Федеральное государственное автономное образовательное учреждение
высшего образования
«**НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ**»

На правах рукописи

ХУДОНОГОВА ЛЮДМИЛА ИГОРЕВНА

КОМПЛЕКСИРОВАНИЕ ИНТЕРВАЛЬНЫХ ИЗМЕРИТЕЛЬНЫХ ДАННЫХ МЕТОДОМ АГРЕГИРОВАНИЯ ПРЕДПОЧТЕНИЙ

Специальность 05.11.13 – Приборы и методы контроля природной среды, веществ, материалов и изделий

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель – доктор технических наук,
профессор Муравьев Сергей Васильевич

Томск – 2017

Содержание

Стр.

ВВЕДЕНИЕ

Актуальность темы. Описание результатов измерений в форме *интервалов*, границы которых определяются найденными экспериментально или заранее заданными значениями неопределенности, широко используется как в теории, так и в практике измерений. Интервальные данные являются распространенной формой данных в таких областях, как распределенные вычисления, базы данных, системы и сети сбора данных и т.д.

Одним из подходов к обработке интервальных данных является *комплексирование данных* (data fusion) – процесс совместной обработки данных о некотором объекте, предоставленных несколькими источниками, с целью получения более полного, объективного и точного знания исследуемой характеристики объекта по сравнению со знанием, полученным из единственного источника.

Процедура комплексирования интервальных данных заключается в формировании такого *результирующего интервала* $[x^* - \varepsilon^*, x^* + \varepsilon^*]$, который согласован (т.е. пересекается) с максимальным количеством исходных интервалов $\{I_k\}$ (не обязательно согласованных между собой) и с максимальной степенью правдоподобия содержит значение, которое может служить представителем всех этих интервалов. *Результатом комплексирования* x^* является средняя точка результирующего интервала с соответствующей *неопределенностью* ε^*.

Существуют различные подходы к комплексированию интервальных данных, среди которых можно выделить методы, основанные на математической статистике и теории вероятностей; теории очевидностей Демпстера-Шафера; одобрительном голосовании; интервальных порядковых числах. Некоторые из этих методов являются чувствительными к несогласованности и/или виду закона распределения входных данных. Недостатком других методов является неединственность получаемых результатов. Кроме того, некоторые подходы требуют для нахождения результата комплексирования x^* дополнительной

входной информации субъективного характера, например, назначения весовых коэффициентов источникам данных.

В связи с этим существует необходимость разработки метода комплексирования интервальных данных, позволяющего на основании неточных, неполных или противоречивых данных определить результат x^* с повышенной точностью, робастностью и достоверностью.

Эти полезные свойства результата комплексирования обеспечивает метод агрегирования предпочтений, основанный на представлении исходных интервалов $\{I_k\}$ на вещественной числовой оси отношениями слабого порядка (ранжированиями) на множестве принадлежащих этим интервалам дискретных значений. Результатом комплексирования x^* служит наилучшее значение в ранжировании консенсуса, найденном для набора ранжирований дискретных значений, соответствующих исходным интервалам.

Острая необходимость в таких методах существует, в частности, в практической области беспроводных сенсорных сетей. *Беспроводная сенсорная сеть* (БСС) представляет собой распределенную, самоорганизующуюся систему сбора, обработки и передачи информации, состоящую из автономных, не требующих специальной установки и обслуживания, устройств. Каждое такое устройство, называемое *узлом*, снабжено *мультисенсором* – набором сенсоров, которые измеряют параметры различных физических полей, сред и объектов в подлежащих мониторингу точках исследуемой области. Комплексирование интервальных измерительных данных мультисенсоров, проведенное методом агрегирования предпочтений может обеспечить повышение точности результатов измерений мультисенсоров и продление их времени жизни.

Целью диссертационной работы является разработка и экспериментальные исследования метода комплексирования интервальных измерительных данных на основе агрегирования предпочтений, устойчивого к виду закона распределения входных данных и обеспечивающего получение значения измеряемой величины с повышенной точностью и достоверностью.

В связи с поставленной целью в работе должны быть решены следующие **задачи**:

- анализ известных методов комплексирования интервальных данных;
- разработка и программная реализация метода комплексирования интервальных данных на основе агрегирования предпочтений принадлежащих этим интервалам дискретных значений и экспериментальные исследования его работоспособности и свойств;
- разработка способа разбиения диапазона актуальных значений, полученного в результате объединения исходных интервалов, для формирования ранжируемых дискретных значений;
- разработка и верификация процедур повышения точности мультисенсоров и снижения энергопотребления узлов в беспроводной сенсорной сети на основе предложенного метода комплексирования интервальных данных.

Методы исследования. Использованы методы теории голосования, теории измерений, теории погрешностей, а также теории вероятностей и математической статистики. Численные экспериментальные исследования проводились с использованием метода Монте-Карло для генерации синтетических измерительных данных с помощью специально разработанного программного обеспечения в среде NI LabVIEW.

Достоверность полученных результатов диссертационной работы подтверждается сравнением свойств разработанных и известных алгоритмов на достаточном объеме исходных данных; совпадением с достаточной точностью аналитических расчетов и результатов численных экспериментов.

Научная новизна

1. Предложен и исследован метод комплексирования интервалов IF&PA, где результатом комплексирования является наилучшее дискретное значение в ранжировании консенсуса, найденном для набора наведенных интервалами ранжирований дискретных значений.
2. Для формирования ранжируемых дискретных значений предложен и экспериментально обоснован способ расчета мощности разбиения диапазона

актуальных значений, полученного в результате объединения исходных интервалов, на основе поправки Шеппарда для дисперсии дискретизированных данных.

3. На основе разработанного метода комплексирования интервалов IF&PA предложен и исследован робастный алгоритм повышения точности результата измерения, где исходные интервальные данные представляют собой неточные и/или неполные показания мультисенсоров беспроводной сенсорной сети.
4. На основе разработанного метода комплексирования интервалов IF&PA предложен и исследован алгоритм выбора подмножества активных узлов в кластере беспроводной сенсорной сети, обеспечивающий снижение энергопотребления (продление времени жизни) узлов.

Практическая ценность работы. Результаты диссертационной работы могут быть использованы для обработки интервальных данных во всех типах систем, где подобные данные имеют место: системы распределенных вычислений, базы данных, сети сбора данных и т.п. Типичными практическими применениями метода IF&PA могут быть: межлабораторые или ключевые сличения; прогнозирование значений фундаментальных констант; проведение сертификационных испытаний (на соответствие); повышение точности сенсоров и выявление отказов сенсорных узлов в беспроводных сенсорных сетях.

Реализация и внедрение результатов работы. Результаты исследований использованы при выполнении следующих НИР:

- грант РНФ 14-19-00926 "Основанный на полимерных оптодах мобильный цветометрический экспресс-анализ природных и техногенных объектов на содержание опасных веществ", 2014-2016 гг.;
- проект № 2.5760.2017/БЧ "Методы повышения точности промышленных робототехнических комплексов" в рамках базовой части государственного задания "Наука" Минобрнауки России, 2017-2019 гг.

Результаты работы также используются: в лаборатории мониторинга окружающей среды Томского государственного университета для обработки данных экологического мониторинга; в учебном процессе на кафедре систем управления и мехатроники Института кибернетики ТПУ. Акты внедрения приложены к диссертационной работе (приложение А).

Положения, выносимые на защиту:

1. Предложенный метод комплексирования интервальных данных на основе агрегирования предпочтений гарантирует получение результата с более высокими точностью, робастностью и достоверностью по сравнению с известными методами.
2. Предложенный способ расчета мощности *n* разбиения диапазона актуальных значений позволяет определить такое значение *n*, при котором с вероятностью 0,95 обеспечивается получение результата комплексирования, наиболее близкого к номинальному значению для всех *n* от 4 до 15.
3. Разработанный на основе метода IF&PA робастный алгоритм повышения точности позволяет снизить неопределенность результата измерения не менее чем в 2-2,3 раза по сравнению с неопределенностью показаний мультисенсоров беспроводной сенсорной сети при возможном непустом подмножестве неисправных сенсоров.
4. Разработанный на основе метода IF&PA алгоритм выбора активного подмножества узлов в кластере сети обеспечивает снижение энергопотребления (продление времени жизни) узлов в кластере в 2-3 раза.

Апробация результатов работы. Основные результаты диссертационной работы докладывались на следующих конференциях: 2nd International Symposium on Computer, Communication, Control and Automation (3CA 2013), Singapore, 2013; XIX и XXI Международная научно-практическая конференция студентов, аспирантов и молодых ученых "Современные техника и технологии", г. Томск, 2013 и 2015 гг.; XII и XIV Всероссийская научно-практическая конференция "Молодежь и современные информационные технологии", г. Томск, 2014 и 2016 гг.; 7th International Congress on Ultra Modern Telecommunications and Control

Systems (ICUMT 2015), Brno, Czech Republic, 2015; XVI Международная научно-техническая конференция "Измерение, контроль, информатизация 2015", г. Барнаул, 2015 г.; VI Научно-практическая конференция с международным участием "Информационно-измерительная техника и технологии", Томск, 2015 г.; XI и XII Международная IEEE Сибирская конференция по управлению и связи (SIBCON), Омск, 2015 г., и Москва, 2016 г.; IV Всероссийский молодежный Форум с международным участием "Инженерия для освоения космоса", Томск, 2016 г.; Joint IMEKO TC1-TC7-TC13 Symposium "Metrology Across the Sciences: Wishful Thinking?", Berkeley, USA, 2016.

Публикации. Основные результаты исследований отражены в 17 публикациях: 3 статьи в ведущих научных журналах и изданиях, рекомендуемых ВАК, в том числе 2 проиндексированы в базах данных Web of Science (WoS) и Scopus; 12 статей в рецензируемых научных журналах и сборниках трудов международных и российских конференций, в том числе 4 проиндексированы в базе данных Scopus и WoS; 2 свидетельства о государственной регистрации программ для ЭВМ.

Структура и объем диссертации. Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы из 121 наименования и приложений. Работа содержит 142 страницы основного текста, включая 34 рисунка и 36 таблиц.

В первой главе представлен анализ отечественных и зарубежных источников по тематике, рассмотрены задачи и классификация методов комплексирования, введено понятие интервальных данных и исследованы существующие методы комплексирования интервальных данных.

Во второй главе рассмотрена задача агрегирования предпочтений и представлен алгоритм ее решения (нахождения ранжирования консенсуса) с помощью правила Кемени. Исследованы особенности агрегирования предпочтений на интервальных данных и введено понятие инранжирования для обозначения ранжирований, наведенных интервалами. Представлен метод комплексирования интервальных данных агрегированием предпочтений.

Приведено описание программного обеспечения, разработанного для проведения численных экспериментальных исследований качества работы предложенного метода, и представлены результаты этих исследований.

В третьей главе исследована проблема нахождения подходящего разбиения диапазона актуальных значений (ДАЗ) для метода IF&PA. Показаны свойства этого разбиения и представлен способ определения мощности разбиения, основанный на поправке Шеппарда для дисперсии дискретизированных значений ДАЗ.

В четвертой главе рассмотрена возможность повышения точности мультисенсоров беспроводной сенсорной сети (БСС) для экологического мониторинга на основе метода комплексирования интервальных данных IF&PA. Предложен робастный алгоритм повышения точности результата измерения мультисенсоров БСС и представлены результаты верификации предложенного алгоритма на входных синтетических данных. Приведены результаты обработки данных реальных БСС алгоритмом повышения точности. Предложен алгоритм выбора подмножества активных узлов в кластере БСС для снижения энергопотребления и проведены численные экспериментальные исследования его работоспособности.

ГЛАВА 1
МЕТОДЫ КОМПЛЕКСИРОВАНИЯ ИНТЕРВАЛЬНЫХ ДАННЫХ

Для получения более точного результата измерений на основании неточных, неполных или противоречивых данных, предоставленных различными источниками, используют методы комплексирования данных. Полученное в результате комплексирования значение обладает большей надежностью по сравнению с отдельными значениями индивидуальных источников, в чем выражается синергетический эффект комплексирования.

В этой главе представлен анализ отечественных и зарубежных источников по тематике, рассмотрены задачи и классификация методов комплексирования, введено понятие интервальных данных и исследованы существующие методы комплексирования интервальных данных.

1.1 Комплексирование данных

Под ***комплексированием данных*** (в англоязычных источниках – ***data fusion***) понимается процесс совместной обработки данных о некотором объекте, предоставленных несколькими источниками, с целью получения более полного, объективного и точного знания исследуемой характеристики объекта, чем знание, полученное из единственного источника. В широком смысле целью комплексирования данных является улучшение качества информации, т.е. получение наиболее полного, краткого и согласованного представления данных [24].

