k$ . Если значение больше, это указывает на то, что  $k$  с большей вероятностью станет центром  $i$ . Информация, отправленная центром кластеризации  $k$  в выборочные данные  $i$ , отражает степень отнесения  $k$  к центру выборочной точки  $i$ . Если значение больше,  $i$ , скорее всего, принадлежит кластеру с центром  $k$ . В итеративном процессе две вышеуказанные информационные матрицы зависят друг от друга и обновляются поочередно. Процесс обновления заключается в следующем [4.8]:

При  $i \neq k$ ,

$$S(i,k) - \max \{A(i,k) + S(i,k)\} \otimes R(i,k) \quad (4.7)$$

При  $i=k$ ,

$$S(k,k) - \max \{A(k,k) + S(k,k)\} \otimes R(k,k) \quad (4.8)$$

Когда количество итераций матрицы притяжений  $R$  и матрицы атрибутов  $A$  превысит максимальное количество итераций, заданное заранее, или изменение  $r(i,k) + a(i,k)$  будет ниже заданного порогового значения, итерационный процесс остановится. После того, как алгоритм кластеризации AP сгруппирует наборы данных, iForest обнаружит выбросы кластеризации

зованных данных.

Предположим, что весь набор данных после классификации алгоритмом AP есть  $\Psi$ ,  $\Psi=(\psi_1, \psi_2, \dots, \psi_n)$ , где  $\psi_q$  представляет q-й кластер.

iForest состоит из  $f$  iTree изолированных деревьев (рис. 4.2), каждое из которых представляет собой бинарную древовидную структуру.

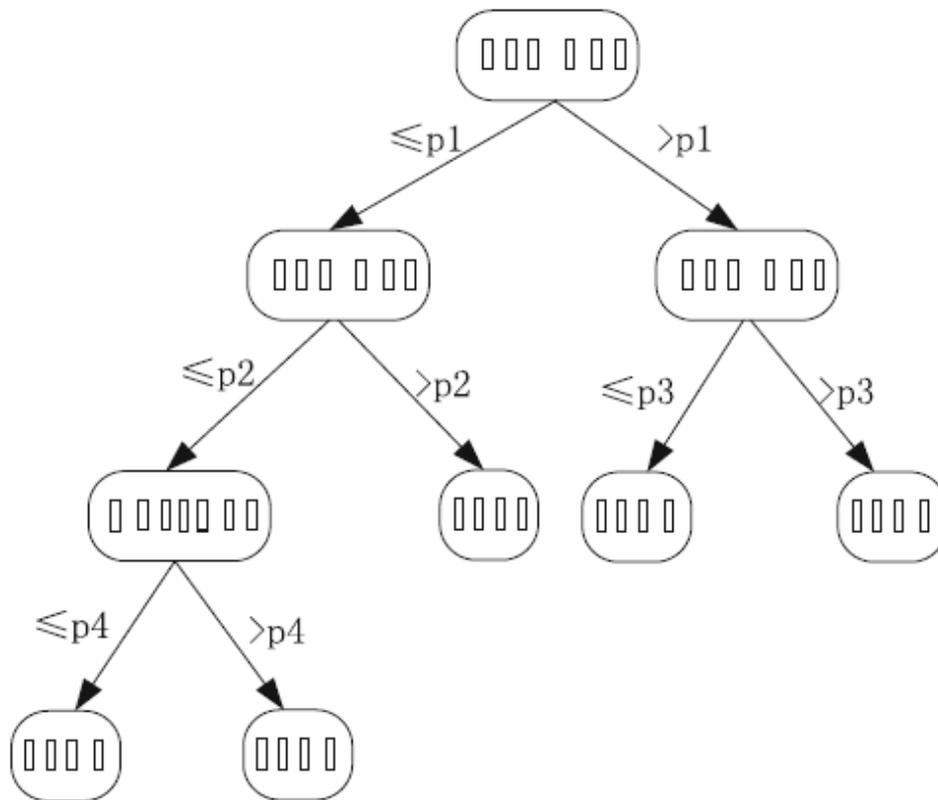


Рис. 4.2. Принципиальная схема конструкции iTree

Процедура обучения iTree выглядит следующим образом:

1) Поместить  $m$  значений набора данных  $D=\{x_1, x_2, \dots, x_m\}$  в корневой узел дерева.

2) Случайным образом указать атрибут  $r$  в текущем  $r$  и случайным образом сгенерировать данные с разделяющим значением  $p$ . Значение  $p$  - это величина между максимальным и минимальным значениями указанного атрибута в текущих данных.

3) Значение  $p$  делит текущее пространство данных на два подпространства, размещает данные, которые меньше  $p$  в указанном атрибуте  $r$ , в

левое подмножество и размещает данные, которые больше или равны  $p$ , в правое подмножество.

4) Шаги рекурсии 2 и 3 генерируют новые узлы до тех пор, пока в узле не останется только один объект данных.

Этапы обучения для  $f$   $iTree$  следующие.

Случайным образом выбирается  $\zeta$  образцов из кластера  $\psi_q$  и используются для создания  $iTree$  в соответствии с описанным выше процессом. Описанные выше действия выполняются  $f$  раз, чтобы получить  $f$  изолированных деревьев, которые образуют набор  $iTree$ . Когда коллекция  $iTree$  собрана, следующим шагом будет этап оценки исключений из данных.

Сначала вычисляется степень  $h(x)$  пути для данных  $x$ , которые необходимо обнаружить. Степень  $h(x)$  пути относится к числу итераций от корневого узла до конца конечного узла. Для дерева  $iTree$  данные  $x$  перемещаются вниз по условию разделения, соответствующему созданию, до тех пор, пока не будет получен доступ к конечному узлу, и не будет записана длина пути  $h(x)$ . Поскольку структура  $iTree$  такая же, как и у дерева бинарного поиска, длина пути к конечному узлу, содержащему данные  $x$ , совпадает с длиной пути неудачного запроса в дереве бинарного поиска, то есть от корневого узла к среднему узлу, достигающему конечного узла, количество пройденных ребер. Производится переход по  $iTree$  и вычисляется оценка исключений для данных  $x$ . В соответствии с оценкой аномалий оцениваются аномальные данные.

#### **4.2. Экспериментальное исследование метода обнаружения аномальных данных в электроэнергетической системе**

Для проверки эффективности разработанного метода предлагается следующая схема проведения сравнительных экспериментов.

#### **4.2.1. Содержание эксперимента**

Эксперимент является имитационным. Экспериментальная группа представляет собой метод обнаружения аномальных данных в системе автоматизации учета электроэнергии, рассмотренный в работе, а экспериментальная группа - традиционный метод обнаружения аномальных данных в системе автоматизации учета электроэнергии. В ходе эксперимента было проведено две серии экспериментов. В первой серии экспериментов сравнивались модель iForest экспериментальной группы и модель RBF сравнительной группы при обнаружении наборов данных, содержащих одни и те же аномальные данные; второй группой экспериментальных показателей сравнения были аномалии экспериментальной группы и метод обнаружения данных сравнительной группы, когда аномальные данные обнаруживаются в наборе экспериментальных данных, скорость воспроизведения и точность метода.

#### **4.2.2. Подготовка к эксперименту**

Данные измерения мощности, выбранные в этом эксперименте, включают информацию о потреблении электроэнергии некоторыми пользователями. Ниже приведены подробные сведения об источнике данных (табл. 4.3).

Таблица 4.3

Подробная информация о наборе экспериментальных данных

<b>№ п/п</b>	<b>Параметр</b>	<b>Тип</b>	<b>Комментарий</b>
1	CJ_MP_ID	Integer	Идентификация точки измерения
2	DATA_DATE	Datetime	Временные рамки данных
3	DATA_SORCE	Integer	Источники данных
4	WRITE_DATE	Datetime	Данные для записи
5	PZ	Float	Общая активная мощность

№ п/п	Параметр	Тип	Комментарий
6	PA	Float	Активная мощность фазы А
7	PB	Float	Активная мощность фазы В
8	PC	Float	Активная мощность фазы С
9	QZ	Float	Общая реактивная мощность
10	QA	Float	Реактивная мощность фазы А
11	QB	Float	Реактивная мощность фазы В
12	QC	Float	Реактивная мощность фазы С
13	MID_I	Float	Ток нулевой последовательности

Формула расчета коэффициента точности метода обнаружения аномальных данных выглядит следующим образом:

$$\text{Точность} = \frac{TP}{TP + FP} \quad (4.9)$$

Формула расчета частоты отзыва данных для метода обнаружения аномальных данных выглядит следующим образом:

$$\text{Отзыв} = \frac{TP}{TP + FN} \quad (4.10)$$

В формулах (4.9) и (4.10) TP - нормальные данные, которые оцениваются как положительные с помощью метода обнаружения. FP - данные, которые метод обнаружения оценивает как положительные, но на самом деле являются аномальными, а FN - это данные, которые метод обнаружения оценивает как аномальные, но на самом деле являются нормальными данными.

Различные наборы данных выполняют обнаружение аномалий один за другим, записывают точность и повторное использование, принимают точность за ординату, а повторное использование за абсциссу и строят кривую P-R. Если кривая P-R одного метода обнаружения включена в другой, то последний лучше первого. Если кривые P-R двух классификаторов

пересекаются, мы можем выбрать точку равновесия, которая является значением при  $p=R$ . Чем больше точка равновесия ВЕР, тем выше эффективность метода обнаружения.

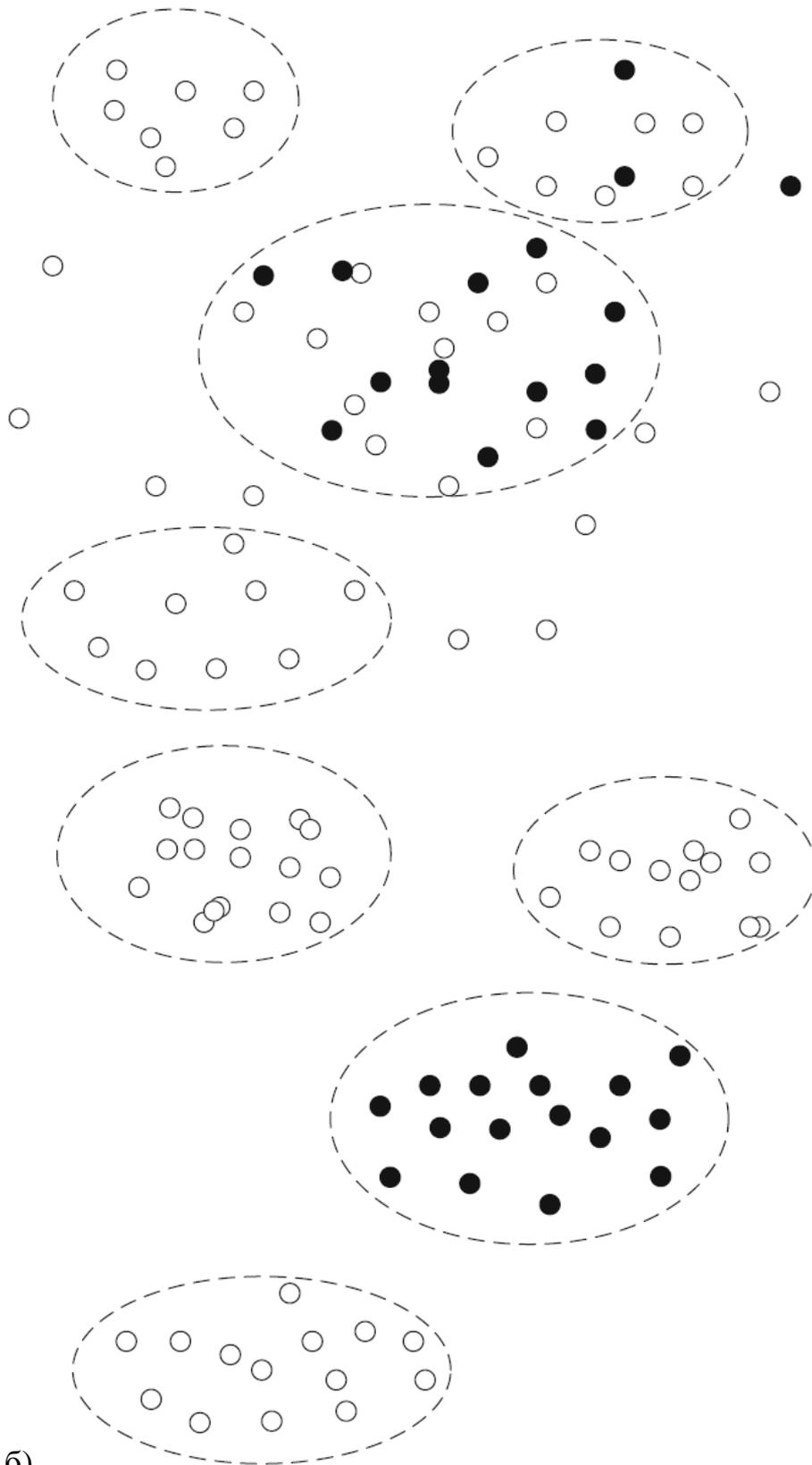
Запишем экспериментальные данные двух экспериментов, обработаем и проанализируем экспериментальные данные и сделаем соответствующий вывод.

### ***4.2.3. Результаты эксперимента***

Чтобы проверить эффективность этого метода, были обнаружены аномальные данные модели iForest и модели RBF контрольной группы, и результаты обнаружения показаны на рис. 4.3.

а) На рис. 4.3 показано, что iForest в основном может обнаруживать аномальные данные, когда для обнаружения используется набор данных, содержащий выбросы, в то время как аномальные данные, обнаруживаемые RBF, намного меньше, чем обнаруживаемые iForest, что указывает на то, что эффект обнаружения аномальных данных модели iForest в системе автоматизации учета электроэнергии лучше.

Сравнивается и анализируется время обнаружения аномальных данных для модели iForest и модели контрольной группы RBF. Результаты сравнения представлены в табл. 4.4.



б)

Рис. 4.3. Сравнение эффективности обнаружения: ○ - верно, ● - выбросы, - - - - функция обнаружения; а) эффективность обнаружения RBF, б) эффективность обнаружения iForest

Таблица 4.4

Сравнение времени обнаружения двумя методами в миллисекундах

<b>Объем данных</b>	<b>iForest</b>	<b>RBF</b>
5000	521	763
10000	772	995
15000	1064	1571
20000	1474	2346
25000	1838	2892
30000	2345	3487
35000	2760	4109
40000	3596	6007

Из анализа рис. 4.3 видно, что когда для обнаружения используется набор данных с выбросами, iForest может в основном обнаруживать выбросы, в то время как выбросы, обнаруженные RBF, намного меньше, чем те, которые были обнаружены iForest, и в результатах обнаружения RBF есть несколько нормальных данных. Согласно табл. 4.4, время обнаружения iForest невелико, а точность обнаружения высока. Это показывает, что в данном эксперименте по обнаружению влияние iForest на аномальные данные лучше.

Однако iForest не может обнаружить локальные выбросы, поэтому на основе первого эксперимента во втором эксперименте тестируется метод обнаружения аномальных данных, сочетающий алгоритмы iForest и кластеризации, рассмотренные с точки зрения данных.

Экспериментальные результаты второго эксперимента приведены в табл. 4.5, а экспериментальные выводы получены путем анализа данных табл. 4.5.

Согласно табл. 4.5, частота обнаружения аномалий в экспериментальной группе составляет 100%, а частота обнаружения аномалий в груп-

пе сравнения составляет около 82,04%. Ясно, экспериментальная группа может точно оценить нормальные данные. Коэффициент точности тестируемого метода в экспериментальной 95,67%. Коэффициент точности метода контрольной группы 74,69%, что значительно ниже, чем у метода обнаружения экспериментальной группы, что указывает на то, что точность обнаружения аномальных данных в экспериментальной группе высока, а количество ложных обнаружений невелико.

Таблица 4.5

Сравнение результатов тестирования между экспериментальной группой и группой сравнения

Номер набора данных	Метод экспериментальной группы		Метод контрольной группы	
	Коэффициент обнаружения аномалий, %	Коэффициент точности, %	Коэффициент обнаружения аномалий, %	Коэффициент точности, %
1	100	95.45	80.81	74.35
2	100	94.13	81.93	73.98
3	100	95.67	75.97	74.11
4	100	91.84	80.30	73.27
5	100	92.63	78.97	74.65
6	100	92.01	79.23	74.00
7	100	94.75	75.25	73.93
8	100	94.75	82.04	73.88
9	100	91.26	77.71	73.33
10	100	94.42	77.66	74.69
11	100	94.69	79.42	73.69
12	100	95.11	78.45	73.83

Таким образом, метод автоматического обнаружения аномальных

данных в системе учета электроэнергии, рассмотренной в работе, занимает меньше времени, обладает более высокой точностью обнаружения и большей практичностью.

Таким образом, в системе автоматизации учета электроэнергии обнаружение аномальных данных может принести больше пользы, чем анализ обычных данных. В работе предлагается использовать гибридный алгоритм обнаружения аномальных данных на основе алгоритма кластеризации и алгоритма iForest в качестве метода обнаружения аномалий. Проведенные эксперименты доказывают, что этот метод исследования по сравнению с традиционными методами требует меньше времени для обнаружения, а результаты обнаружения более надежны и что этот метод осуществим.

### **4.3. Система динамического мониторинга потерь больших объемов данных в мобильной сети на базе Интернета вещей**

#### ***4.3.1. Разработка архитектуры системы динамического мониторинга потерь больших объемов данных в мобильной сети на базе Интернета вещей***

Технология Интернета вещей заключается в подключении любого объекта к сети с помощью информационно-измерительного оборудования в соответствии с согласованным протоколом, а также в обмене и передаче информации с помощью информационных носителей для реализации интеллектуального контроля и других функций [4.19, 4.20]. В системе мониторинга используется беспроводная сенсорная сеть, спроектированы многоканальные узлы базовых станций, а сенсорные узлы расположены в разных каналах для осуществления мониторинга обмена данными. Расположение узлов в системе мониторинга, использующей технологию Интернета вещей, показано ниже (рис. 4.4).

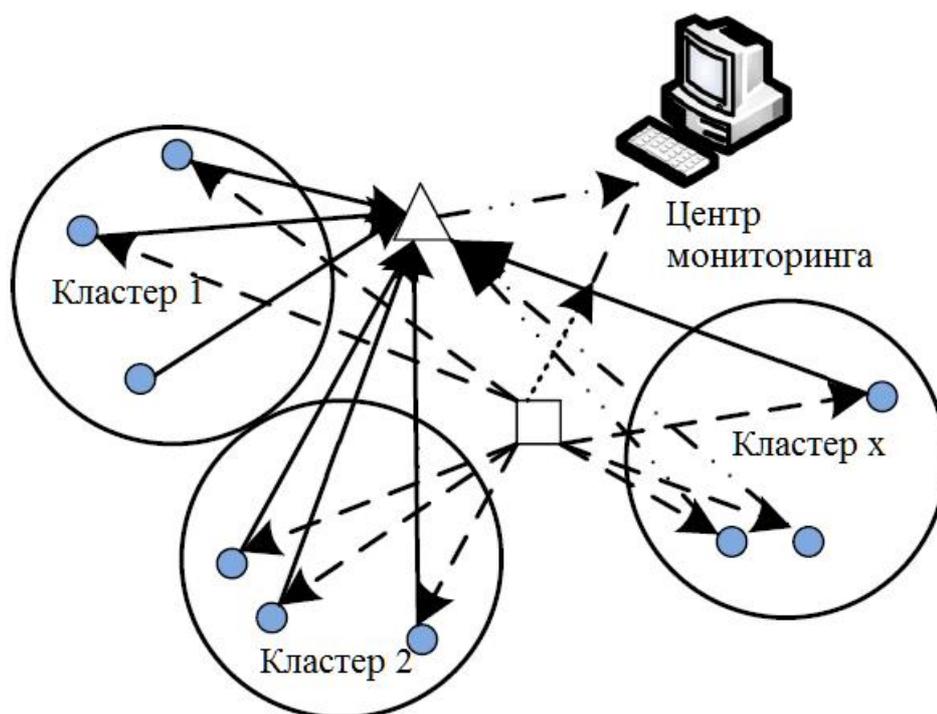


Рис. 4.4. Архитектура системы мониторинга:  $\dashv\dashv\dashv$  маршрутизация беспроводных каналов,  $\dashv\dashv$  синхронная маршрутизация,  $\cdots$  USB,  $\cdots\cdots$  последовательный канал,  $\triangle$  многоканальный узел базовой станции,  $\square$  узел управления,  $\bullet$  узел датчика

Многоканальный узел базовой станции и центр мониторинга соединены через шину USB 2.0, которая может удовлетворить требование одновременной загрузки 8 беспроводных каналов данных в центр мониторинга на высокой скорости и в полном объеме [4.21, 4.22]. Скорость передачи данных по сети удваивается по сравнению с первоначальными 250 Кбит/с для одного канала и может достигать 2 Мбит/с. Скорость шины USB 2.0 может достигать более 12 Мбит/с, что может решить проблему ограничения пропускной способности, вызванную использованием последовательного порта.