Простым примером полезного эффекта комплексирования данных может служить результат классической обработки повторных измерений, которая появилась задолго до введения сравнительно нового термина «комплексирование данных». Действительно, с точки зрения математической статистики, результат x одного измерения характеризуется стандартным отклонением S, в то время как оценка математического ожидания $\overline{x}$ в виде среднего арифметического n нормально распределенных результатов измерений характеризуется стандартным

отклонением $S(\overline{x}) = S / \sqrt{n}$. В этом и заключается единственный, но полезный эффект от проведения многократных измерений вместо одного.

Получение точного и достоверного описания объекта посредством объединения данных из разных источников является непростой задачей и требует применения специально разработанных ***методов комплексирования***. Такие методы условно можно разделить на приведенные ниже три группы в зависимости от того, какие исходные данные предоставляются источниками информации [40].

1. ***Взаимодополняющее (комплементарное) комплексирование:*** информация от нескольких источников представляет собой взаимодополняющие фрагменты некоторой ситуации и используется для получения более полной информации об этой ситуации. Так, на рисунке 1.1 источники $И_1$ и $И_2$ предоставляют различные фрагменты информации a и b, которые затем комплексируются в полное описание ($a \div b$) [72]. Примером может служить сеть камер, в которой информация о некотором пространстве предоставляется несколькими камерами с различными полями обзора [47].

2. ***Избыточное комплексирование:*** два или более источников предоставляют данные об одной и той же характеристике объекта, которые комплексируются для получения более надежной оценки этой характеристики. Как показано на рисунке 1.1, источники $И_2$ и $И_3$ передают одну и ту же информацию b, а комплексирование позволяет достичь более точного представления этой информации (b). Такие методы применяется для повышения надежности, точности и достоверности данных, и часто используется для обеспечения отказоустойчивости и робастности системы. К данной группе можно отнести, например, снижение уровня шума за счет комплексирования изображений, поступающих с различных камер с частично совпадающим полем обзора.

3. ***Кооперативное комплексирование:*** полученная информация от разных источников объединяется в новую информацию, обычно более сложную, чем исходная. Источники $И_4$ и $И_5$ на рисунке 1.1 предоставляют различную

информацию c_1 и c_2, которая в результате комплексирования становится качественно другой, лучше описывающей объект, чем c_1 и c_2 по отдельности. Типичным примером является определение местоположения объекта на основании данных об угле отклонения и расстоянии от заданной точки [16].

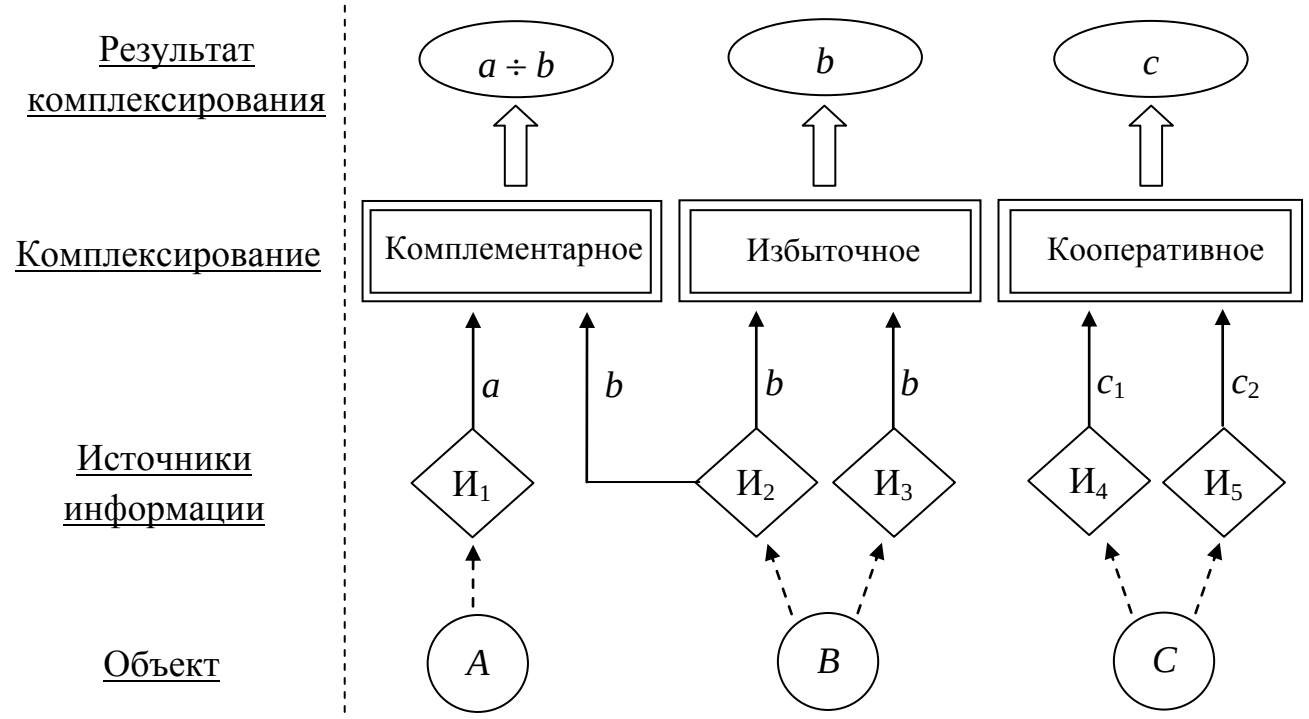


Рисунок 1.1 – Классификация методов комплексирования в зависимости от предоставляемой информации

В список методов комплексирования данных обычно включают методы математической статистики и теории вероятностей [91], теории нечетких множеств [17], теории очевидностей Демпстера-Шафера [60], байесовский подход [99], фильтр Калмана [50], различные методы из области искусственного интеллекта [46], методы голосования (агрегирования предпочтений) [73, 87].

В настоящее время комплексирование данных находит широкое применение во многих областях: робототехника [19], дистанционные измерения [26], управление системой движения [95], отслеживание и распознавание объектов [62], военная деятельность [32], инженерия [42], аэрокосмические системы [57], медицина [22], метрология [53] и др.

1.2 Модели комплексирования данных

Модели, служащие основой при разработке технических требований, структуры и предложений по использованию систем комплексирования данных, определяют основные стадии процесса комплексирования в зависимости от уровня абстракции данных [16]. Рассмотрим две наиболее широко используемые модели.

1.2.1 JDL модель

Данная модель считается самой известной в области комплексирования данных. Она была предложена исследовательской группой «Joint Directors of Laboratories» (JDL) совместно с Министерством обороны США [112]. Модель состоит из пяти уровней обработки информации от источников, базы данных и шины данных, соединяющей все компоненты, как показано на рисунке 1.2. Модель включает следующие компоненты:

- источники: предоставляют входные данные для обработки. Могут быть сенсорами, базами данных, априорными знаниями или мнениями людей.
- система управления базой данных: обеспечивает хранение данных. Эта функция является крайне важной, поскольку комплексирование предполагает использование большого объема разнородных данных.
- интерфейс пользователя: обеспечивает ввод команд и запросов человеком, а также вывод уведомлений о результатах комплексирования с использованием визуальной и звуковой информации.

JDL модель определяет следующие уровни обработки данных:

- уровень 0 (предварительная обработка источников): уменьшает объем данных для последующих уровней обработки за счет распределения данных между подходящими процессами и выбора подходящих источников.
- уровень 1 (оценка объекта): преобразует данные в согласованную структуру. Типичные задачи этого уровня включают идентификацию и локализацию объекта. Примером может служить отслеживание цели: разнородные данные

(изображения, величины углов, акустические сигналы) преобразуются для получения информации о местонахождении цели.

- уровень 2 (оценка ситуации): обеспечивает контекстное описание отношений между объектами и исследуемым событием. Основной целью обычно является определение значимых объектов и событий. На данном уровне для оценки ситуации используется априорная информация и сведения об окружающей среде.
- уровень 3 (оценка угрозы): оценивает текущую ситуацию и на ее основе прогнозирует возможное возникновение угроз, уязвимостей и возможностей.
- уровень 4 (оценка процесса): следит за работой системы, управляет порядком выполнения заданий и распределяет ресурсы в зависимости от заданных целей.

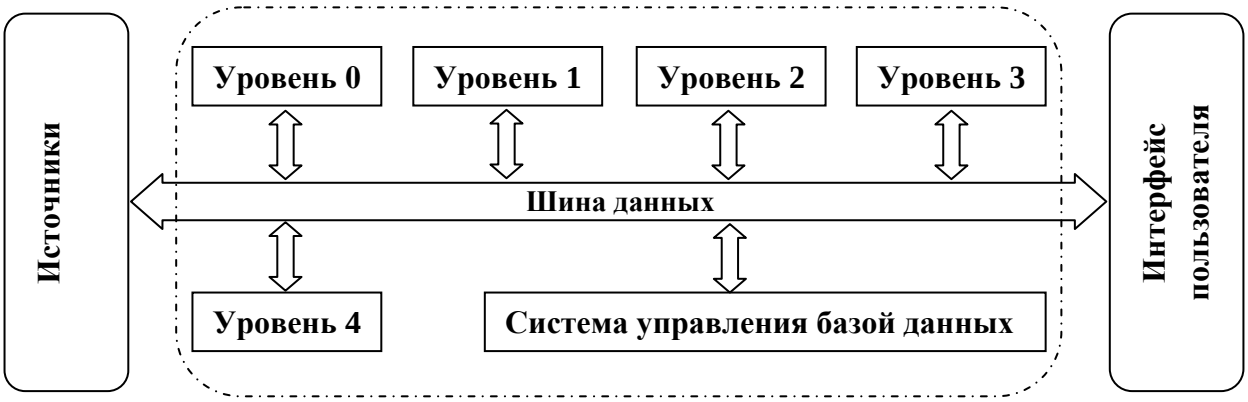


Рисунок 1.2 – JDL модель

В таблице 1.1 представлена декомпозиция уровней JDL модели с примерами соответствующих каждому уровню целей и методов комплексирования [45].

Таблица 1.1 – Пример декомпозиции уровней JDL модели с соответствующими целями и методами комплексирования

Уровень	Цель	Метод
1.Оценка объекта	Приведение данных к стандартному виду	• Преобразование координат • Согласование единиц
	Взаимосвязь объекта и данных	• Байесовский подход • Метод ближайших соседей

Продолжение таблицы 1.1

	Оценка позиционных и кинематических свойств объекта	• Фильтр Калмана • Метод максимального правдоподобия • Гибридные методы
	Распознавание объекта	• Алгоритмы кластеризации • Распознавание образов
2.Оценка ситуации	• Агрегирование объектов • Интерпретация событий/действий • Контекстная интерпретация	• Нечеткая логика • Экспертные системы • Нейронные сети
3. Оценка угрозы	• Предсказание намерений • Многомерный прогноз	• Нейронные сети • Голосование
4.Оценка процесса	• Оценка производительности • Управление процессом	• Теория полезности • Линейное программирование • Системы, основанные на знаниях

1.2.2 DFD модель

Модель, предложенная Belur V. Dasarathy [33], определяет элементы процесса комплексирования, основываясь на входах и выходах системы. DFD модель представлена на рисунке 1.3. Входными данными могут быть необработанные данные, а выходными – некоторое решение. Модель определяет пять категорий комплексирования:

- данные на входе – данные на выходе (Д-Д): комплексирование необработанных данных, результатом которого являются такие же данные, обычно более точные или надежные.
- данные на входе – свойство на выходе (Д-С): данные от источников используются для извлечения свойств, описывающих объект или ситуацию.
- свойство на входе – свойство на выходе (С-С): ряд свойств улучшаются или уточняются либо используется для извлечения новых свойств.

- свойство на входе – решение на выходе (С-Р): ряд свойств объекта или ситуации используется для получения символьного представления или решения.
- решение на входе – решение на выходе (Р-Р): ряд решений комплексируется для получения новых или улучшенных решений.

DFD модель использует трехуровневую иерархическую структуру, в которой одному уровню соответствует входная и выходная информация одного и того же класса (данные, свойство или решение). Так, категория Д-Д осуществляется на низком уровне, а комплексирование категорий С-С и Р-Р выполняются на среднем и высоком уровнях соответственно. Межуровневое комплексирование реализуется при переходе от одного класса информации к другому: от данных к свойству (Д-С) и от свойства к решению (С-Р) [58].