Аппаратная структура проектируемого многоканального узла базовой станции показана на рис. 4.5.

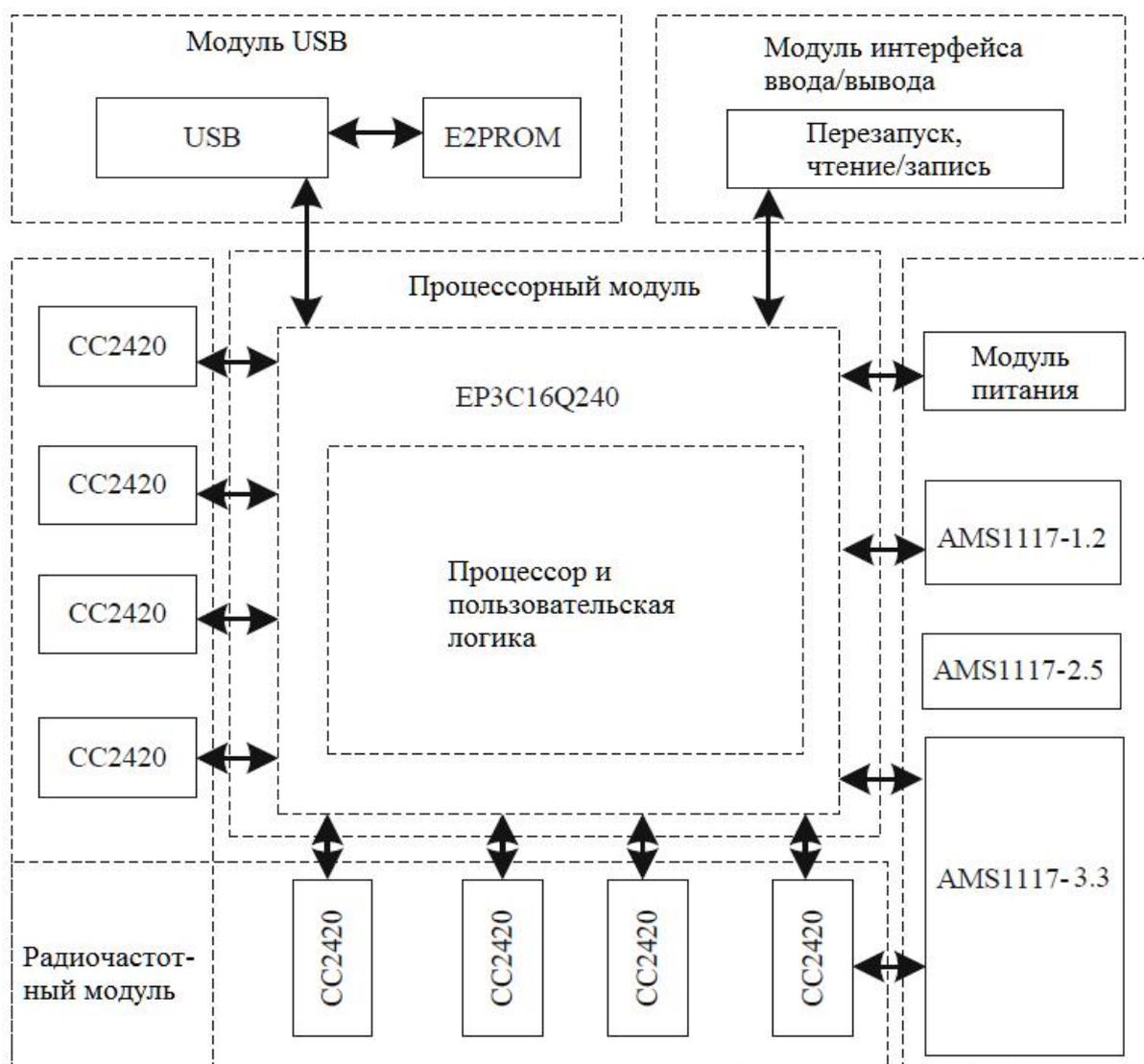


Рис. 4.5. Аппаратная структура многоканального узла базовой станции

Аппаратная часть многоканального узла базовой станции включает в себя микросхему FPGA, микросхему USB, 8 независимых радиочастотных модулей CC2420, электрически стираемую программируемую микросхему постоянной памяти, модуль управления питанием, модуль синхронизации и периферийные схемы каждого чипа. Среди них 8 радиочастотных модулей, каждый вывод USB-чипа и часть микросхемы FPGA могут быть подключены к выводу ввода-вывода, а функциональная конфигурация доступного вывода ввода-вывода микросхемы FPGA может быть реализована

с помощью программы, чтобы реализовать работу модуля PLD на радиочастотном модуле и модуле связи. Радиочастотный модуль CC2420 отвечает за прием и отправку беспроводных данных, а его первоначальная настройка и работа контролируются микросхемой FPGA. Несколько радиочастотных модулей CC2420 используются для параллельного приема данных по разным каналам, и между ними нет конкурирующих помех. После включения питания и инициализации USB-чипа поток данных загружается под управлением микросхемы FPGA.

FPGA - это продукт дальнейшего развития, основанный на программируемой матричной логике, общей матричной логике GAL, сложных программируемых логических устройствах и других программируемых устройствах. Она представляет собой полуавтоматическую схему в области специальных интегральных схем, которая не только устраняет недостатки пользовательских схем, но и преодолевает недостатки ограниченного числа вентилях оригинальных программируемых устройств. Учитывая универсальность конструкции узлов многоканальной базовой станции, при выборе микросхемы FPGA в основном ориентируются на продукцию Altera, которая производит микросхемы FPGA общего назначения. Материалы исследования включают руководство по выбору микросхем FPGA, руководство по эксплуатации микросхем и онлайн-предложения, предоставленные компанией Altera. Исследования показывают, что основными продуктами Altera являются микросхемы низкого класса (серия Cyclone IV), микросхемы среднего класса (серия ARIA) и т.д. С дальнейшим внедрением продуктов серии Cyclone серия Cyclone может обеспечить больше доступных контактов и ячеек памяти и еще больше снизить статическое энергопотребление, что позволяет выполнять функцию многоканального узла базовой станции, а стоимость чипов серии Arria значительно ниже. Согласно проведенному анализу, узлу многоканальной базовой станции требуется больше доступных выводов ввода-вывода и логических блоков,

поэтому необходимо выбрать соответствующую микросхему FPGA в качестве ядра обработки данных для многоканальных узлов базовой станции серий Cyclone II и Cyclone III. Сравнение конкретных параметров серий Cyclone II и Cyclone III показано в табл. 4.6.

Таблица 4.6

Параметры микросхем серий Cyclone II и Cyclone III

Модельный логический блок - модуль оперативной памяти PLL, доступные контакты ввода-вывода				
EP2C5	4.608	119.808	2	142
EP2C8	8.256	165.888	2	138
EP2C20	18.752	239.616	4	142
EP3C5	5136	423936	2	106
EP3C10	10320	423936	2	106
EP3C16	15408	516096	4	160
EP3C25	24624	608256	4	148

Учитывая фактическое применение мониторинга, необходимо завершить настройку функции USB-чипа, управлять USB-чипом, чтобы реализовать функцию загрузки данных с высокой скоростью; получать статус RF-модуля в режиме реального времени и управлять 8 RF-модулями, чтобы завершить операцию параллельного приема данных с высокой скоростью. скорость зависит от статуса приема пакета данных; инкапсулировать и верифицировать полученные пакеты данных, а также использовать механизм HDLC для анализа пакетов данных, чтобы гарантировать, что программа главного компьютера может быть выполнена с высокой скоростью. Для обеспечения корректности пакетов данных, загружаемых в центр мониторинга, используется механизм проверки CRC. В заключение, в качестве основного чипа многоканального узла базовой станции выбран чип EP3c16q240 серии Cyclone III.

#### ***4.3.2. Разработка программного обеспечения системы динамического мониторинга потерь больших объемов данных в мобильной сети***

На основе описанного выше аппаратного обеспечения разработана программная часть системы мониторинга. Динамический мониторинг потерь больших объемов данных в мобильной сети можно разделить на три уровня: мониторинг кластера, мониторинг производительности узлов и мониторинг работы заданий (рис. 4.6).

Управление узлами и планирование их работы является важной частью кластеров больших данных. От того, хорошо ли работает кластер, напрямую зависит эффективность всего кластера, поэтому система должна отслеживать различные информационные ресурсы каждого узла и управлять ими в режиме реального времени. Администратор может настроить кластер в режиме реального времени в соответствии с производительностью кластера, устранить узкое место в кластере и обеспечить бесперебойную работу задания.

Агент мониторинга кластера в основном отвечает за сбор индексных данных мониторинга кластера. Для мониторинга кластера можно просматривать информацию о состоянии кластера, общее количество компьютеров в кластере, пользователей кластера, группы хостов и другую информацию. Сбор данных о производительности кластера в основном осуществляется с помощью командной строки. Например, время безотказной работы: просмотр загрузки компьютера, вывод количества процессов, ожидающих ресурсов процессора, и количества процессов, заблокированных в режиме бесперебойного ввода-вывода; `mpstat-pall`: отображение загрузки каждого процессора; `pidstat1`: выходные данные процессора процесса непрерывного вывода не будут охватывать данные, полученные ранее.



Рис. 4.6. Мониторинг потерь больших объемов данных

Подключаемый модуль кластерного мониторинга в основном отвечает за получение данных индикатора мониторинга, отправляемых процессом агента. При мониторинге кластера процесс внедрения подключаемого модуля заключается в использовании механизма расширения Icing для ин-

капсуляции команды для получения данных о производительности кластера и использовании пользовательского механизма расширения подключаемого модуля Icing для формирования подключаемого модуля, изменения файла конфигурации Icing и подключения подключаемого модуля Deployed на сервере Icing.

Мониторинг производительности узлов в основном предназначен для мониторинга базовых узлов и сбора различных показателей производительности, чтобы менеджеры могли своевременно и всесторонне получать информацию о производительности узлов кластера. Модуль мониторинга производительности узлов в основном разработан также с использованием Icing. Агент мониторинга производительности узла отвечает за сбор данных о производительности узла и обработку данных, включая модуль сбора данных мониторинга производительности и модуль обработки данных мониторинга производительности.

Модуль сбора данных мониторинга включает в себя мониторинг общих услуг и мониторинг частных сервисов. Для мониторинга общих услуг можно получить доступ, используя некоторые стандартные протоколы сети общего пользования, такие как HTTP, POP3, IMAP, FTP и SSH. Некоторые собственные подключаемые модули Icing могут напрямую собирать данные о производительности. В результате нет необходимости устанавливать агенты на узлы для мониторинга этих служб. Частная служба является противоположностью общедоступной. Доступ к частной службе невозможен через общедоступную сеть. Прямой доступ к информации через сеть невозможен. Поэтому при мониторинге частных сервисов на контролируемом объекте должен быть установлен соответствующий агент.