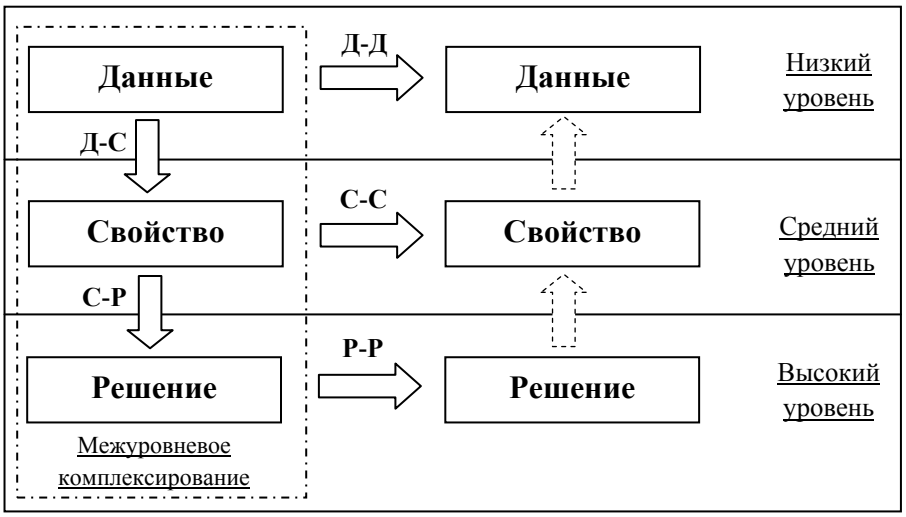


Рисунок 1.3 – DFD модель

Таким образом, JDL модель рассматривает комплексирование с точки зрения системного подхода и дает целостное представление о структуре системы, в то время как DFD модель обеспечивает детальное определение задач комплексирования с помощью ожидаемых входных и выходных данных и способствует установлению связей между этими задачами и данными.

В следующем параграфе рассмотрены основные проблемы, возникающие при комплексировании данных.

1.3 Проблемы комплексирования данных

Реализация системы комплексирования данных сталкивается с необходимостью преодоления ряда трудностей, основные из которых описаны ниже.

Проблема неточных данных. Данные, предоставляемые различными источниками, зачастую обладают низкой точностью вследствие влияния процедуры получения данных, собственных свойств источников, внешних факторов или природы самих данных. В таких случаях возникающая при оценке данных неопределенность приводит к получению заведомого неверного результата.

Проблема противоречивых данных возникает, когда несколько источников предоставляют данные, не согласующиеся между собой, причем степень несогласованности может варьироваться от низкой («выброс», частичное противоречие, затрагивающее единственный источник) до крайне высокой (несогласованность большинства представленных данных). Подобный конфликт может являться следствием внутренних свойств источника данных (например, субъективность эксперта или погрешность сенсора) или же влияния внешних факторов (например, условий окружающей среды).

Проблема разнородных данных. Данные от источников могут быть представлены в виде качественно различной (разнородной) информации, характеризующей один и тот же объект исследования. Например, измерение таких параметров как температура воздуха, концентрация углекислого газа и задымленность позволяет зарегистрировать возникновение пожара в помещении. Оценка осложняется тем, что помимо количественных значений, параметры могут быть выражены также и качественными показателями (например, цвет или форма объекта).

Проблема неполных данных. Ситуация, когда предоставленных данных недостаточно для получения удовлетворительного знания об объекте, возникает нередко и связана прежде всего с невозможностью получения информации от всех источников. Например, при измерениях с помощью беспроводной сенсорной

сети, выход из строя сенсоров, радиомодуля узла или перегрузка радиоканала могут являться причинами того, что часть требуемых данных не будет доставлена на центральный узел.

Проблема большого объема данных. Данные, поставляемые источниками, могут отличаться значительным объемом, что затрудняет их дальнейшую передачу и обработку. Данная проблема особенно актуальна, если речь идет об информации, получаемой с помощью технических средств. Передача больших объемов данных всегда связана с увеличением временных, энергетических и вычислительных ресурсов системы. Процедуры комплексирования должна обеспечивать уменьшение размера данных с сохранением значимой информации.

Несмотря на то, что представленные проблемы интенсивно изучались на протяжении долгого времени [45, 55, 65], их окончательное решение еще далеко от завершения, в связи с чем исследования в этом направлении остаются актуальными.

1.4 Комплексирование интервальных данных

1.4.1 Интервалы

Бесконечное множество действительных чисел, заключенных между двумя точками вещественной числовой оси, будем называть ***интервалом.*** Каждый интервал I характеризуется ***нижней границей*** l, ***верхней границей*** u и ***средней точкой*** x, так что $I = [l, u]$; $l < x$; $x = 0{,}5 \cdot (u + l)$; $l, u, x \in \mathbb{R}$ (рисунок 1.4).

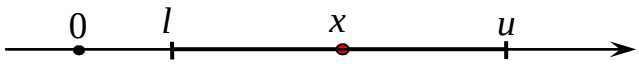


Рисунок 1.4 – Интервал на вещественной числовой оси

Часто средняя точка x представляет собой результат некоторого измерения, интервал неопределенности которого имеет вид $[x - 0{,}5 \cdot (u - l),\ x + 0{,}5 \cdot (u - l)] = [x - \varepsilon,\ x + \varepsilon]$. В этом случае интервал может быть представлен парой $\langle x, \varepsilon \rangle$.

Будем рассматривать набор из m замкнутых интервалов $\{I_k\}$, $k = 1, \ldots, m$, на вещественной числовой оси как исходную информацию для комплексирования. Каждая пара интервалов не обязательно имеет непустое пересечение, т.е. могут найтись два таких интервала I_j и I_k, $j \neq k$; $j, k = 1, \ldots, m$, что $I_j \cap I_k \neq \varnothing$. Такие интервалы будем называть ***несогласованными*.**

1.4.2 Методы комплексирования интервалов

Комплексирование интервальных данных представляет собой отдельную область исследования, которая характеризуется собственными методами и алгоритмами, обусловленными спецификой интервальных данных. В данной работе под ***измерительными интервальными данными*** будем понимать результаты измерений, представленные в форме интервалов на вещественной оси (см. рисунок 1.4). Процедура комплексирования интервальных данных заключается в формировании такого ***результирующего интервала*** $[x^* - \varepsilon^*, x^* + \varepsilon^*]$, который согласован с максимальным количеством исходных интервалов (не обязательно согласованных между собой) и с максимальной степенью правдоподобия содержит значение, которое может служить представителем всех этих интервалов. ***Результатом комплексирования*** x^* является средняя точка результирующего интервала с соответствующей неопределенностью ε^*.

Использование интервальных данных эффективно в следующих случаях:

- измерительная задача имеет неединственное решение, и несколько значений являются одинаково подходящими для представления результата;
- неопределенность, возникающая в процессе измерений, оказывает сильное влияние на результат.

Так, представление результата измерений интервальными данными характерно, в частности, для беспроводных сенсорных сетей [65], межлабораторных сличений [31], прогнозирования значений фундаментальных констант [64], проведения испытаний на соответствие [89] и др. Например, в беспроводной сенсорной сети точность сенсоров существенно зависит от влияния

таких факторов, как условия окружающей среды и заряд модуля питания узла. В такой ситуации нельзя гарантировать, что измеренное значение величины будет являться точным представителем действительного значения. Для учета возможных внешних или внутренних воздействий, измеренному значению x приписываются границы неопределенности ε, и результат измерений представляется интервалом $[x - \varepsilon, x + \varepsilon]$, в котором с заданной вероятностью находится действительное значение измеряемой величины.

Операции, сходные с арифметическими и применимые к интервалам, рассматриваются специальной областью математики – интервальным анализом. Этот подход полезен при исследовании величин, значения которых известны только приблизительно или искажены погрешностями округления [51, 69].

Некоторые из перечисленных в параграфе 1.1 методов комплексирования, используемые, когда результат представлен единственным значением, были адаптированы для работы с интервальными данными. В соответствии с результатами проведенного анализа литературных источников, основные подходы к комплексированию интервальных данных включают в себя методы, основанные на теории очевидностей Демпстера-Шафера [60, 63], теории вероятностей и математической статистике [61, 120], нечеткой логике [100], одобрительном голосовании [18, 21, 28, 87], статистическом подходе определения наименьшего согласованного подмножества интервалов и последующего расчета средневзвешенного значения [29, 41], а также интервальных порядковых числах [39, 44, 110].

1.4.3 Комплексирование интервальных данных с приписанной доверительной вероятностью

В работах [61, 120] представлен вероятностный подход, основанный на комплексировании интервалов с учетом их ***доверительной вероятности***.

Каждый k-й источник из множества $S = \{s_1, s_2, \ldots, s_m\}$ предоставляет информацию о некой величине X в виде закрытого интервала $I_k = [l_k, u_k]$, $l_k \le u_k$ и $k = 1, \ldots, m$, вместе с информацией о приписанной этому интервалу

доверительной вероятности $\alpha_k = P_k(X \in I_k)$. Руководствуясь тем, что значение величины X лежит либо внутри интервала I_k, либо за его пределами, т.е. внутри дополнения интервала $I_k^c = (-\infty, l_k) \cup (u_k, \infty)$, можно сказать, что источник фактически предоставляет два интервала I_k и I_k^c, предполагая принадлежность значения X только одному из них. Соответственно, доверительная вероятность интервала I_k^c выражается как $\alpha_k^c = P(X \in I_k^c) = 1 - \alpha_k$.

Тогда исходными данными для комплексирования является множество интервалов $I = \{I_1^{p_1}, ..., I_m^{p_m}\}$ вместе с множеством соответствующих доверительных вероятностей $\alpha = \{\alpha_1^{p_1}, ..., \alpha_m^{p_m}\}$, где p_k принимает значение 1, если предполагается, что X принадлежит интервалу, и 0 в обратном случае.

На множестве I находят каждое непустое пересечение интервалов, после чего определяют все возможные объединения полученных пересечений:

$$I' = \bigcup_{j=1}^{N} \bigcap_{k=1}^{m} I_k^{p_k(j)}, \tag{1.1}$$

где N – число пересечений интервалов. Таким образом, выходом данного процесса является множество новых интервалов I', мощность которого зависит от значений исходных интервалов множества I. Для установления доверительных вероятностей для интервалов I', найденных по формуле (1.1), используется выражение:

$$\alpha' = P(X \in I) = \frac{1}{c} \prod_{k=1}^{m} \alpha_k^{p_k}, \tag{1.2}$$

где параметр c определяется как

$$c = \sum_{I \neq \varnothing} \prod_{k=1}^{m} \alpha_k^{p_k}. \tag{1.3}$$

После получения множества I' с соответствующими каждому новому интервалу доверительными вероятностями необходимо выбрать единственный подходящий результирующий интервал I_r на основании применения какого-либо оптимизационного критерия.

В работе [120] представлены два возможных критерия:

а) минимизация ширины h интервала при допустимом минимальном значении доверительной вероятности α_{min};

б) максимизация доверительной вероятности α при допустимом максимальном значении ширины интервала h_{max}.

Реализация вероятностного метода может быть показана на следующем примере [120].

Пример 1.1. Пусть три источника предоставляют данные об исследуемой величине в виде интервалов I_1 = [8, 11], I_2 = [9, 12] и I_3 = [10, 13] с приписанными доверительными вероятностями α_1 = 0,80, α_2 = 0,83 и α_3 = 0,85 соответственно. Тогда исходными данными для комплексирования является множество интервалов $I = \{I_1^{p_1}, ..., I_m^{p_m}\}$ и доверительных вероятностей $\alpha = \{\alpha_1^{p_1}, ..., \alpha_m^{p_m}\}$, приведенных в таблице 1.2.

Таблица 1.2 – Исходные интервальные данные с соответствующими доверительными вероятностями

k	***p***	***I***	**α**
1	1	[8, 11]	0,80
	0	$(-\infty, 8) \cup (11, \infty)$	0,20
2	1	[9, 12]	0,83
	0	$(-\infty, 9) \cup (12, \infty)$	0,17
3	1	[10, 13]	0,85
	0	$(-\infty, 10) \cup (13, \infty)$	0,15

В соответствии с формулами (1.1)-(1.3), данные из таблицы 1.2 комплексируются для получения новых интервалов I', каждому из которых соответствует новая доверительная вероятность α', как показано в таблице 1.3.

Таблица 1.3 – Новые интервалы с соответствующими доверительными вероятностями

I′	$(-\infty, 8) \cup (13, \infty)$	[8, 9)	[9, 10)	[10, 11]	(11, 12]	(12, 13]
α′	0,0059	0,0237	0,1159	0,6567	0,1642	0,0336

Результаты выбора единственного результирующего интервала I_r на основании двух представленных выше оптимизационных критериев демонстрируются в таблице 1.4.♦

Таблица 1.4 – Результирующие интервалы с соответствующими доверительными вероятностями

α_{min}	0,6	0,8	0,9
h_{max}	1	2	3
I_r	[10, 11]	[10, 12]	[9, 12]
α_r	0,6567	0,8209	0,9368

Рассмотренный вероятностный метод отличается простотой реализации и быстродействием. Недостаток метода заключается в необходимости получения сведений о доверительной вероятности, приписываемой конкретному интервалу, что на практике не всегда возможно. Авторы также неоднократно подчеркивают, что предложенный метод применим лишь в случае, если исходные интервальные данные получены от независимых источников. Кроме того, авторами не приводится обоснование выбора наиболее подходящего из предложенных оптимизационных критериев.

1.4.4 Теория Демпстера-Шафера для интервальных данных

Работы [60, 63] посвящены методам комплексирования данных, в основе которых лежит теория очевидностей Демпстера-Шафера, адаптированная для применения к интервальным данным.

Пусть X – конечное множество всех рассматриваемых утверждений, тогда $Y = 2^X$ – совокупность всех подмножеств множества X, включая пустое множество. Множество Y состоит из n подмножеств A_i, $i = 1, \ldots, n$. Каждый k-й источник из множества $S = \{s_1, s_2, \ldots, s_m\}$, $k = 1, \ldots, m$, предоставляет интервал значений A_i в форме $I_{ki} = [l_{ki}, u_{ki}]$ при $l_{ki} \le u_{ki}$. На основании полученных интервалов для всех n подмножеств от всех m источников формируется набор I интервалов:

$$I = \begin{cases} I_1 = [I_{11}, I_{12}, ..., I_{1n}] \\ I_2 = [I_{21}, I_{22}, ..., I_{2n}] \\ ... \\ I_m = [I_{m1}, I_{m2}, ..., I_{mn}] \end{cases}, \quad (1.4)$$

где I_k – множество n интервалов, предоставленных k-м источником, и каждый элемент также является интервалом.

Будем называть **м*ассой*** подмножества A такую массу $m(A)$, что для любого A значение $m(A) \geq 0$ и $\sum_{A \subseteq Y} m(A) = 1$, при этом $m(\varnothing) = 0$.

Для вычисления возможности события A на основе приписанных масс используется ***функция доверия*** $bel(A)$, определяемая как сумма всех масс собственных подмножеств A:

$$bel(A) = \sum_{B \subseteq A} m(B). \quad (1.5)$$

Объединение подмножеств A_1, A_2, …, A_k с массой m_1 и B_1, B_2, …, B_i с массой m_2, если их функции доверия bel_1 и bel_2 соответственно, осуществляется в соответствии с ***правилом объединения***:

$$\begin{cases} m(A) = \dfrac{\sum\limits_{A_k \cap B_i = A} m_1(A_k) m_2(B_i)}{\sum\limits_{A_k \cap B_i \neq \varnothing} m_1(A_k) m_2(B_i)}. \\ m(\varnothing) = 0 \end{cases} \quad (1.6)$$

Для интервальных данных правило объединения (1.6) несколько трансформируется, объединяя верхние и нижние границы интервалов соответственно. Правило объединения, представленное в [63], реализуется в три этапа:

1. ***Нормирование.*** Нижние границы интервалов l_{ki} нормируются следующим образом:

$$\hat{l}_{ki} = \frac{l_{ki}}{\sum\limits_{i=1}^{n} l_{ki}}, \quad (1.7)$$

так что $\sum_{i=1}^{n} \hat{l}_{ki} = 1$. Тогда нижние границы интервалов *k*-го источника относительно *n* элементов могут быть записаны в виде:

$$\hat{l}_k = [\hat{l}_{k1}, \hat{l}_{k2}, ..., \hat{l}_{kn}]. \quad (1.8)$$

Верхние границы нормируются аналогичным образом, поэтому здесь и далее будем приводить формулы только для нижних границ. Таким образом, элементы $\hat{l}_k$ и $\hat{u}_k$ являются скалярными значениями и рассматриваются как массы для формулы (1.6).

2. ***Объединение.*** В соответствии с (1.6), нормированные значения границ интервалов объединяются для получения масс $m^l(A_r)$ и $m^u(A_r)$, $r = 1, \ldots, n$.

3. ***Преобразование интервалов.*** Для получения результирующих интервалов $\hat{I}_r$ границы $m^l(A_r)$ и $m^u(A_r)$ преобразуются по формулам (1.9)-(1.11):

$$\hat{l}(A_r) = \min(m^l(A_r), m^u(A_r)), \quad (1.9)$$

$$\hat{u}(A_r) = \max(m^l(A_r), m^u(A_r)), \quad (1.10)$$

$$\hat{I}(A_r) = [\hat{l}(A_r), \hat{u}(A_r)]. \quad (1.11)$$

Выбор единственного результирующего интервала основывается на цели проведения процедуры комплексирования. Так, если целью является идентификация объекта, предлагается использовать следующее правило:

$$\begin{aligned} m(A_1) &= \max\{m(A_j), A_j \subset X\} \\ m(A_2) &= \max\{m(A_j), A_j \subset X, A_j \neq A_1\} \end{aligned}, \quad (1.12)$$

если

$$\begin{cases} m(A_1) - m(A_2) > \varepsilon_1 \\ m(U) < \varepsilon_2 \\ m(A_1) > m(U) \end{cases}, \quad (1.13)$$

где A_1 представляет собой искомый результат, $m(A_j) = \hat{u}(A_j)$ при j = 1, …, n, ε_1 и ε_2 – заданные пороговые значения, а U – подмножество неопределенности множества Y.

Пример 1.2. В работе [63] представлен пример обработки интервальных данных с помощью метода Демпстера-Шафера. В задаче идентификации объекта два источника предоставляют данные о четырех типах (t_1 – t_4) объекта в виде доверительных интервалов I_1 и I_2, приведенных в таблице 1.5.

Таблица 1.5 – Интервальные данные, полученные от двух источников

Интервал	**Тип объекта**			
	t_1	t_2	t_3	t_4
I_1	[0,3; 0,35]	[0,4; 0,45]	[0,1; 0,2]	[0,1;0,2]
I_2	[0,6; 0,8]	[0,25; 0,3]	[0,1; 0,2]	[0,1; 0,2]

В соответствии с формулами (1.9)-(1.11) для объединения полученных интервалов были получены результирующие интервалы, показанные в таблице 1.6.

Таблица 1.6 – Результирующие интервалы

Интервал	**Тип объекта**			
	t_1	t_2	t_3	t_4
$\hat{I}$	[0,53; 0,57]	[0,30; 0,33]	[0,07; 0,13]	[0,03; 0,04]

Выражения (1.12) и (1.13) позволяют при заранее заданных пороговых значениях $\varepsilon_1 = \varepsilon_2 = 0{,}1$ на основании данных таблицы 1.6 определить единственный результирующий интервал [0,53; 0,57] и таким образом установить, что результатом идентификации является тип t_1.♦

К достоинствам метода Демпстера-Шафера можно отнести его работоспособность в случае неполных данных (см. п. 1.3), т.е. когда информация от некоторых источников отсутствует. Главный недостаток метода заключается в том, что правило объединения (1.6) придает особое значение согласованности между многочисленными источниками и игнорирует все конфликтующие данные с помощью нормирования. Правомерность использования правила (1.6)

подвергается серьёзным сомнениям специалистами [107-109] в случае значительных несоответствий между источниками информации.

1.4.5 Одобрительное голосование

Голосование является одним из эффективных методов решения проблемы согласования неточных, неполных или противоречивых данных, полученных из разных источников [88]. Под голосованием понимается процедура измерения уровня согласованности мнений m избирателей (экспертов) относительно ряда n кандидатов (альтернатив). Задачей любой процедуры голосования является нахождение консенсуса, наилучшим образом представляющего обобщенное мнение избирателей. В работах [18, 21, 28, 86-88] представлен подход к комплексированию интервальных данных на основе одобрительного голосования.

Одобрительным голосованием (approval voting) называется система, в которой каждый избиратель выбирает не единственную подходящую альтернативу, а голосует за любое число (одну, несколько или все) предпочтительных для него альтернатив. Все выбранные избирателем альтернативы, т.е. ***одобренное множество***, считаются равнозначными. Подсчет голосов показывает, сколько избирателей «поддержало» каждую альтернативу, и победителем становится альтернатива с наибольшим числом голосов [88]. ***Интервальное голосование*** является частным случаем одобрительного голосования, в котором одобренное множество каждого k-го избирателя представляет собой закрытый интервал $I_k = [l_k, u_k]$ значений на вещественной оси, и все значения из этого интервала (альтернативы) являются одинаково предпочтительными с точки зрения избирателя. Множество всех избирателей вместе с предоставленными ими интервалами I называют ***линейным сообществом*** S [21]. Сообщество S считается ***согласованным,*** если при наличии по крайней мере трех избирателей ($m \geq 3$) существует такая пара избирателей j и k, чьи интервалы пересекаются, в результате образовывая интервал пересечения I', т.е. $I_j \cap I_k = I'$.

Для любого S ***уровень согласованности*** интервала пересечения $q(I')$ отражает количество избирателей, одобривших интервал I'. Очевидно, что уровень согласованности сообщества $q(S)$ определяется максимальным значением $q(I')$ среди всех интервалов пересечения и совпадает с уровнем согласованности результирующего интервала $I_г$:

$$q(S) = \max_{I' \in R} q(I') = q(I_г). \qquad (1.14)$$

В большинстве случаев для определения результирующего интервала $I_г$, представляющего собой наилучший компромисс представленных мнений, применяется одно из правил голосования, в частности, ***правило относительного большинства*** или ***абсолютного большинства.***

Согласно правилу относительного большинства, подсчитывается число избирателей, проголосовавших за каждую из альтернатив, и победителем становится альтернатива, получившая наибольшее количество голосов. В отличие от алгоритмов, где мнение избирателя представляется единственной альтернативой, интервальное голосование не требует прямого подсчета или попарного сравнения всех предпочитаемых альтернатив, за счет чего его временная сложность значительно снижается. Результирующий интервал находится как ***пересечение*** наибольшего количества исходных интервалов. Таким образом, результатом голосования становится интервал $I_г$, характеризующийся максимальным уровнем согласованности $q(I_г)$. При необходимости получения на выходе процедуры комплексирования единственного числового значения за результат принимается средняя точка результирующего интервала.

На рисунке 1.5 приведен пример интервального голосования с применением правила относительного большинства. Каждый k-ый избиратель предоставляет исходный интервал I_k, $k = 1, \ldots, 5$, содержащий предпочтительное множество альтернатив. На основании полученного набора интервалов определяется результирующий интервал $I_г$ с максимальным уровнем согласованности.

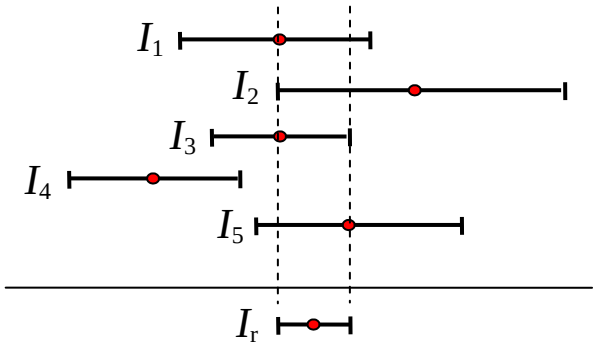


Рисунок 1.5 – Интервальное голосование с использованием правила относительного большинства (уровень согласованности 4)

В соответствии с правилом абсолютного большинства, победителем считается альтернатива, собравшая больше половины голосов. Результирующий интервал определяется посредством последовательного исключения всех интервалов пересечения, уровень согласованности которых составляет менее 50 % от числа избирателей.

Так, для интервалов, представленных на рисунке 1.6, начиная с крайней левой границы и двигаясь направо, увеличиваем значение счетчика (исходное значение которого равно нулю) каждый раз, когда пересекаем нижнюю границу какого-либо исходного интервала, и уменьшаем значение счетчика в случае пересечения верхней границы. Когда значение счетчика достигает числа $\lceil m/2 \rceil$ (3 голоса для 5 исходных интервалов для рисунка 1.6), это означает, что найдена нижняя граница результирующего интервала I_r. Аналогичным образом, начиная отсчет с правого края, определяем и верхнюю границу [86].

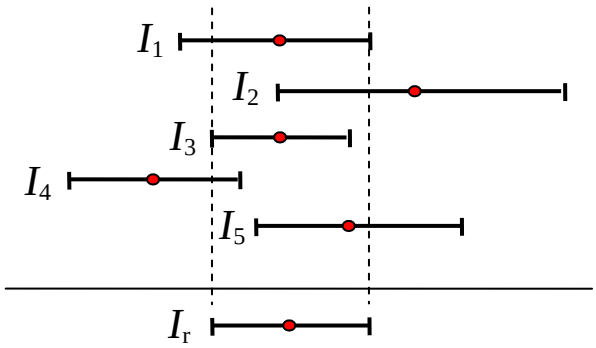


Рисунок 1.6 – Интервальное голосование с использованием правила абсолютного большинства (уровень согласованности 3)

К достоинствам одобрительного голосования можно отнести простоту и невысокую временную сложность алгоритмов. Однако нахождение результирующего интервала с использованием правил относительного и абсолютного большинства может сопровождаться появлением ***парадоксов***. Предположим, что при голосовании получены интервалы от шести избирателей, как показано на рисунке 1.7. Применяя правило относительного большинства, можно найти два непересекающихся между собой интервала $I_{г1}$ и $I_{г2}$, каждый из которых характеризуется одним и тем же уровнем согласованности. В таком случае неясно, какой интервал пересечения следует выбрать в качестве результирующего, поскольку нет причин считать один из них предпочтительнее другого. В свою очередь, при использовании правила абсолютного большинства результирующий интервал часто оказывается достаточно широким, зачастую даже шире исходных интервалов, что приводит к получению результата с бóльшей неопределенностью.