В модуле получения данных мониторинга производительности, после того как агент Icing возвращает данные индекса производительности, Icing поддерживает два метода обработки данных о производительности, то есть использование командной строки для обработки или записи непо-

средственно в определенный файл, который можно настроить с помощью файла конфигурации. Если используется первый способ, определяется обработчик данных о производительности подключаемого модуля в файле конфигурации Icing. После запуска подключаемого модуля запускается обработчик данных о производительности. В это время обработчик производительности получает данные из переменных среды. Icing запускает подключаемый модуль через командную строку для выполнения соответствующей проверки состояния, а затем перехватывает стандартный выходной поток подключаемого модуля для получения результатов после выполнения. Результат выполнения Icing включает в себя, по крайней мере, одну строку удобочитаемого текста для представления текущего состояния, и данные о производительности, относящиеся к подключаемому модулю, также включаются в результат. Если используется второй метод, Icing может напрямую поместить полученные данные о производительности в файл данных о производительности, использовать компонент для обработки файла или настроить команду в Icing для периодической обработки файла данных о производительности.

Мониторинг работы заданий является важной частью системы мониторинга, а объектом мониторинга являются данные о состоянии работы мобильной сети. Мониторинг заданий включает в себя модуль сбора данных, модуль обработки данных и модуль отправки данных. Объектом сбора данных являются журналы, созданные во время работы мобильной сети. IDEA упаковывает исполняемый шелл-код и зависимые jar-пакеты через submit, генерирует соответствующие jar-файлы и отправляет их в кластер через командную строку. Что касается задач, сохраняются журналы выполнения в рабочем каталоге и создается новая папка для каждой задачи, чтобы сохранить файлы журналов и зависимые пакеты jar. Информацию о состоянии больших данных мобильной сети можно получить из журнала.

После завершения сбора данных они фильтруются и извлекаются,

сохраняются в определенной структуре данных и заполняются агентом, находящимся в узле кластера. В модуле обработки данных пользователь вызывает функциональный модуль обработки данных после сбора данных о производительности. Сначала задача обработки данных мобильной сети генерирует журналы, которые хранятся в рабочей папке. Используя метод сбора журналов, основанный на анализе текста, Readfile анализирует и обрабатывает собранные журналы и получает идентификатор задания рабочей задачи Readformfile, обрабатывает собранные журналы, извлекает показатели выполнения каждой задачи, включая показатели мониторинга, такие как пользователь, время запуска, продолжительность, информация о состоянии выполнения, и сохраняет их в заранее определенной структуре данных. В SNMPSend все показатели мониторинга объединяются для формирования пакета данных SNMP, готового к отправке.

Модуль отправки данных использует механизм, управляемый событиями, для отправки информации мониторинга по протоколу SNMP. Он не требует опроса сервера мониторинга, экономит пропускную способность и использование сервера, а также обеспечивает производительность системы.

#### **4.4. Экспериментальное исследование системы динамического мониторинга потерь больших данных в мобильной сети**

Для проверки эффекта временной синхронизации системы динамического мониторинга потерь больших объемов данных в мобильной сети была создана экспериментальная платформа для проверки функции временной синхронизации. В ходе эксперимента система мониторинга используется для сбора сигналов фиксированной частоты, генерируемых генератором сигналов в режиме реального времени. После проведения эксперимента путем анализа собранных данных вычисляется ошибка синхронизации по времени сбора данных сенсорным узлом, а затем сравнивается

эффективность синхронизации сбора данных традиционной системы мониторинга и разработанной системы динамического мониторинга потерь больших данных в мобильной сети.

#### ***4.4.1. Архитектура платформы эксперимента***

Установлено, что большинство экспериментов основаны на компьютерном моделировании. Кроме того, на основе экспериментальной платформы беспроводной сенсорной сети для мониторинга состояния конструкций исследователи проверили эффект синхронизации при подключении к сети. Экспериментальный метод, который они использовали, состоял в том, чтобы с помощью генератора сигналов генерировать сигнал фиксированной частоты для сбора сенсорной сетью, а затем анализировать данные, собранные всеми узлами. На основе этого метода разработан экспериментальный метод, основанный на получении стандартной формы сигнала, для проверки эффективности синхронизации системы мониторинга.

Оборудование, использованное в эксперименте, включало 16 сенсорных узлов Telosb, узел двойного радиочастотного ретранслятора, узел управления, генератор сигналов произвольной формы AFG3021 и ноутбук. Генератор сигналов произвольной формы AFG3021 имеет точность вывода 14 бит. Из них 16 узлов разделены на 8 различных каналов. В ходе эксперимента генератор сигналов произвольной формы AFG3021 был настроен на генерацию стандартного тестируемого сигнала для сбора 16 узлов. Форма волны, генерируемая измеренным сигналом, задается как сигнал треугольной формы, а выходной конец подключается к выводам ADC0 и GND в 16 узлах с несколькими выводами. Диапазон напряжений выходного сигнала составляет от 0 до 2.5, чтобы соответствовать допустимому диапазону входного напряжения канала дискретизации AD узла Telosb и избежать повреждения узла Telosb. Угол наклона восходящего фронта выходной треугольной волны установлен равным 300 В/с. Поскольку точ-

ность дискретизации AD узла Telosb составляет 12 бит, диапазон значений дискретизации AD может быть рассчитан от 0 до 4960. Опорное напряжение, выбранное узлом Telosb, составляет 2,5 В, а формула расчета разрешения по напряжению выглядит следующим образом:

$$P = \frac{1}{m} \cdot 2.5V \quad (4.10)$$

В формуле  $\mu$  представляет собой значение выборки AD. Согласно приведенной выше формуле, когда значение выборки AD изменяется на 1, измеренное значение напряжения соответственно изменяется на  $P$ . Следовательно, когда значение выборки AD изменяется на 1, соответствующее изменение времени нарастания выходного сигнала рассчитывается следующим образом:

$$t = \frac{P}{K} \quad (4.11)$$

где  $k$  - угол наклона восходящего фронта выходного сигнала.

При одновременном сборе данных несколькими сенсорными узлами, если время сбора данных полностью синхронизировано, данные, собранные с использованием одних и тех же данных о серийном номере, должны быть согласованными. При возникновении последовательной ошибки во время сбора данных узлом вы можете использовать значения выборки AD, собранные разными узлами в одном цикле сбора, для вычисления значения напряжения, собранного в это время, и среднего значения напряжения сбора данных для всех узлов датчиков. Затем значение напряжения, полученное каждым узлом, вычитается из среднего значения напряжения, и ошибка синхронизации времени узла относительно среднего времени выборки вычисляется по формуле (4.12). Формула выглядит следующим образом:

$$Dt = \frac{DV}{k} \quad (4.12)$$

В формуле  $\Delta V$  представляет собой разницу между полученным значением напряжения и средним значением напряжения.

Весь эксперимент длился 2 ч. Перед началом эксперимента включался генератор сигналов произвольной формы, точно отрегулирована форма выходного сигнала в соответствии с установленным значением, подключен выходной вывод положительного напряжения к выводам ADC0 всех сенсорных узлов и подключен выходной вывод отрицательного напряжения к выводам GND всех сенсорных узлов. После начала эксперимента узел управления главным компьютером отправляет команду запуска всем сенсорным узлам, чтобы начать мониторинг сети. В ходе рабочего процесса узел управления синхронизирует время сбора данных со всех сенсорных узлов в сети в соответствии со своей собственной программой. Узел базовой станции в системе мониторинга получает пакет данных, отправленный сенсорным узлом, в режиме реального времени и загружает его на верхний компьютер для хранения, ожидая окончания эксперимента для анализа.

#### ***4.4.2. Результаты экспериментов и их анализ***

После завершения эксперимента, в соответствии с номером пакета данных, хранящегося в верхнем компьютере, случайным образом выбрано три момента времени, записаны значения выборки AD, собранные всеми узлами датчиков в эти три момента времени, и преобразованы в значения напряжения, как показано в табл. 4.7.

В табл. 4.7 приведены экспериментальные данные измерений предлагаемой системы динамического мониторинга потерь больших данных в мобильной сети на основе физической сети, и эффект синхронизации не может быть виден интуитивно. По этой причине данные измерений двух других систем мониторинга не приводятся последовательно, поэтому можно напрямую сравнить результаты измерений временной ошибки, джиттера и эффекта синхронизации. Ниже приведены экспериментальные результаты традиционной системы мониторинга на базе ZigBee (табл. 4.8).

Таблица 4.7

Данные экспериментальных измерений синхронного эффекта

Номер узла	Измеренное напряжение		
	1	2	3
3	1.1684	1.2019	1.3412
6	1.1721	1.2341	1.3681
9	1.1745	1.2471	1.3124
11	1.1617	1.2036	1.3369
13	1.1702	1.2571	1.3154
15	1.1629	1.2347	1.3274
17	1.1794	1.2367	1.3096
19	1.1624	1.2903	1.3264
21	1.1454	1.2036	1.3746
22	1.1564	1.2461	1.3091
23	1.1754	1.2412	1.3325

Таблица 4.8

Анализ эффекта синхронизации системы мониторинга на базе ZigBee

Показатель	Измеренное напряжение		
	1	2	3
Максимум	1.2936	1.4216	1.6047
Минимум	1.1021	1.3011	1.4214
Максимальная разница	0.1915	0.1205	0.1833
Среднее значение	1.1974	1.3824	1.5691
Ошибка синхронизации	232	304	357

Колебание временной погрешности представлено на рис. 4.7.