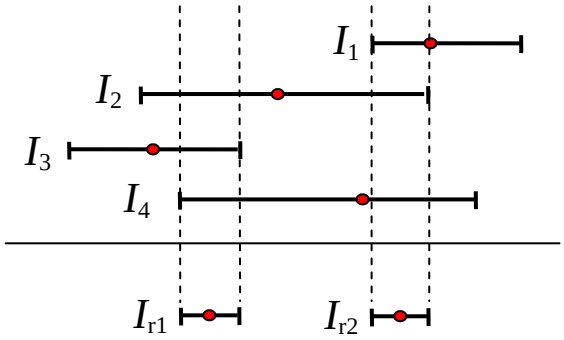


Рисунок 1.7 – Парадокс при использовании правила относительного большинства

1.4.6 Интервальные порядковые числа

В работах [39, 44] предложен подход, который рассматривает задачу ***агрегирования предпочтений*** на множестве альтернатив $X = \{x_1, x_2, \ldots, x_n\}$. Под задачей агрегирования предпочтений [43, 78] понимают нахождение для m ранжирований из n альтернатив ***ранжирования консенсуса***, определяющего порядок альтернатив с учетом их порядка в каждом из m исходных ранжирований [6]. Заметим, что данный подход не относится к методам комплексирования

интервальных данных, но представляет собой заслуживающий внимания метод обработки интервалов в задаче агрегирования предпочтений. Более подробно процедура агрегирования предпочтений и ее особенности рассматриваются в главе 2.

Значение $\overline{r_i} \in \mathrm{N}$, где N – множество всех натуральных чисел, представляет ранг альтернативы x_i в ряду упорядоченных по предпочтению альтернатив. Вследствие неполноты и неопределенности данных, информация о предпочтениях множества источников данных $S = \{s_1, s_2, \ldots, s_m\}$ представляется в виде так называемых ***интервальных порядковых чисел***. Интервальным порядковым числом (ИПЧ) называется число $\overline{r}$, обозначаемое как $\overline{r} = [r^L, r^U]$, где $r^L, r^U \in \mathrm{N}$, представляющее собой ряд дискретных порядковых значений $r \in \mathrm{N}$, таких что $r^L \le r \le r^U$, Здесь r^L и r^U – нижняя и верхняя граница ИПЧ $\overline{r}$ соответственно. При этом считается, что альтернативе x может быть с одинаковой вероятностью приписан любой ранг из ряда $[r^L, r^L+1, ..., r^U-1, r^U]$. На рисунке 1.8 в качестве примера показаны два ИПЧ $\overline{a} = [2,3]$ и $\overline{b} = [3,5]$. На основании множества ИПЧ $R = \sum_{i=1}^{n} \overline{r_i}$, полученных от каждого из m источников, формируется итоговое ранжирование, где каждой альтернативе соответствует конкретный и единственный ранг. Таким образом, подход рассматривает ИПЧ в качестве ***средства*** для агрегирования предпочтений.

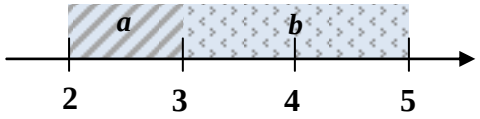


Рисунок 1.8 – Интервальные порядковые числа $\overline{a} = [2,3]$ и $\overline{b} = [3,5]$ на вещественной оси

Для сравнения двух ИПЧ $\overline{a}$ и $\overline{b}$ вводится понятие «степень вероятности». Пусть множество всех порядков пар $Y = \{(i, j) \in [a^L, a^U] \times [b^L, b^U] : i, j \in \mathrm{N}\}$ обладает мощностью $t = |Y|$, а множества $Y_{\overline{a} \succ \overline{b}} = \{(i, j) \in Y : i > j\}$ и $Y_{\overline{a} = \overline{b}} = \{(i, i) \in Y\}$ –

мощностями $d = \left|Y_{\bar{a} \succ \bar{b}}\right|$ и $g = \left|Y_{\bar{a} = \bar{b}}\right|$ соответственно. Тогда степень вероятности того, что $\bar{a}$ предпочтительнее $\bar{b}$, определяется по следующей формуле:

$$P(\bar{a} \succ \bar{b}) = \frac{(d + 0{,}5g)}{t}. \qquad (1.15)$$

На основании рассчитанной степени вероятности и вектора весовых коэффициентов ω = (ω$_1$, ω$_2$, …,ω$_m$), приписываемых источникам, рассчитывается матрица предпочтений $A = (a_{ij})$, при этом

$$a_{ij} = \sum_{k=1}^{m} \omega_k P(\bar{r}_{kj} \succ \bar{r}_{ki}), \qquad (1.16)$$

где ω$_k$ – весовой коэффициент *k*-го источника, $\bar{r}_{kj}$ и $\bar{r}_{ki}$ – ранги, присвоенные *k*-ым источником *i*-ой и *j*-ой альтернативам соответственно, представленные в форме ИПЧ, при *i, j* = 1, …, *n*.

Нахождение итогового ранжирования может основываться, в частности, на свойствах нечеткой логики или теории графов.

Нечеткие отношения предпочтения. С помощью полученной матрицы предпочтений *A* определяется ***вектор предпочтений*** $\upsilon = (\upsilon_1, \upsilon_2, ..., \upsilon_n)$, где υ_i отражает степень предпочтения альтернативы x_i. Бинарное отношение, определенное в *A*, рассматривается как ***нечеткое отношение предпочтения***, в связи с чем вектор υ может быть найден как положительное l_p-решение оптимизационной задачи:

$$\min\{\sum_{i,j=1}^{n} \left|(1 - a_{ij})\upsilon_i - a_{ij}\upsilon_j\right|^p + |\sum_{i=1}^{n} \upsilon_i - 1|^p\}, \qquad (1.17)$$

где *p* – размерность l_p-пространства. При *p* = 1 задача (1.17) представляет собой задачу линейного программирования и решается с помощью симплекс-метода, а в случае *p* = 2 – посредством метода наименьших квадратов [39].

Итоговое ранжирование альтернатив x_i получают сортировкой значений υ_i вектора предпочтений в порядке убывания.

Теория графов. Определение итогового ранжирования основывается на представлении матрицы предпочтений *A* в виде взвешенного ориентированного

графа, где множество вершин является множеством альтернатив X, и a_{ij} выражает вес соответствующей дуги (x_i, x_j). Тогда на базе матрицы A формируется матрица $S = (s_{ij})$, объединяющая результаты попарных сравнений так называемой «силы» альтернатив следующим образом:

$$s_{ij} = \frac{a_{ij}}{\sum_{k=1}^{n} a_{kj}}, \tag{1.18}$$

где $i, j = 1, \ldots, n$. Итоговое ранжирование определяется неотрицательным ***ранговым вектором*** $r = (r_1, r_2, ..., r_n)$, в котором r_i отражает силу альтернативы x_i из матрицы S. Сила альтернативы оценивается из ее взаимодействия с другими альтернативами, принимая во внимания их силы. Тогда ранг альтернативы x_i может быть получен на основе следующего соотношения:

$$r_i = \sum_{j=1}^{n} s_{ij} r_j, \tag{1.19}$$

при $i = 1, \ldots, n$. Ранговый вектор r является неотрицательным собственным вектором и может быть найден с помощью итерационного степенного метода [52].

Результативность предложенного метода может быть продемонстрирована на примере, представленном в работе [44].

Пример 1.3. Множество источников $S = \{s_1, s_2, \ldots, s_5\}$, которым приписаны одинаковые весовые коэффициенты $\omega_k = 0{,}2$ при $k = 1, \ldots, 5$, определяют лучшую альтернативу из ряда $X = \{x_1, x_2, x_3, x_4\}$. Данные от всех источников в виде ИПЧ представлены в таблице 1.7.

Таблица 1.7 – Интервальные порядковые числа, предоставленные источниками

Источник	**Альтернатива**			
	x_1	x_2	x_3	x_4
s_1	[2, 3]	[1, 1]	[2, 4]	[3, 4]
s_2	[3, 4]	[1, 2]	[1, 2]	[3, 4]
s_3	[4, 4]	[2, 3]	[2, 3]	[1, 2]
s_4	[1, 1]	[2, 3]	[4, 4]	[2, 3]
s_5	[2, 3]	[1, 3]	[1, 2]	[4, 4]

В соответствии с (1.16), рассчитывается матрица предпочтений A:

$$A=\begin{pmatrix} 0,5 & 0,2666 & 0,3584 & 0,675 \\ 0,7334 & 0,5 & 0,6666 & 0,725 \\ 0,6416 & 0,3333 & 0,5 & 0,5584 \\ 0,325 & 0,275 & 0,4416 & 0,5 \end{pmatrix}.$$

При определении результата с использованием способа 1 (при p = 2) вектор предпочтений имеет вид:

$$\upsilon = (0,1682; 0,4536; 0,2327; 0,1454),$$

в то время как применение способа 2 дает следующий ранговый вектор:

$$r = (0,2139; 0,3375; 0,2529; 0,1957)^t.$$

Тогда итоговое ранжирование, полученное обоими способами, записывается как $x_2 \succ x_3 \succ x_1 \succ x_4$, и таким образом, наилучшей альтернативой признается x_2. ♦

Представленный метод агрегирования предпочтений на множестве интервальных порядковых чисел характеризуется четкой логикой и хорошим математическим обоснованием способов нахождения итогового ранжирования. С другой стороны, недостатком данного подхода является необходимость назначения весов источникам информации, что всегда представляет собой субъективную процедуру и при неверном распределении весовых коэффициентов может в значительной степени повлиять на результат комплексирования.

Выводы к главе 1

1. Проведен анализ отечественных и зарубежных источников, посвященных комплексированию данных. Приведена классификация методов комплексирования и рассмотрены проблемы, решаемые с их помощью.

2. Рассмотрены основные методы комплексирования интервальных данных, основанные на вероятностном подходе, теории очевидностей Демпстера-Шафера, одобрительном голосовании и интервальных порядковых числах. Анализ показал, что исследованные методы не могут обеспечить одновременно единственность результата комплексирования и его устойчивость к несогласованности входных данных и при этом требуют дополнительной входной информации субъективного характера.

3. Проведенный анализ литературных источников показал необходимость разработки метода комплексирования интервальных данных, позволяющего на основании неточных, неполных или противоречивых данных определить единственный результат комплексирования с повышенной точностью, робастностью и достоверностью.

ГЛАВА 2
КОМПЛЕКСИРОВАНИЕ ИНТЕРВАЛЬНЫХ ДАННЫХ АГРЕГИРОВАНИЕМ ПРЕДПОЧТЕНИЙ

В этой главе рассмотрена задача агрегирования предпочтений и представлен алгоритм ее решения (нахождения ранжирования консенсуса) с помощью правила Кемени. Исследованы особенности агрегирования предпочтений на интервальных данных и введено понятие инранжирования для обозначения ранжирований, наведенных интервалами. Представлен метод комплексирования интервальных данных агрегированием предпочтений. Приведено описание программного обеспечения, разработанного для проведения численных экспериментальных исследований качества работы предложенного метода, и представлены результаты этих исследований.

2.1 Агрегирование предпочтений

2.1.1 Исходные понятия

Пусть $A = \{a_1, a_2, ..., a_n\}$ – множество альтернатив, которые необходимо ранжировать по степени проявления некоторого признака. Отношением ***строгого порядка*** ρ на множестве A называется антирефлексивное ($\overline{a_i \rho a_i}$ для всех i), антисимметричное (если $a_i \rho a_i$ и $a_j \rho a_i \Rightarrow a_i = a_j$ для всех i, j) и транзитивное (если $a_i \rho a_j$ и $a_j \rho a_k \Rightarrow a_i \rho a_k$ для всех i, j, k) бинарное отношение. Для обозначения отношения $a_i \rho a_j$, в котором a_i строго предпочтительнее a_j, будем использовать запись вида $a_i \succ a_j$. Отношение ***безразличия*** τ на множестве A определяется как рефлексивное ($a_i \tau a_i$ для всех i) и симметричное ($a_i \tau a_j \Rightarrow a_j \tau a_i$ для всех i, j) бинарное отношение. В дальнейшем безразличие будем трактовать как ***эквивалентность***. Эквивалентность альтернатив a_i и a_j обозначается в виде $a_i \sim a_j$.

Отношением предпочтения на множестве A будем называть бинарное отношение λ, представляющее собой объединение отношений строгого порядка и эквивалентности, т.е.

$$\lambda = \rho \cup \tau . \qquad (2.1)$$

Таким образом, λ является отношением ***слабого порядка***, обладающим следующими свойствами:

- рефлексивность ($a_i \, \lambda \, a_i$ для всех $a_i \in A$);
- транзитивность (если $a_i \lambda a_j$ и $a_j \lambda a_k \Rightarrow a_i \lambda a_k$ для всех i, j, k);
- полнота (выполняется либо $a_i \, \lambda \, a_j$, либо $a_j \, \lambda \, a_i$ для всех $a_i, a_j \in A$).