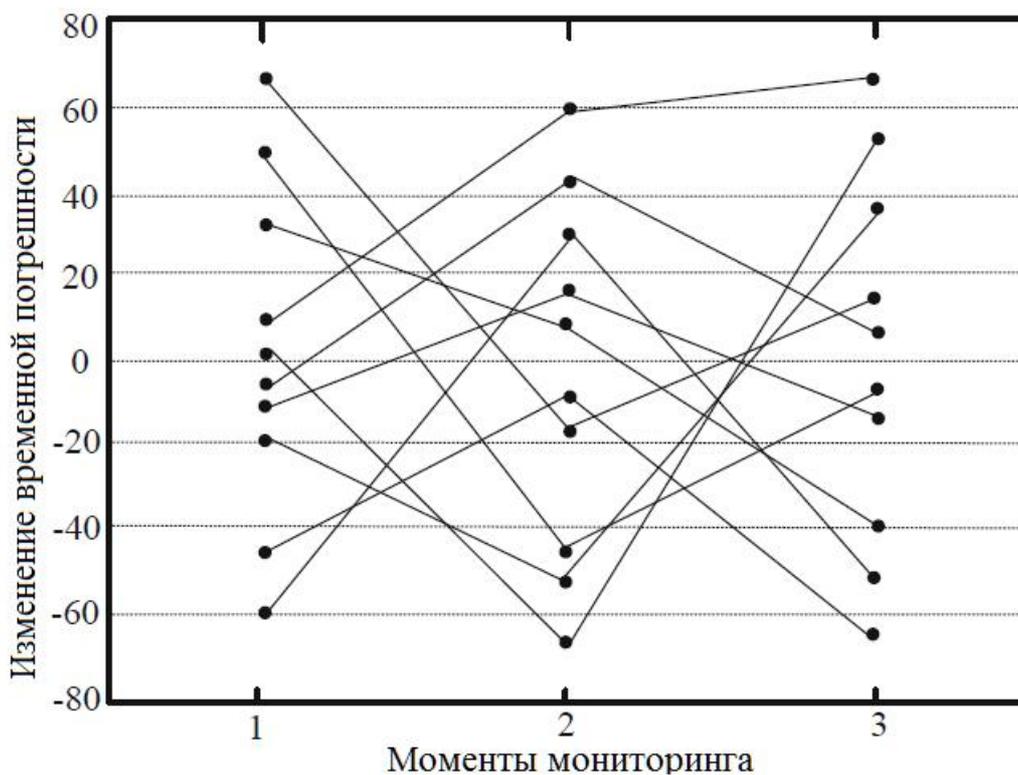


Рис. 4.7. Изменение временной погрешности системы на базе ZigBee

Ниже приведены результаты экспериментов с веб-системой мониторинга (табл. 4.9).

Таблица 4.9

Анализ влияния синхронизации на веб-систему мониторинга

Показатель	Измеренное напряжение		
	1	2	3
Максимум	1.3147	1.5007	1.6214
Минимум	1.1724	1.2374	1.4025
Максимальная разница	0.1423	0.2633	0.2189
Среднее значение	1.2371	1.3725	1.506
Ошибка синхронизации	201	262	304

Колебания временной ошибки представлены на рис. 4.8.

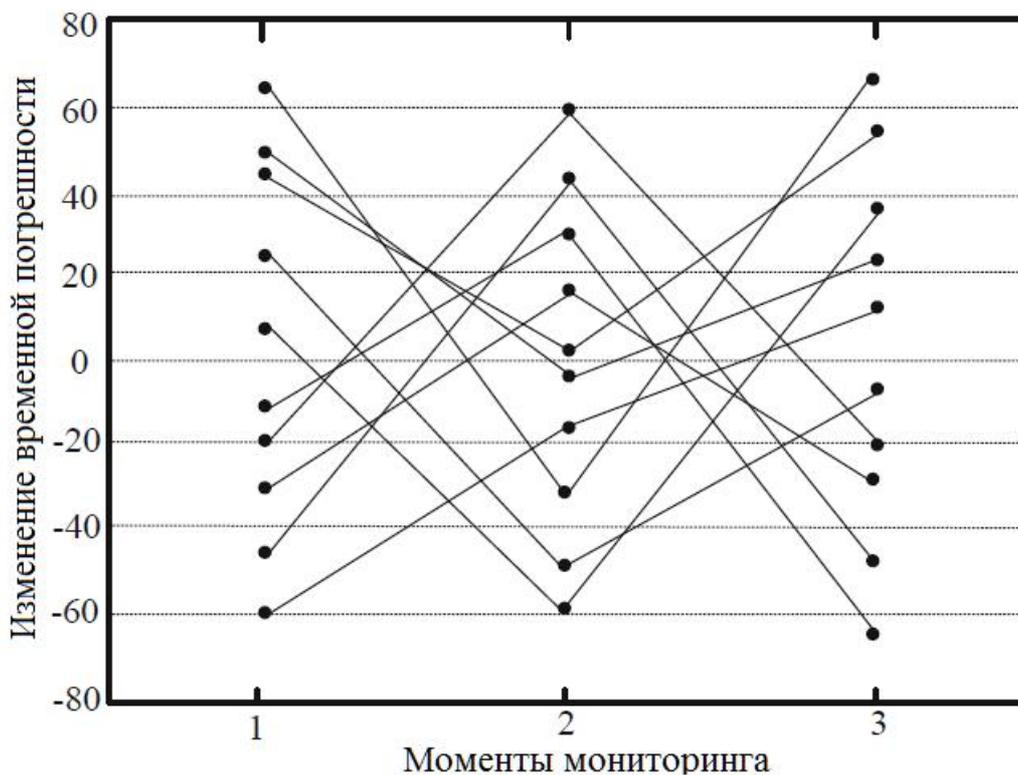


Рис. 4.8. Изменение временной погрешности веб-системы

Ниже приведены результаты экспериментов с разработанной системой мониторинга на основе физической сети (табл. 4.10).

Таблица 4.10

Анализ эффекта синхронизации системы мониторинга на основе физической сети

Показатель	Измеренное напряжение		
	1	2	3
Максимум	1.1794	1.2903	1.3746
Минимум	1.1454	1.2019	1.3091
Максимальная разница	0.034	0.0884	0.0655
Среднее значение	1.1663	1.2360	1.332
Ошибка синхронизации	29 мкс	38 мкс	49 мкс

Ошибка изменения времени представлена на рис. 4.9.

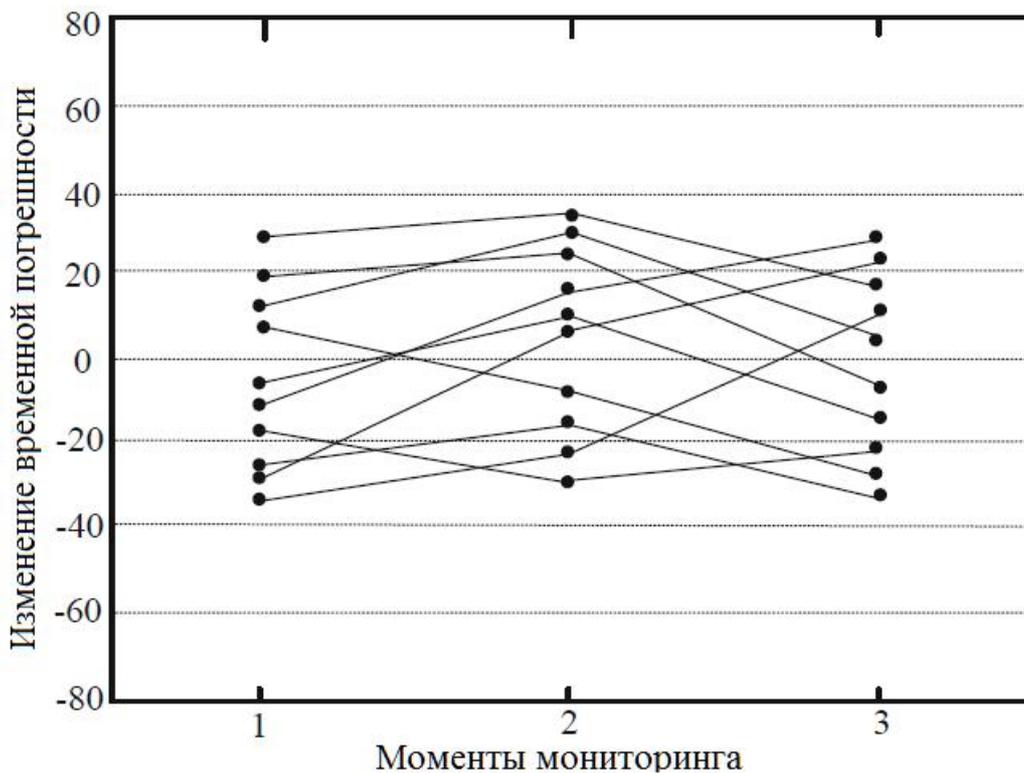


Рис. 4.9. Изменение системной ошибки времени на основе Интернета вещей

Точки на рис. 4.9 представляют собой узлы. Наблюдая за приведенными выше результатами, можно сделать вывод о том, что диапазон изменения временной погрешности относительного среднего значения на рис. 4.9 уже, чем у двух других результатов, и погрешность синхронизации, показанная в таблице анализа эффекта синхронизации измеренного напряжения, всегда находится в пределах 90 мкс, в то время как погрешность синхронизации системы мониторинга, основанной на ZigBee и Web, превышает 200 мкс, что значительно превышает стандартный диапазон временных ошибок. В заключение отметим, что разработанная система динамического мониторинга потерь больших объемов данных в мобильной сети на основе Интернета вещей имеет меньшую ошибку синхронизации по времени. В пределах допустимого диапазона нормального применения системы система обладает хорошей функцией синхронизации сбора данных.

Для дальнейшей проверки эффективности системы, описанной в статье, традиционная система на базе ZigBee и система Интернета вещей используются для сравнения и анализа точности динамического мониторинга потерь больших данных в мобильной сети. Результаты сравнения показаны на рис. 4.10.

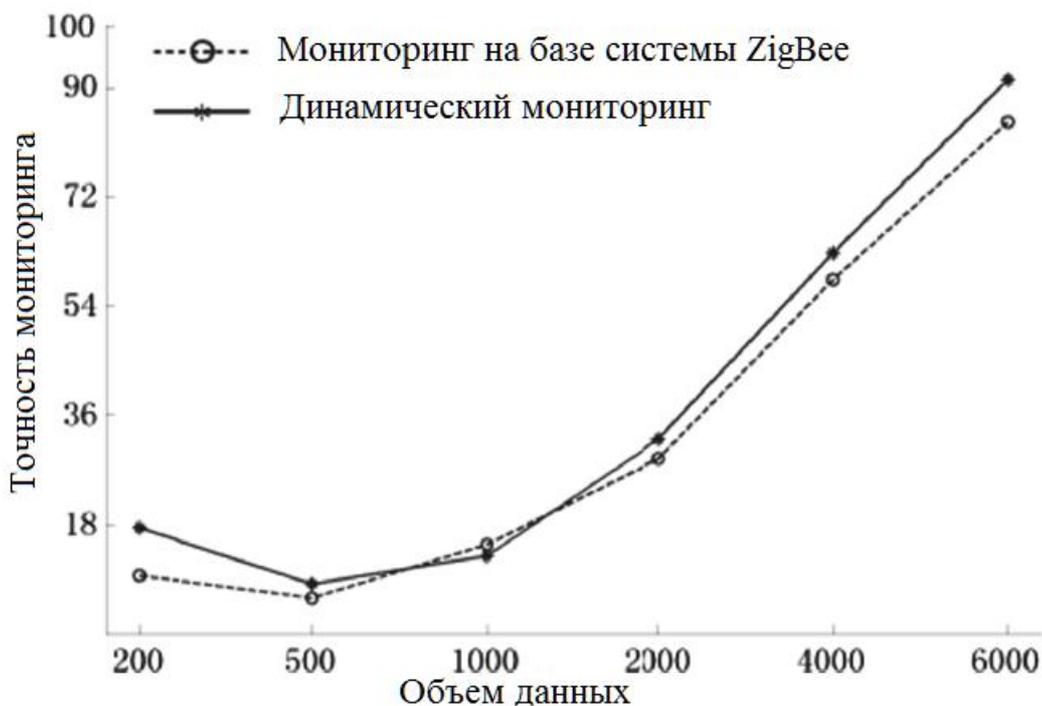


Рис. 4.10. Результаты сравнения точности мониторинга

Согласно рис. 4.10, максимальная точность динамического мониторинга потерь больших данных в мобильной сети может достигать 98%, в то время как максимальная точность динамического мониторинга потерь больших данных в мобильной сети на базе системы ZigBee составляет всего 82%, что показывает, что точность динамического мониторинга потерь больших данных в мобильной сети превосходит традиционную систему мониторинга потерь больших данных в мобильных сетях на основе системы ZigBee.