Отношение предпочтения может быть представлено в виде ***ранжирования*** альтернатив множества A:

$$\lambda_k : a_1 \succ a_2 \ldots \sim a_s \sim a_t \succ \ldots \sim a_n. \qquad (2.2)$$

Пусть заданы m ранжирований n альтернатив. Тогда множество Λ, состоящее из m ранжирований, т.е.

$$\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}, \qquad (2.3)$$

называется ***профилем предпочтения*** для заданных m и n.

Пример 2.1. Пусть система пожарной сигнализации оповещает о возникновении пожара на одном из этажей здания на основании информации, получаемой от пяти датчиков ($n = 5$), расположенных в различных помещениях [13]. Каждый датчик измеряет три параметра окружающей среды ($m = 3$): температуру (λ_1), задымленность (λ_2) и концентрацию угарного газа (λ_3). Тогда может быть сформирован профиль предпочтения $\Lambda = \{\lambda_1, \lambda_2, \lambda_3\}$ пяти помещений ($a_1, a_2, ..., a_5$) для трех параметров:

$$\begin{aligned} &\lambda_1 : a_5 \succ a_2 \succ a_1 \succ a_3 \succ a_4, \\ &\lambda_2 : a_2 \succ a_1 \sim a_5 \succ a_4 \succ a_3, \\ &\lambda_3 : a_1 \succ a_2 \succ a_5 \succ a_4 \sim a_3. \end{aligned} \qquad (2.4)$$

Для компактности записи будем обозначать альтернативы в ранжировании их индексами. Тогда профиль Λ (2.4) принимает вид:

$$\begin{aligned} &5 \succ 2 \succ 1 \succ 3 \succ 4, \\ &2 \succ 1 \sim 5 \succ 4 \succ 3, \\ &1 \succ 2 \succ 5 \succ 4 \sim 3. \end{aligned} \qquad (2.5)$$

Исходя из полученного профиля, можно определить, в каких помещениях необходимо в первую очередь обеспечить эвакуацию и ликвидацию пламени, с помощью агрегирования предпочтений. ♦

Под ***агрегированием предпочтений*** будем понимать определение для m ранжирований n альтернатив единственного отношения предпочтения β, представляющего наилучший компромисс между ранжированиями исходного профиля. Такое ранжирование β назовем ***ранжированием консенсуса*** [80].

2.1.2 Агрегирование предпочтений на основе правила Кемени

Метод нахождения ранжирования (или отношения) консенсуса агрегированием предпочтений относится к методам ***голосования*** (см. п. 1.4.5), в котором множество A – это множество кандидатов (альтернатив), а Λ – множество избирателей. То, каким образом будет определяться ранжирование консенсуса, зависит от конкретного ***правила голосования***.

Существует множество различных правил голосования [5, 54, 94, 97], среди которых лучшим для нахождения ранжирования консенсуса признано ***правило Кондорсе*** [81, 106]. Однако определяемое им ранжирование консенсуса не обязательно транзитивно, т.е. может существовать некоторое ранжирование консенсуса β, в котором $a_i \succ a_j$ и $a_j \succ a_k$, в то время как $a_k \succ a_i$ при $a_i, a_j, a_k \in \beta$. Такой ***парадокс Кондорсе*** встречается достаточно часто; вероятность его появления выше 50 % при $3 \le m \le \infty$ и $2 \le n \le 10$, если m четное, но наличие эквивалентностей в профиле несколько снижает эту вероятность [81]. Подходящим способом преодоления парадокса Кондорсе считается ***правило Кемени*** [54].

Пусть пространство Π является множеством всех $n!$ линейных (строгих) отношений порядка $\succ$ на множестве A. Каждый линейный порядок соответствует одной из ***перестановок*** первых n натуральных чисел N_n. Правило Кемени позволяет находить ранжирование консенсуса β как линейный порядок альтернатив $\beta \in \Pi$ такой, что определенное в терминах числа парных

несоответствий между ранжированиями расстояние Кемени $D(\beta, \Lambda)$ между β и профилем Λ минимально [71]:

$$\beta = \arg\min_{\lambda \in \Pi} D(\lambda, \Lambda). \quad (2.6)$$

Профиль Λ будем описывать $(n \times n)$ ***матрицей профиля*** $P = [p_{ij}]$:

$$p_{ij} = \sum_{k=1}^{m} d_{ij}^{k}, \quad (2.7)$$

где

$$d_{ij}^{k} = \begin{cases} 0, \text{ если } a_i^k \succ a_j^k \\ 1, \text{ если } a_i^k \sim a_j^k \\ 2, \text{ если } a_i^k \prec a_j^k \end{cases}. \quad (2.8)$$

Одно из свойств матрицы P описывается выражением

$$(p_{ij} + p_{ji})/2 = m \text{ для } i, j = 1, \ldots, n, \quad (2.9)$$

т.е. значение $p_{ij}/2$ можно трактовать, как число предпочтений альтернативы a_j относительно a_i.

Тогда расстояние Кемени $D(\lambda, \Lambda)$ определяется суммой элементов верхней треугольной подматрицы матрицы P, и формула для расстояния $D(\lambda, \Lambda)$ между ранжированием λ и профилем Λ приобретает вид:

$$D(\lambda, \Lambda) = \sum_{i<j} p_{ij}. \quad (2.10)$$

Полученное в соответствии с (2.6) ранжирование консенсуса β называют ***ранжированием*** (или ***медианой***) ***Кемени***. Заметим, что порядок элементов a_i в ранжировании λ соответствует перестановке соответствующих строк и столбцов матрицы P. Иными словами, метод заключается в нахождении такой перестановки строк и столбцов матрицы профиля P, что сумма элементов ее верхней треугольной подматрицы минимальна.

Правило Кемени занимает особое место среди правил голосования благодаря тому, что оно имеет глубокое аксиоматическое обоснование и исключает возможность возникновения парадоксов голосования. Однако при этом оно обладает двумя недостатками.

Первый недостаток состоит в том, что число найденных оптимальных решений (ранжирований консенсуса, представляющих собой строгие порядки), особенно при четных n, может значительно превышать единицу и, таким образом, ограничивать возможности практических применений правила Кемени [76, 78]. Множественные оптимальные решения $\{\beta_1, \beta_2, \ldots, \beta_{N_{\text{kem}}}\}$ порождают парадоксальную ситуацию, когда неопределенность после решения задачи значительно превышает исходную неопределенность.

Для разрешения этой проблемы можно применять специальное преобразование множественных оптимальных строгих порядков в единственный слабый порядок [74].

Будем называть профиль предпочтения $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ ***входным профилем*** задачи о ранжировании Кемени, а множество всех полученных ранжирований консенсуса $B = \{\beta_1, \beta_2, \ldots, \beta_{N_{\text{kem}}}\}$ – ***выходным профилем***. Чтобы получить из выходного профиля единственное итоговое ранжирование $\beta_{\text{fin}} = \Phi(\beta_1, \beta_2, \ldots, \beta_{N_{\text{kem}}})$, воспользуемся следующим введенным в [74] ***правилом свертки***:

$\forall\ i, j$ **if** отношения $a_i \succ a_j$ и $a_i \prec a_j$ встречаются одинаковое количество раз во всех ранжированиях консенсуса
then $a_i \sim a_j$ в β_{fin}
else в β_{fin} включается то отношение строгого предпочтения, которое встречается среди оптимальных решений наибольшее число раз. (2.11)

Таким образом, все ранжирования выходного профиля B являются отношениями строгого порядка, в то время как итоговое ранжирование β_{fin} становится слабым порядком.

Второй недостаток заключается в том, что задача о ранжировании Кемени (2.6) является NP-полной, т.е. не может быть решена за время, пропорциональное полиному от размерности n задачи [3, 36, 81]. В общем случае задача может быть решена методом полного перебора [3, 80], приводящим к экспоненциальному росту времени решения в зависимости от размерности задачи. Однако, для небольшой, но подходящей для практического применения,

размерности задачи $n < 30$, существуют алгоритмы точного решения задачи о ранжировании Кемени за приемлемое время (не более 1-2 с) [20, 23, 30, 80].

В данном диссертационном исследовании для решения задачи о ранжировании Кемени будем использовать реализующий метод ветвей и границ рекурсивный алгоритм RECURSALL, разработанный в научной группе под руководством профессора Муравьева С.В. [76]. Алгоритм позволяет находить все возможные ранжирования Кемени для заданного входного профиля предпочтений и их свертку (2.11).

2.2 Комплексирование интервалов агрегированием предпочтений

2.2.1 Диапазон актуальных значений

Целью комплексирования интервалов является нахождение такой точки x^* на вещественной оси, которая принадлежит максимальному количеству интервалов из $\{I_k\}$ и может служить представителем всех этих интервалов. В этой связи необходимо иметь возможность представлять интервалы множествами, состоящими из конечного числа элементов, т.к., вообще говоря, каждый замкнутый интервал на вещественной оси содержит бесконечное число точек. Для этой цели введем понятие ***диапазона актуальных значений*** (ДАЗ) $A = \{a_1, a_2, \ldots, a_n\}$, на котором существует унаследованное от вещественной оси ***отношение полного порядка*** $a_1 < a_2 < \ldots < a_n$, т.е. транзитивное, антисимметричное и линейное бинарное отношение.

ДАЗ можно было бы сформировать путем объединения интервалов $\{I_k\}$ и разбиения результата объединения на равные подынтервалы. Однако наличие несогласованных интервалов может привести к разрывам интервала, получаемого в результате объединения. Поэтому для исключения разрывов процесс перехода от набора интервалов $\{I_k\}_{k=1}^m$ на вещественной числовой оси к дискретному множеству A осуществляется в три этапа, как показано на рисунке 2.1.

Этап 1. Формирование ДАЗ $[a_1, a_n]$ из исходных интервалов $\{I_k\}_{k=1}^m$.

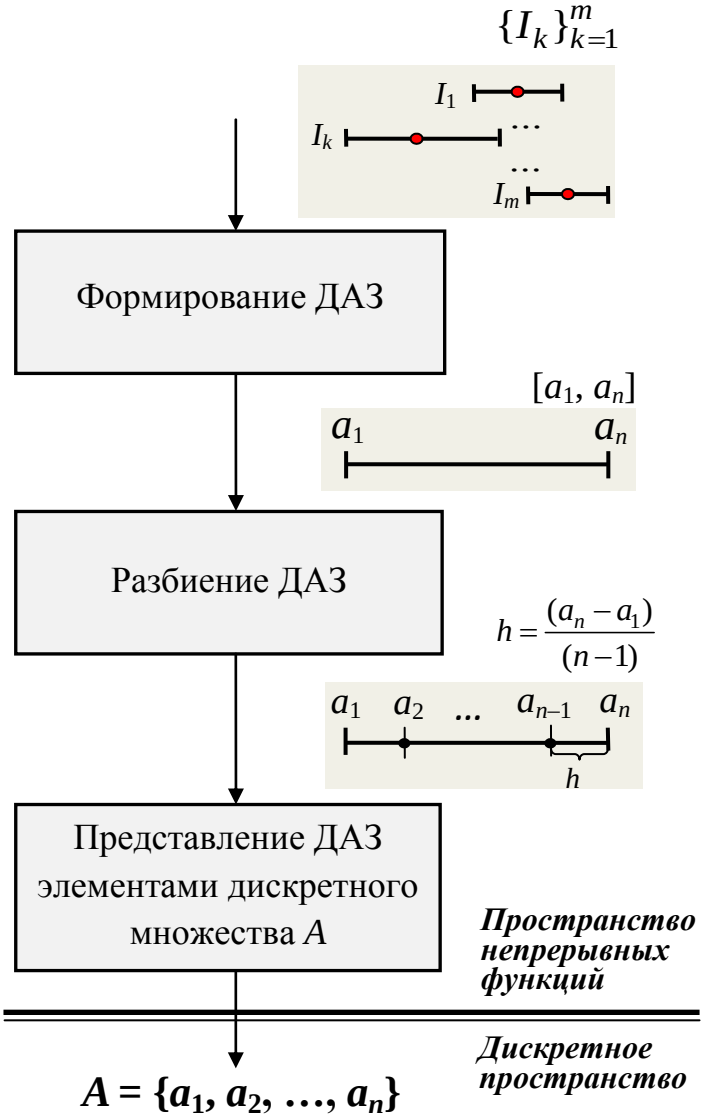


Рисунок 2.1 – Три этапа перехода от непрерывного пространства значений к дискретному при формировании ДАЗ

В качестве нижней границы a_1 выбирается наименьшая нижняя граница l_k для всех интервалов, т.е.

$$a_1 = \min\{l_k \mid k = 1, \ldots, m\}, \tag{2.12}$$

а в качестве верхней границы a_n – наибольшая верхняя граница u_k этих интервалов, т.е.

$$a_n = \max\{u_k \mid k = 1, \ldots, m\}. \tag{2.13}$$

Этап 2. Разбиение ДАЗ на n – 1 равных подынтервалов длиной h, где

$$h = \frac{a_n - a_1}{n-1}. \tag{2.14}$$

Длину h будем называть ***нормой*** разбиения. После разбиения норма определяется формулой

$$h = |a_i - a_{i-1}|, \tag{2.15}$$

где a_i – правая граница i-го подынтервала, $i = 2, \ldots, n$.

Этап 3. Представление ДАЗ элементами дискретного множества $A = \{a_1, a_2, \ldots, a_n\}$, где $A \subset [a_1, a_n]$, а i-й элемент множества определяется как

$$a_i = a_{i-1} + h, \tag{2.16}$$

при $i = 2, \ldots, n$. Число дискретных значений множества A будем называть ***мощностью*** разбиения ДАЗ. Проблема выбора рационального значения мощности n рассмотрена в главе 3.

Таким образом, результатом разбиения ДАЗ является дискретное множество $A_n = \{a_1, a_2, \ldots, a_n\}$, которое может быть использовано для формирования ранжирований, представляющих исходные интервалы $\{I_k\}$.

2.2.2 Инранжирования: ранжирования, наведенные интервалами

Для любого отдельно взятого интервала из множества $\{I_k\}$ можно рассматривать диапазон актуальных значений A как объединение двух непересекающихся множеств: множества A_k, включающего все те элементы A, которые принадлежат интервалу I_k, и дополнения $\bar{A}_k$, включающего все остальные элементы A, т.е. $A = A_k \cup \bar{A}_k$, $A_k \cap \bar{A}_k = \varnothing$, для любого $k = 1, \ldots, m$.

Тогда для любого $k = 1, \ldots, m$, некоторое k-е ранжирование λ_k, наведенное интервалом I_k и состоящее из элементов множества A, должно удовлетворять следующим условиям при $i, j = 1, \ldots, n$:

$$\begin{cases} \text{(i)} \quad a_i \in A_k \wedge a_j \notin A_k \Rightarrow a_i \succ a_j; \\ \text{(ii)} \quad a_i, a_j \in A_k \vee a_i, a_j \notin A_k \Rightarrow a_i \sim a_j; \\ \text{(iii)} \quad a_i \notin A_k \wedge a_j \in A_k \Rightarrow a_i \prec a_j. \end{cases} \tag{2.17}$$

Заметим, что k-е ранжирование состоит из двух классов эквивалентности, образованных элементами введенных выше множеств A_k и $\bar{A}_k$. При этом

элементы класса A_k строго предпочитаются элементам класса $\bar{A}_k$, т.е. всегда $A_k \succ \bar{A}_k$. Следовательно, каждое ранжирование содержит единственный символ строгого порядка $\succ$ и $n - 2$ символов эквивалентности $\sim$.

Пример 2.2. Одно из возможных ранжирований для $n = 5$ имеет вид $\lambda_k = \{a_2 \sim a_3 \succ a_1 \sim a_4 \sim a_5\}$, где $A_k = \{a_2 \sim a_3\}$, $\bar{A}_k = \{a_1 \sim a_4 \sim a_5\}$ и $A_k \succ \bar{A}_k = \{a_2 \sim a_3\} \succ \{a_1 \sim a_4 \sim a_5\}$ (рисунок 2.2). ♦

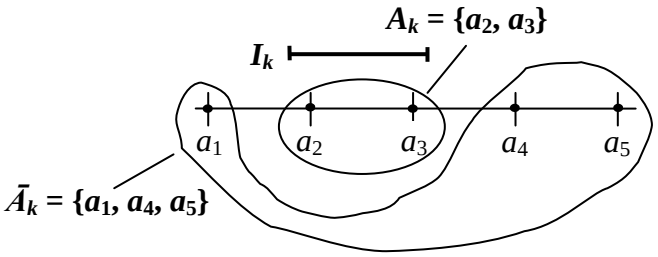


Рисунок 2.2 – Пример разбиения множества A на два подмножества A_k и $\bar{A}_k$ для интервала I_k при $n = 5$

Последовательность элементов $\{a_i\}$ множества A является ***строго монотонной***, т.к. $a_i < a_{i+1}$ для всех $i \in \mathbb{N}$. Ясно, что класс $A_k \subseteq A$ может включать только последовательные наборы элементов из A без пропусков, т.е. индексы этих элементов представляют собой ***отрезок натурального ряда***. Это означает, что разность индексов любой пары соседних элементов a_i и a_j из A_k не может быть больше 1, т.е. справедливо условие

$$\text{(iv) } a_i, a_j \in A_k - \text{соседние элементы} \Rightarrow j \equiv i + 1. \qquad (2.18)$$

Ранжирования, удовлетворяющие условиям (2.17)-(2.18), будем называть ***ранжированиями, наведенными интервалами***, или, в краткой форме, ***инранжированиями***. Таким образом, набор интервалов $\{I_k\}$, $k = 1, \ldots, m$, может быть представлен в соответствии с выражениями (2.17)-(2.18) профилем предпочтений $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_m\}$, где любое λ_k является инранжированием.

Нарушение условия (2.18) приводит к существованию ***запрещенных ранжирований***, для которых тем не менее выполняются условия (2.17). Пространства всех инранжирований и запрещенных ранжирований для $n = 1, \ldots, 5$ приведены в таблице 2.1, где $|A_k| = 1, \ldots, n$ представляет мощность множества A_k.

Для краткости записи символы $\succ$ и $\sim$ опущены, элементы ранжирований обозначены только их индексами, а соответствующие индексы элементов множеств A_k и $\overline{A}_k$ ограничены скобками: например, $\{a_3 \sim a_4 \succ a_1 \sim a_2 \sim a_5\} \Leftrightarrow$ (34)(125).

Таблица 2.1 – Пространства инранжирований и запрещенных ранжирований для $n = 1, \ldots, 5$ с соответствующими мощностями

n	$\lvert A_k \rvert$	Инранжирования	Мощность T_n	Запрещенные ранжирования	Мощность F_n
1	1	1	**1**	∅	**0**
2	1	12 21	**3**	∅	**0**
	2	(12)		∅	
3	1	1(23) 2(13) 3(12)	**6**	∅	**1**
	2	(12)3 (23)1		(13)2	
	3	(123)		∅	
4	1	1(234) 2(134) 3(124) 4(123)	**10**	∅	**5**
	2	(12)(34) (23)(14) (34)(12)		(14)(23) (13)(24) (24)(13)	
	3	(123)4 (234)1		(124)3 (134)2	
	4	(1234)		∅	
5	1	1(2345) 2(1345) 3(1245) 4(1235) 5(1234)	**15**	∅	**16**
	2	(12)(345) (23)(145) (34)(125) (45)(123)		(15)(234) (14)(235) (13)(245) (24)(135) (25)(134) (35)(234)	
	3	(123)(45) (234)(15) (345)(12)		(124)(35) (125)(34) (134)(25) (135)(24) (145)(23)	

Продолжение таблицы 2.1

n	$\lvert A_k \rvert$	Инранжирования	Мощность T_n	Запрещенные ранжирования	Мощность F_n
5	3		**15**	(235)(14) (245)(13)	**16**
	4	(1234)5 (2345)1		(1235)4 (1245)3 (1345)2	
	5	(12345)		∅	

Из рассмотрения таблицы 2.1 (см. также рисунок 2.3) следует, что мощности множеств инранжирований в зависимости от числа разбиений n имеют вид последовательности 1, 3, 6, 10, 15, 21, 28, 36, 45, 55, 66, 78, 91, 105, 120, … (последовательность A000217 в OEIS) [103]. Элементы этой последовательности называются ***треугольными числами*** T_n (рисунок 2.3) и обладают рядом интересных свойств [98, 111].

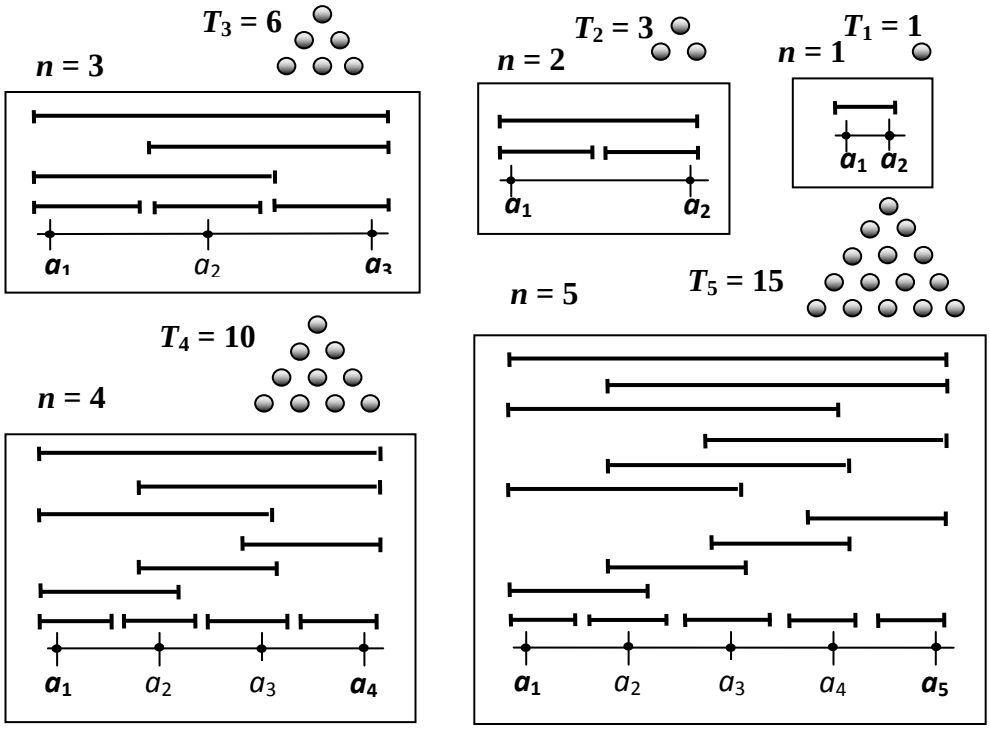


Рисунок 2.3 – Все возможные интервалы и соответствующие треугольные числа для $n = 1, \ldots, 5$

Треугольные числа определяются по формуле:

$$T_n = n(n+1)/2. \qquad (2.19)$$

В свою очередь, мощности F_n запрещенных ранжирований в зависимости от n описываются последовательностью 0, 0, 1, 5, 16, 42, 99, 219, 466, 968, 1981, 4017, 8100, 16278, … (последовательность A002662 в OEIS) [104]. Мощность F_n рассчитывается как

$$F_n = 2^n - 1 - n(n+1)/2. \tag{2.20}$$

Разрешенные комбинации индексов элементов множества A_k при фиксированном n описываются выражением:

$$A_k \in \bigcup_{j=0}^{n-1} \bigcup_{i=1}^{n-j} \{a_{i+j}\}. \tag{2.21}$$

Таблица 2.2 иллюстрирует справедливость формулы (2.21) при n = 5.

Таблица 2.2 – Разрешенные комбинации индексов элементов для n = 5

j	**Элементы A_k**
1	$\{a_i\}$, $i = 1, \ldots, n$,
2	$\{a_i\} \cup \{a_{i+1}\}$, $i = 1, \ldots, (n-1)$,
3	$\{a_i\} \cup \{a_{i+1}\} \cup \{a_{i+2}\}$, $i = 1, \ldots, (n-2)$,
4	$\{a_i\} \cup \{a_{i+1}\} \cup \{a_{i+2}\} \cup \{a_{i+3}\}$, $i = 1, \ldots, (n-3)$,
5	$\{a_i\} \cup \{a_{i+1}\} \cup \{a_{i+2}\} \cup \{a_{i+3}\} \cup \{a_{i+4}\}$, $i = 1, \ldots, (n-4)$

Общее выражение, описывающее пространство инранжирований, имеет следующий вид:

$$\lambda_k \in \bigcup_{j=0}^{n-1} \left\{ \left(\bigcup_{i=1}^{n-j} \{a_{i+j}\} \right) \succ \left(A_k \setminus \bigcup_{i=1}^{n-j} \{a_{i+j}\} \right) \right\}. \tag{2.22}$$

2.2.3 Метод комплексирования интервалов IF&PA

Алгоритм нахождения ранжирования консенсуса по правилу Кемени с учетом особенностей, описанных в п. 2.2, был использован при разработке метода комплексирования интервальных данных агрегированием предпочтений, который будем сокращенно именовать **IF&PA** (interval fusion with preference aggregation). Разработанный метод IF&PA относится к группе методов избыточного комплексирования и осуществляется на первом уровне обработки данных (оценка объекта) в соответствии с JDL моделью (см. пп.1.1-1.2).

Метод IF&PA работает на множестве дискретных значений A (полученных согласно описанной в п. 2.2.1 процедуре), которое позволяет представлять интервалы ранжированиями, и включает в себя 4 основных этапа, представленных ниже.