Большие данные в мобильных сетях - это тенденция общественного развития, требующая повышенного внимания к проблемам сетевой безо-

пасности. Используя характеристики Интернета вещей и больших данных, в работе представлена система динамического мониторинга потерь больших данных в мобильных сетях на основе Интернета вещей. Используя чип er3c16q240 в качестве основного управляющего чипа многоканального узла базовой станции, спроектировано аппаратное обеспечение системы, а системное программное обеспечение разработано с использованием функций мониторинга кластера, мониторинга производительности узлов и мониторинга выполнения заданий, что завершает разработку системы динамического мониторинга потерь больших объемов данных в мобильной сети на основе Интернета вещей. Решены проблемы, существующие в традиционной системе мониторинга, и заложен хороший фундамент для будущего развития системы мониторинга [4.23].

#### **4.5. Выводы к главе 4**

1. В системе автоматизации учета электроэнергии обнаружение аномальных данных может принести больше пользы, чем анализ обычных данных. Предложен гибридный алгоритм обнаружения аномальных данных на основе алгоритма кластеризации и алгоритма iForest в качестве метода обнаружения аномалий. Проведенные эксперименты доказывают, что этот метод исследования по сравнению с традиционными методами требует меньше времени для обнаружения, а результаты обнаружения более надежны и что этот метод осуществим.

2. В работе представлена система динамического мониторинга потерь больших данных в мобильных сетях на основе Интернета вещей. Спроектировано аппаратное обеспечение системы, а системное программное обеспечение разработано с использованием функций мониторинга кластера, мониторинга производительности узлов и мониторинга выполнения заданий, что завершает разработку системы динамического мониторинга потерь больших объемов данных в мобильной сети на основе Интернета

вещей.

3. Максимальная точность динамического мониторинга потерь больших данных в мобильной сети может достигать 98%, в то время как максимальная точность динамического мониторинга потерь больших данных в мобильной сети на базе системы ZigBee составляет всего 82%, что показывает, что точность динамического мониторинга потерь больших данных в мобильной сети превосходит традиционную систему мониторинга потерь больших данных в мобильных сетях на основе системы ZigBee.

#### **Литература к главе 4**

- 4.1. Yang X., Qu Y., Pang H. et al. Power metering pipeline fault warning technology based on deep learning algorithm// *Electron. Des. Eng.* 28(04), 153–157 (2020)
- 4.2. Gao S., Li C. An improved spectral clustering algorithm for anomaly detection of power data// *Comput. Simul.* 36(11), 239–242 + 304 (2019)
- 4.3. Yang J., Zeng X., Yao L. et al. Research on abnormal electricity monitoring based on large data mining// *Autom. Instrum.* 08, 219–222 (2019)
- 4.4. Tong X., Yu S. Fault detection algorithm for transmission lines based on random matrix spectrum analysis// *Autom. Electr. Power Syst.* 43(10), 101–115 (2019)
- 4.5. Xu G., Ning B., Zhong Y. Automatic matching of voltage blackout events in metering automation system// *Electron. Test* 04, 111–112 (2019)
- 4.6. Huang J., Dai B., Zhang L. et al. Study on dynamic identification of abnormal data of electric energy measurement device// *Guangxi Electr. Power* 41(04), 53–55 + 64 (2018)
- 4.7. Chen Q., Zheng K., Kang C. et al. Detection methods of abnormal electricity consumption behaviors: review and prospect// *Autom. Electri. Power Syst.* 42(17), 189– 199 (2018)
- 4.8. Liu S., Glowatz M., Zappatore M. et al. *E-Learning, e-Education, and Online Training*. Springer, Heidelberg (2018), 374 p.
- 4.9. Zhang L., Shao F. Design of abnormal risk monitoring system for network big data platform// *Mod. Electron. Tech.* 41(22), 143–146 (2018)
- 4.10. Du S. Application of online monitoring system in urban rail transit based on the big data// *Urban Mass Transit* 21(S2), 30–33 (2018)
- 4.11. Chen, Z. Sun J. Simulation of mobile network information transmission security defense under big data// *Comput. Simul.* 35(05), 207–210 (2018)
- 4.12. Zhang Y., Liu K., Yang L. et al. Platform construction and data processing application technology in coal industry monitoring big data// *Coal*

Sci. Technol. 47(03), 75–80 (2019)

4.13. Leng X., Chen G., Jiang Y. et al. Data specification and processing in big-data analysis system for monitoring and operation of smart grid// Autom. Electr. Power Syst. 42(19), 169–178 (2018)

4.14. Liu S., Lu M., Li H. et al. Prediction of gene expression patterns with generalized linear regression model// Front. Genet. 10, 120 (2019)

4.15. Liu W., Zong L., Xing C. et al. Design of the overall information collection of wind farm monitoring system based on EDPF-CP system// Renew. Energy Resour. 36(08), 1204–1208 (2018)

4.16. Deng Z., Cui J., Liang Z. Condition monitoring system of mine hoist based on storage test// Coal Technol. 38(05), 179–181 (2019)

4.17. Deng M. Regulation and protection of personal data in the context of big data// J. Beijing Univ. Posts Telecommun. 21(01), 19–25 (2019)

4.18. Zheng P., Shuai L., Arun S., Khan M. Visual attention feature (VAF): a novel strategy for visual tracking based on cloud platform in intelligent surveillance systems// J. Parallel Distrib. Comput. 120, 182–194 (2018)

4.19. Fu X., Gao Y., Luo B. et al. Security threats to HADOOP: data leakage attacks and investigation// IEEE Netw. PP(2), 12–16 (2017)

4.20. Buesing H., Vogt C., Ebigbo A. et al. Numerical study on CO<sub>2</sub> leakage detection using electrical streaming potential (SP) data// Water Resour. Res. 53(1), 455–469 (2017)

4.21. Liu S., Liu D., Srivastava G. et al. Overview and methods of correlation filter algorithms in object tracking// Complex Intell. Syst., 2021. Vol. 7, pp. 1895–1917.

4.22. Blackford J., Artioli Y., Clark J. et al. Monitoring of offshore geological carbon storage integrity: implications of natural variability in the marine system and the assessment of anomaly detection criteria// Int. J. Greenhouse Gas Control 64, 99–112 (2017)

4.23. Lu M., Liu S. Nucleosome positioning based on generalized relative

entropy// Soft. Comput. 23(19), pp. 9175–9188 (2018).

## Заключение

Целью работы являлось создание моделей анализа и алгоритмов управления процессами разработки специального энергоэффективного программного обеспечения перемещения данных в беспроводных информационных системах.

В процессе выполнения диссертационного исследования получены следующие основные результаты:

В процессе выполнения диссертационного исследования получены следующие основные результаты:

1. Создана модель снижения энергопотребления процессов перемещения данных в беспроводных информационных системах, обеспечивающая повышение эффективности узла слияния по мере увеличения количества кластеризуемых узлов, а также их индивидуальных уровней энергии.

2. Предложена модификация стратегии управления данными на основе агентного подхода, позволяющая повысить энергоэффективность маршрутизации в крупномасштабных беспроводных сенсорных сетях.

3. Разработан алгоритм улучшения показателей энергоэффективной связи объектов беспроводных информационных систем, обеспечивающий сокращение общего времени, необходимого для завершения процесса кластеризации.

4. Разработан гибридный итерационный алгоритм поиска аномальных данных, сочетающий применение матрицы сходства аномальных признаков данных и кластерный анализ с использованием обучения iForest, обеспечивающий улучшение точности обнаружения в среднем на 22%.

5. Разработана структура программного комплекса динамического мониторинга потерь больших объемов данных с многоуровневым сбором информации (уровни кластера, производительности узлов и работы заданий), обеспечивающая максимальную точность мониторинга до 98%,

6. Элементы программного обеспечения зарегистрированы в ФИПС.

### **Рекомендации и перспективы дальнейшей разработки темы**

Результаты исследования рекомендуются к применению в задачах исследования и разработки программного обеспечения в беспроводных информационных системах.

Дальнейшая разработка темы будет направлена на практическую реализацию теоретических и алгоритмических результатов, интеграцию в проекты наиболее распространенных распределенных систем. Развитие результатов будет направлено на дальнейшее снижение энергопотребления беспроводных коммуникационных систем.

## Список использованных источников

1. Анализ методов и моделей распределения местоположения узлов в сетях для оценки эффективности работы беспроводных датчиков / Али Хусейн, В.В. Лавлинский // Моделирование систем и процессов. – Вып. 2.– 2020. – С.27-34.
2. Атласов Д.И., Сотников Д.В., Васми Ихаб А Васми, Хуссейн Али Иед, Линкина А.В. Типовой интерфейс облачных вычислений. Свидетельство о регистрации программы для ЭВМ № 2025681822 от 18.08.2025. - М.: Роспатент, 2025.
3. Имитационная модель распределения местоположения узлов в сетях для оценки эффективности работы беспроводных датчиков / Али Хусейн, В.В. Лавлинский // Моделирование систем и процессов. – Т. 13, Вып. 2. - 2020. – С.35-42
4. Хуссейн А.И. Математическое и программное обеспечение для повышения энергоэффективности протокола управления потоками данных со скрытой Марковской моделью в беспроводных сенсорных сетях// Информатика. Экономика. Управление - Informatics. Economics. Management, 4(2), 2025, pp. 3007–3013. <https://doi.org/10.47813/2782-5280-2025-4-2-3007-3013>.
5. Хуссейн Али Иед, Баталов Д.И., Красновский Е.Е. Модель интеграции данных для беспроводных сенсорных сетей, направленная на повышение энергоэффективности// Системы управления и информационные технологии, №3(97), 2024. С. 93-100.
6. Хуссейн Али Иед, Рындин Н.А. Моделирование управления межобъектным взаимодействием в беспроводных сенсорных сетях// Сб. тр. VI Всеросс. НПК «Информационные технологии в экономике и управлении». – Махачкала, 2024. С. 96-102.
7. Хуссейн Али Иед, Рындин Н.А. Особенности использования параметров управления потоками данных для стохастического поиска крат-

чайших путей// Экономика и менеджмент систем управления, №1(55), 2025. – С. 90-100

8. Хуссейн Али Иед, Рындин Н.А. Стохастический поиск кратчайших путей с использованием параметров управления потоками данных на основе протокола динамической маршрутизации для мобильных AD-НОС сетей// Вестник Воронежского государственного технического университета. – 2025. – Т. 21, № 4. – С. 78-86.

9. Хуссейн Али Иед, Рындин Н.А., Божко Л.М., Мутина Е.И. Моделирование мониторинга аномалий в системе автоматизации измерения мощности// Системы управления и информационные технологии, №2.1(100), 2025. С. 74-80.

10. Хуссейн Али Иед, Рындин Н.А., Питолин А.В. Архитектура интеллектуальной системы динамического мониторинга потерь больших объемов данных в мобильной сети// Системы управления и информационные технологии, №3(101), 2025. С. 69-77.

11. Хуссейн Али Иед. Численное исследование протокола LEACH в качестве полураспределенного протокола кластеризации для беспроводных сенсорных сетей// Информационные технологии моделирования и управления, №3(137), 2024. – С. 199-208.

12. Хуссейн А. Исследование эффективности работы беспроводных датчиков на основе местоположения узлов в сетях// Интеллектуальные информационные системы: тр. Междунар. НПК: в 2 ч.. Ч. 2. - Воронеж: ВГТУ, 2021. - С. 153-157

13. Хуссейн А.Е. Математическая модель оценки эффективности работы беспроводных датчиков на основе местоположения узлов в сетях// Оптимизация и моделирование в автоматизированных системах: тр. Междунар. молодежной научной шк. – Воронеж: ВГТУ, 2022. – С. 116-121.

14. Afzaal A.A., Kamal A. An intelligent neural-WSN based schema for energy resources forecast// *International Journal of Advanced Science and Technology*, 2011, vol. 33, pp. 121130.
15. Akyildiz IF, Weilian S, Sankarasubramaniam Y, Cayirci E (2002) *Wireless sensor networks: a survey*. *Comput Netw* 38(4):393–422.
16. Almesaeed R, Jedidi A (2021) Dynamic directional routing for mobile wireless sensor networks. *Ad Hoc Netw* 110:102301.
17. Anand S, Rafeeqe KM (2022) Enhancing the security in wireless sensor network using hidden markov model. *Soft computing for security applications*. Springer, Berlin, pp 409–423.
18. Bahbahani MS, Alsusa E (2017) A cooperative clustering protocol with duty cycling for energy harvesting enabled wireless sensor networks. *IEEE Trans Wirel Commun* 17(1):101–111.
19. Balamurugan A. An energy efficient fitness based routing protocol in WSN, *ICTACT Journal on Communication Technology*, 2014, vol. 5, no. 1, pp. 894-899.
20. Basheer A, Sha K (2017) Cluster-based quality-aware adaptive data compression for streaming data. *J Data Inf Qual (JDIQ)* 9(1):1–33.
21. Ben-Othman J., Bashir Y. Energy efficient and QoS based routing protocol for WSN// *Journal of Parallel and Distributed Computing*, 2010, vol. 70, no. 8, pp. 849-857.
22. Blackford J., Artioli Y., Clark J. et al. Monitoring of offshore geological carbon storage integrity: implications of natural variability in the marine system and the assessment of anomaly detection criteria// *Int. J. Greenhouse Gas Control* 64, 99–112 (2017)
23. Bokare M., Ralegaonkar A. WSN: A Promising Approach for Distributed Sensing Task// *Excel Journal of Engineering Technology and Management Science*, 2012, vol. 1, no. 1, pp. 1-9.

24.Boonsawat V., Jurarat E., Kulwadee B., Somsak K. ‘XBee WSN for temperature monitoring// The second conference on application research and development, Chon Buri, Thailand. 2010, 117.

25.Boudries A, Aliouat M, Siarry P (2014) Detection and replacement of a failing node in the wireless sensors networks. *Comput Electr Eng* 40(2):421–432.

26.Budianu C, Ben-David S, Tong L (2006) Estimation of the number of operating sensors in large-scale sensor networks with mobile access. *IEEE Trans Signal Process* 54(5):1703–1715.

27.Buesing H., Vogt C., Ebigbo A. et al. Numerical study on CO<sub>2</sub> leakage detection using electrical streaming potential (SP) data// *Water Resour. Res.* 53(1), 455–469 (2017)

28.Capo-Chichi EP, Guyennet H, Friedt JM (2009) K-rle: a new data compression algorithm for wireless sensor network. In: 2009 third IEEE international conference on sensor technologies and applications, pp 502–507.

29.Chakraborty A., Swarup K.M., Mrinal K.N. A Genetic algorithm inspired routing protocol for WSN// *International Journal of Computational Intelligence Theory and Practice*, 2011,vol. 6, no. 1, pp. 1-8.

30.Chaudhary S., Singh N., Pathak A., Vatsa A.K. Energy Efficient Techniques for Data aggregation and collection in WSN// *Int. J. Comp. Sci. Eng. Appl.*, 2012, vol. 2, no. 4, pp.37-40.

31.Chen Q., Zheng K., Kang C. et al. Detection methods of abnormal electricity consumption behaviors: review and prospect// *Autom. Electri. Power Syst.* 42(17), 189– 199 (2018)

32.Chen R.-C., Haung Y.-F., Hsieh C.-F. ‘Ranger intrusion detection system for WSN with sybil attack based on ontology// *New Aspects of Applied Informatics, Biomedical Electronics and Informatics and Communications, WSEAS International Conference on Applied Informatics and Communications*, 2010, pp. 176-180.

- 33.Chen, Z. Sun J. Simulation of mobile network information transmission security defense under big data// *Comput. Simul.* 35(05), 207–210 (2018)
- 34.Chitra S.M., Vinoba V. Padmavathy T.Y. Link Reliability Routing Protocol in WSN using Game Theory Approach// *Proceedings of the International Conference on Applied Mathematics and Theoretical Computer Science*, 2013, pp. 235-238.
- 35.Cobb M. Powering High Speed Analog-to-Digital Converters with Switching Power Supplies// *Power Management Design Line*, Analog Devices, 2009, May 27.
- 36.Deng M. Regulation and protection of personal data in the context of big data// *J. Beijing Univ. Posts Telecommun.* 21(01), 19–25 (2019)
- 37.Deng Z., Cui J., Liang Z. Condition monitoring system of mine hoist based on storage test// *Coal Technol.* 38(05), 179–181 (2019)
- 38.Du S. Application of online monitoring system in urban rail transit based on the big data// *Urban Mass Transit* 21(S2), 30–33 (2018)
- 39.Elrahim A.G.A., Hussein E.A., Salwa R.E., Magdy I.M. An energy aware WSN geographic routing protocol// *Universal Journal of Computer Science and Engineering Technology*, 2010, vol. 1, no. 2, pp. 105-111.
- 40.Enami N., Reza A.M., Kouros D., Hoseini M. Neural network based energy efficiency in WSN: A survey// *International Journal of Computer Science & Engineering Survey*, 2010, vol. 1, no. 1, pp. 39-53.
- 41.Eshaftri M, Al-Dubai AY, Romdhani I, Yassien MB (2015) A new energy efficient cluster based protocol for wireless sensor networks. In: 2015 IEEE federated conference on computer science and information systems (FedCSIS), pp 1209–1214.
- 42.Fan G.J., Jin S.Y. Coverage problem in WSN: A survey// *Journal of networks*, 2010, vol. 5, no. 9, pp. 1033-1040.

43. Felemban E. 'Advanced border intrusion detection and surveillance using WSN technology', *Int. J. Communications, Network and System Sciences*, 2013, no. 6, pp. 251-259.

44. Flouri K., Beferull-Lozano B., Tsakalides P. Energy-Efficient Distributed Support Vector Machines for Wireless Sensor Networks// *Proc. 2006 European Workshop on Wireless Sensor Networks (EWSN '06)*, Zurich, Switzerland, 2006, February 13-15.

45. Fu X, Pace P, Aloï G, Li W, Fortino G (2021) Toward robust and energy-efficient clustering wireless sensor networks: A doublestage scale-free topology evolution model. *Comput Netw* 200:108521.

46. Fu X., Gao Y., Luo B. et al. Security threats to HADOOP: data leakage attacks and investigation// *IEEE Netw.* PP(2), 12–16 (2017)

47. Gao S., Li C. An improved spectral clustering algorithm for anomaly detection of power data// *Comput. Simul.* 36(11), 239–242 + 304 (2019)

48. Gupta P, Kumar PR (1999) Critical power for asymptotic connectivity in wireless networks. *Stochastic analysis, control, optimization and applications*. Springer, Berlin, pp 547–566.

49. Hamzah A, Shurman M, Al-Jarrah O, Taqieddin E (2019) Energy-efficient fuzzy-logic-based clustering technique for hierarchical routing protocols in wireless sensor networks. *Sensors* 19(3):561.

50. Hu P, Zhou Z, Liu Q, Li F (2007) The hmm-based modeling for the energy level prediction in wireless sensor networks. In: *2007 2nd IEEE conference on industrial electronics and applications*, pp 2253–2258.

51. Huang J., Dai B., Zhang L. et al. Study on dynamic identification of abnormal data of electric energy measurement device// *Guangxi Electr. Power* 41(04), 53–55 + 64 (2018)

52. Huang X, Acero A, Hon H-W, Reddy R (2001) A guide to theory, algorithm, and system development, spoken language processing. Prentice Hall PTR, New Jersey.

53.Husein A.E., Ryndin N.A. Development of a probabilistic method for managing data flows based on a hidden Markov model// Modern informatization problems in simulation and social technologies (MIP-2025'SCT): Proc. of the XXX-th Int. Open Science Conf. - Yelm, WA, USA: Science Book Publishing House, 2025. – pp. 54-64.

54.Husein A.E., Ryndin N.A. Stochastic data flow management process to minimize energy consumption// Modern informatization problems in simulation and social technologies (MIP-2026'SCT): Proceedings of the XXXI-th International Open Science Conference (Yelm, WA, USA, January 2026). - Yelm, WA, USA: Science Book Publishing House, 2026. – Pp. 65-75.

55.Issariyakul T, Hossain E (2009) Introduction to network simulator 2 (ns2). Introduction to network simulator NS2. Springer, Berlin, pp 1–18.

56.Kalkha H, Satori H, Satori K (2016) Performance evaluation of AODV and leach routing protocol. Adv Inf Technol Theory Appl 1(1):112–118.

57.Kalkha H, Satori H, Satori K (2017) A dynamic clustering approach for maximizing scalability in wireless sensor network. Trans Mach Learn Artif Intell. <https://doi.org/10.14738/tmlai.54.3328>.

58.Kalkha H, Satori H, Satori K (2019) Preventing black hole attack in wireless sensor network using hmm. Proc Comput Sci 148:552–561

59.Kirankumar B.B., Ajayakumar C.K., Poornima M.C., Gururaj K. Energy Efficient Multipath Routing Protocol with Guaranteed Data Delivery for WSN// International Journal of Information and Education Technology, 2014, vol. 4, no.5, pp. 430-435.

60.Kravets O.Ja., Ali Husein, Getmanskaia D.V., Mustafa Jaafar Baqer Al-Imari, Chernoiivanenko I.A., Redkin Yu.V., Varlamov O.O. Automata model of a system with distributed ledger technology based on a blockchain// International Journal on Information Technologies and Security, vol. 17, no. 4, 2025, pp. 79-86. <https://doi.org/10.59035/VSVB7916>. WOS: 001622781400008.

- 61.Kumar A.S., Kanthan L.N., Priya S.N., Amala G, Revathi A., Revathi R. A Survey on Flow Balanced Routing in WSN// International Journal Of Technology Enhancements And Emerging Engineering Research, 2014, vol. 2, no. 9.
- 62.Kumar S, Tiwari SN, Hegde RM (2015) Sensor node tracking using semi-supervised hidden markov models. *Ad Hoc Netw* 33:55–70.
- 63.Leng X., Chen G., Jiang Y. et al. Data specification and processing in big-data analysis system for monitoring and operation of smart grid// *Autom. Electr. Power Syst.* 42(19), 169– 178 (2018)
- 64.Li J, Liu D (2015) Dpso-based clustering routing algorithm for energy harvesting wireless sensor networks. In: 2015 IEEE international conference on wireless communications and signal processing (WCSP), pp 1–5.
- 65.Liu S, Srivastava R, Koksal CE, Sinha P (2009) Pushback: a hidden markov model based scheme for energy efficient data transmission in sensor networks. *Ad Hoc Netw* 7(5):973–986.
- 66.Liu S., Glowatz M., Zappatore M. et al. *E-Learning, e-Education, and Online Training*. Springer, Heidelberg (2018), 374 p.
- 67.Liu S., Liu D., Srivastava G. et al. Overview and methods of correlation filter algorithms in object tracking// *Complex Intell. Syst.*, 2021. Vol. 7, pp. 1895–1917.
- 68.Liu S., Lu M., Li H. et al. Prediction of gene expression patterns with generalized linear regression model// *Front. Genet.* 10, 120 (2019)
- 69.Liu W., Zong L., Xing C. et al. Design of the overall information collection of wind farm monitoring system based on EDPF-CP system// *Renew. Energy Resour.* 36(08), 1204–1208 (2018)
- 70.Lu M., Liu S. Nucleosome positioning based on generalized relative entropy// *Soft. Comput.* 23(19), pp. 9175–9188 (2018).

71. Malek A-G, Chunlin L, Zhiyong Y, Hasan AHN, Xiaoqing Z (2012) Improved the energy of ad hoc on-demand distance vector routing protocol. *IERI Proc* 2:355–361.

72. Medina Carlos, Segura Jose´ C, de la Torre Angel (2013) Accurate time synchronization of ultrasonic TOF measurements in IEEE 802.15.4 based wireless sensor networks. *Ad Hoc Netw* 11(1):442–452.

73. Mini RAF, Loureiro AAF, Nath B (2004) The distinctive design characteristic of a wireless sensor network: the energy map. *Comput Commun* 27(10):935–945.

74. Mozaffari M, Safarinejadian B, Shasadeghi M (2020) A novel mobile agent-based distributed evidential expectation maximization algorithm for uncertain sensor networks. *Trans Inst Meas Control* 43(7):1609–1619.

75. Nazli Tekin and Vehbi Cagri Gungor (2020) Analysis of compressive sensing and energy harvesting for wireless multimedia sensor networks. *Ad Hoc Netw* 103:102164.

76. Obado V, Djouani K, Hamam Y (2012) Hidden markov model for shortest paths testing to detect a wormhole attack in a localized wireless sensor network. *Proc Comput Sci* 10:1010–1017.

77. Ortega AP, Ramchurn SD, Tran-Thanh L, Merrett GV (2021) Partner selection in self-organised wireless sensor networks for opportunistic energy negotiation: a multi-armed bandit based approach. *Ad Hoc Netw* 112:102354.

78. Park GY, Kim H, Jeong HW, Youn HY (2013) A novel cluster head selection method based on k-means algorithm for energy efficient wireless sensor network. In: 2013 27th IEEE international conference on advanced information networking and applications workshops, pp 910–915.

79. Rabiner LR (1989) A tutorial on hidden markov models and selected applications in speech recognition. *Proc IEEE* 77(2):257–286.

80. Rawat P, Singh KD, Chaouchi H, Bonnin JM (2014) Wireless sensor networks: a survey on recent developments and potential synergies. *J Supercomput* 68(1):1–48.

81. Razzaque MA, Bleakley C, Dobson S (2013) Compression in wireless sensor networks: a survey and comparative evaluation. *ACM Trans Sens Netw (TOSN)* 10(1):1–44.

82. Rohit Kumar and Joy Chandra Mukherjee (2021) On-demand vehicle-assisted charging in wireless rechargeable sensor networks. *Ad Hoc Netw* 112:102389.

83. Russel A, Moundounga A, Satori H, Satori K (2020). An overview of routing techniques in wsns. In: 2020 IEEE fourth international conference on intelligent computing in data sciences (ICDS), pp 1–7.

84. Saihi M, Boussaid B, Zouinkhi A, Abdelkrim N (2015) Distributed fault detection based on hmm for wireless sensor networks. In: 2015 IEEE 4th international conference on systems and control (ICSC), pp 189–193.

85. Shiny A.V.A., Nagarajan V. Energy Efficient Routing Protocol for Mobile WSN// *International Journal of Computer Applications*, 2012, vol. 43, no. 2, pp. 1-5.

86. Singh DP, Bhateja V, Soni SK (2014) Prolonging the lifetime of wireless sensor networks using prediction based data reduction scheme. In: 2014 IEEE international conference on signal processing and integrated networks (SPIN), pp 420–425.

87. Sohraby K, Minoli D, Znati T (2007) *Wireless sensor networks: technology, protocols, and applications*. John Wiley and sons, New Jersey.

88. Song F, Zhu M, Zhou Y, You I, Zhang H (2019) Smart collaborative tracking for ubiquitous power IoT in edge-cloud interplay domain. *IEEE Int Things J* 7(7):6046–6055.

89. Stahlbuhk T, Shrader B, Modiano E (2019) Learning algorithms for scheduling in wireless networks with unknown channel statistics. *Ad Hoc Netw* 85:131–144.

90. Swain RR, Dash T, Khilar PM (2019) A complete diagnosis of faulty sensor modules in a wireless sensor network. *Ad Hoc Netw* 93:101924.

91. Tabatabaei S (2020) A novel fault tolerance energy-aware clustering method via social spider optimization (sso) and fuzzy logic and mobile sink in wireless sensor networks (wsns). *Comput Syst Sci Eng* 35(6):477–494.

92. Tong X., Yu S. Fault detection algorithm for transmission lines based on random matrix spectrum analysis// *Autom. Electr. Power Syst.* 43(10), 101–115 (2019)

93. Vouma Lekoundji J-B (2014) *Mode`les de Markov cache´s*, PhD thesis, Universite´ du Que´bec a` Montre´al.

94. Wang C-F, Shih J-D, Pan B-H, Tin-Yu W (2014) A network lifetime enhancement method for sink relocation and its analysis in wireless sensor networks. *IEEE Sens J* 14(6):1932–1943.

95. Wang Q, Hassanein H, Takahara G (2004) Stochastic modeling of distributed, dynamic, randomized clustering protocols for wireless sensor networks. In: *Workshops on mobile and wireless networking/high performance scientific, engineering computing/network design and architecture/optical networks control and management. Ad Hoc and Sensor Networks/Compil*, pp 456–463.

96. Xiaofei X, Zhang Z, Chen Y, Li L (2020) Hmm-based predictive model for enhancing data quality in WSN. *Int J Comput Appl* 42(4):351–359.

97. Xu G., Ning B., Zhong Y. Automatic matching of voltage blackout events in metering automation system// *Electron. Test* 04, 111–112 (2019)

98. Yang J., Zeng X., Yao L. et al. Research on abnormal electricity monitoring based on large data mining// *Autom. Instrum.* 08, 219–222 (2019)

99. Yang S, Ma L, Jia S, Qin D (2019) A novel markov model-based low-power and secure multihop routing mechanism. *J Sens* 2019(2019):1–11.

100. Yang X., Qu Y., Pang H. et al. Power metering pipeline fault warning technology based on deep learning algorithm// *Electron. Des. Eng.* 28(04), 153–157 (2020)
101. Yen JY (1971) Finding the k shortest loopless paths in a network. *Manag Sci* 17(11):712–716.
102. Zhang L., Shao F. Design of abnormal risk monitoring system for network big data platform// *Mod. Electron. Tech.* 41(22), 143–146 (2018)
103. Zhang Y., Liu K., Yang L. et al. Platform construction and data processing application technology in coal industry monitoring big data// *Coal Sci. Technol.* 47(03), 75–80 (2019)
104. Zheng P., Shuai L., Arun S., Khan M. Visual attention feature (VAF): a novel strategy for visual tracking based on cloud platform in intelligent surveillance systems// *J. Parallel Distrib. Comput.* 120, 182–194 (2018)
105. Zurita G, Shukla AK, Pino JA, Merigo' JM, Lobos-Ossando'n V, Muhuri PK (2020) A bibliometric overview of the journal of network and computer applications between 1997 and 2019. *J Netw Comput Appl* 165:102695.