Этап 1. Формирование диапазона актуальных значений $A = \{a_1, a_2, ..., a_n\}$.

На этом этапе происходит формирование ДАЗ из набора исходных интервалов $\{I_k\}$, $k = 1, \ldots, m$, как описано в п. 2.2.1, расчет нормы h и разбиение ДАЗ на $n - 1$ равных подынтервала длиной h для получения множества дискретных значений $A = \{a_1, a_2, ..., a_n\}$.

Этап 2. Представление интервалов инранжированиями и построение профиля предпочтений $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_m\}$.

На основании исходных интервалов $\{I_k\}$ в соответствии с формулами (2.17)-(2.18) формируются инранжирования λ_k и профиль предпочтения Λ, состоящий из m инранжирований.

Этап 3. Определение значения x^* как лучшей альтернативы в ранжировании консенсуса для профиля Λ.

Этап включает в себя применение рекурсивного алгоритма ветвей и границ RECURSALL для поиска всех ранжирований консенсуса β по правилу Кемени (см. п. 2.1.2), свертку найденных ранжирований в единственное ранжирование консенсуса β_{fin} и выбор наиболее предпочтительной альтернативы a_i в полученном ранжировании в качестве результата комплексирования x^*. При наличии в β_{fin} нескольких наилучших альтернатив за результат x^* принимается их медиана.

Этап 4. Расчет неопределенности ε^* значения x^*.

Находим и исключаем из множества $\{I_k\}$ интервалы, не содержащие значение x^*. При этом мощность множества согласованных интервалов равна m_{con}, а неопределенность ε^* результата комплексирования x^* определяется как

наименьшее из двух значений: максимальная нижняя граница l_k и минимальная верхняя граница u_k исходных интервалов, т.е.:

$$\varepsilon^* = \min\left(\max_{k=1,\ldots,m_{\text{con}}} \{l_k \le x^*\}, \min_{k=1,\ldots,m_{\text{con}}} \{u_k \ge x^*\} \right). \qquad (2.23)$$

Пример работы метода для набора из семи исходных интервалов продемонстрирован на рисунке 2.4.

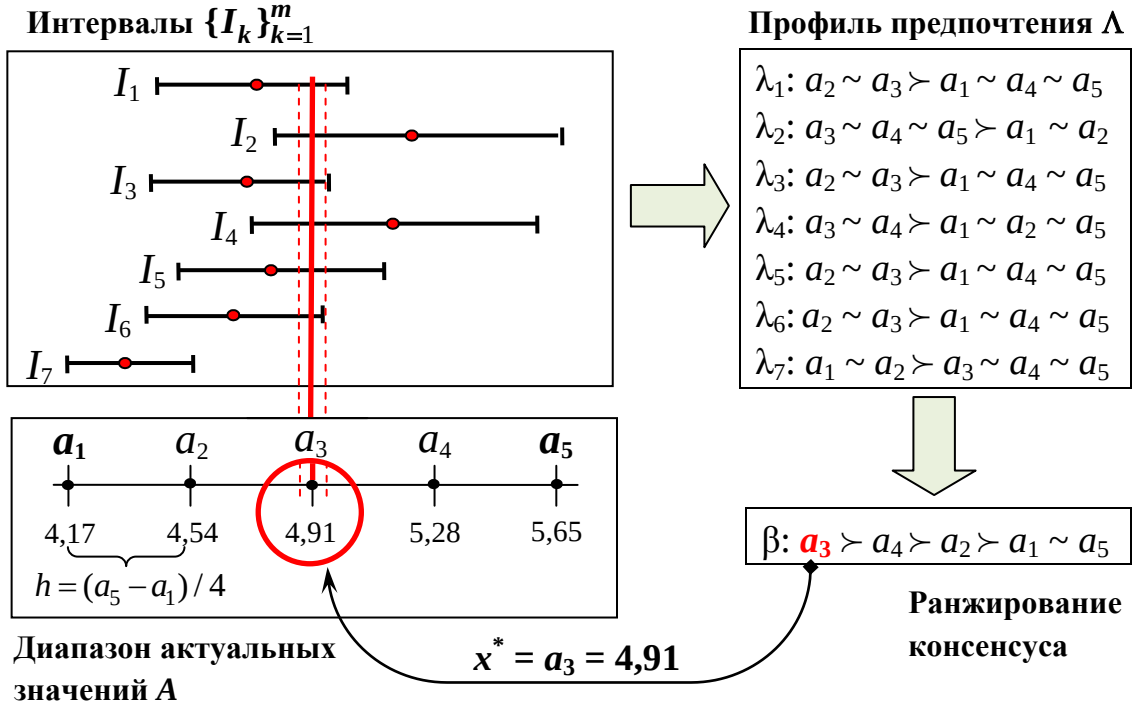


Рисунок 2.4 – Пример набора исходных интервалов; разбиение ДАЗ на 4 подынтервала; его дискретное представление $A = \{a_1, a_2, \ldots, a_5\}$; профиль предпочтения $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_7\}$; результат комплексирования $x^* = 4{,}91$

Метод IF&PA позволяет эффективно решать ряд проблем, возникающих при реализации систем комплексирования (см. п. 1.3). Так, для решения проблемы неточных данных используется избыточность предоставляемых данных, что дает возможность уменьшить неопределенность измерений. Проблема противоречивых данных решается нахождением ранжирования консенсуса, представляющего собой наилучший компромисс исходных данных. Метод также показывает хорошую работоспособность в случае неполных данных, т.к. отсутствие информации от нескольких источников не влияет на работу алгоритма нахождения результата комплексирования x^*.

2.3 Программное обеспечение для численных экспериментов

Для экспериментального исследования предложенного метода IF&PA было разработано специализированное программное обеспечение (ПО) IntFusion в среде программирования NI LabVIEW [12]. IntFusion реализует три метода комплексирования интервальных данных:

- IF&PA,
- одобрительное голосование на основе правила относительного большинства (ОБ) и
- одобрительное голосование на основе правила абсолютного большинства (АБ).

Особенности методов ОБ и АБ рассмотрены в главе 1, п. 1.4.5.

IntFusion выполняет следующие функции:

- генерация интервальных данных по нормальному и равномерному закону распределения;
- определение результата комплексирования x^* и его неопределенности ε^* методами IF&PA, ОБ и АБ;
- отображение и сохранение исходных данных и результатов их обработки.

Заметим, что данное ПО представляет собой универсальный инструмент для комплексирования интервальных данных, однако при внесении небольших изменений может быть адаптирован для применения в конкретных практических приложениях (например, для повышения точности сенсоров в беспроводной сенсорной сети, см. п. 4.2.1).

Структура ПО IntFusion показана на рисунке 2.5. ПО IntFusion включает в себя три основных модуля:

- модуль генерации данных;
- модуль обработки данных;
- модуль визуализации и архивации данных.

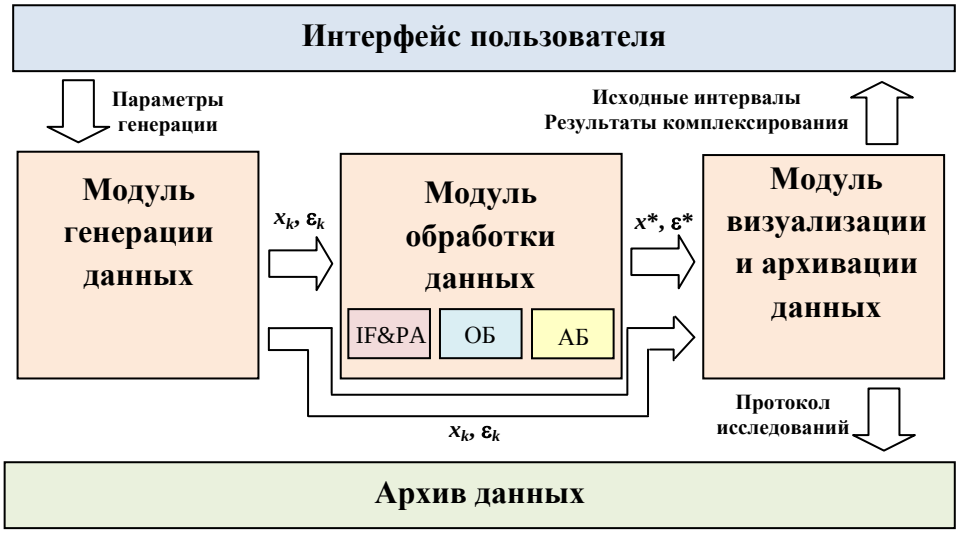


Рисунок 2.5 – Модульная структура программного обеспечения

2.3.1 Интерфейс программы

Внешний вид экранного интерфейса пользователя (лицевой панели) IntFusion представлен на рисунке 2.6.

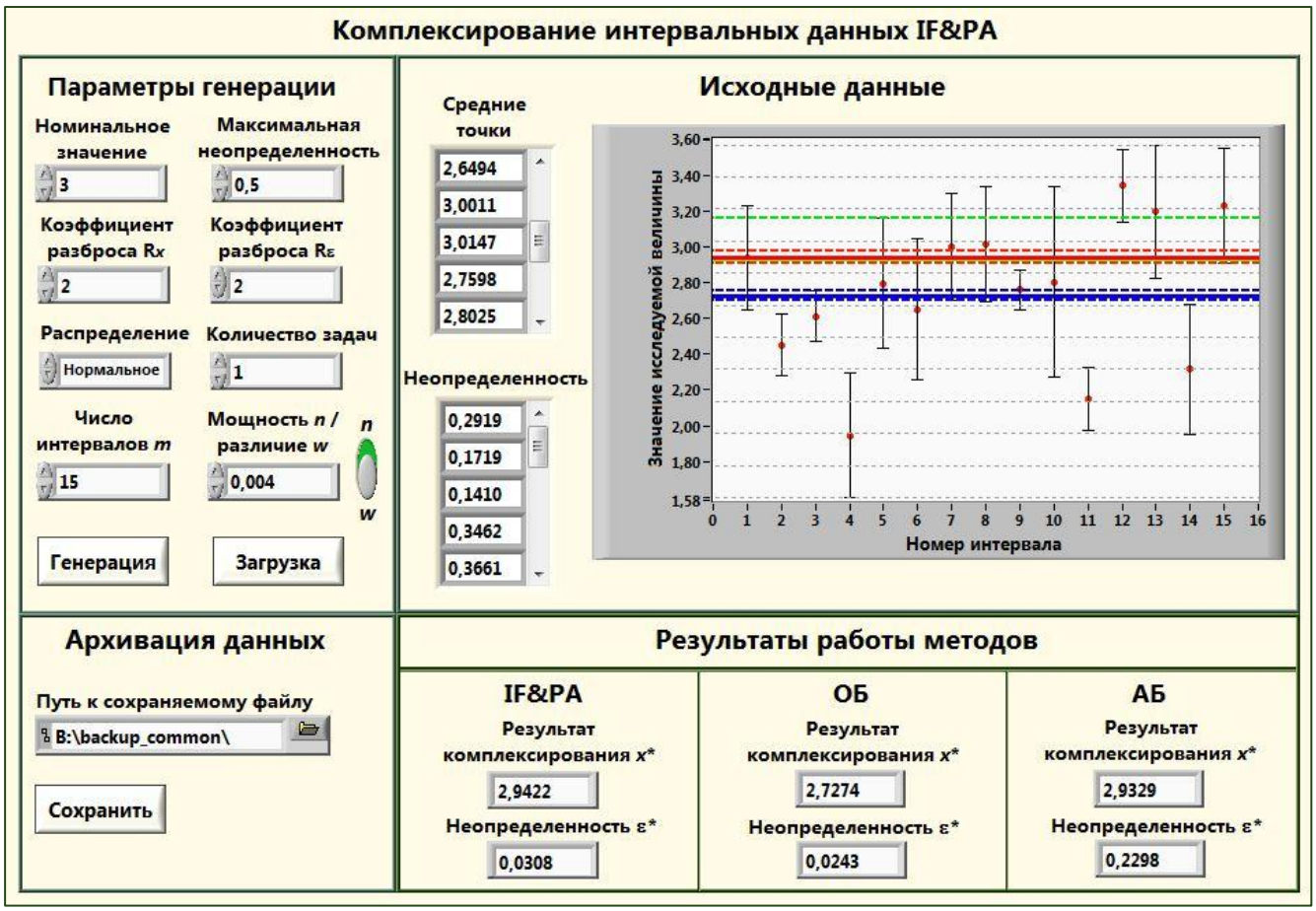


Рисунок 2.6 – Внешний вид интерфейса пользователя ПО IntFusion

Лицевая панель состоит из четырех основных функциональных окон.

В окне «Параметры генерации» осуществляется ввод пользователем следующих параметров генерации данных (см. п. 2.3.2):

- номинальное значение $x_{\text{ном}}$;
- максимальная неопределенность Δx;
- вид распределения;
- число интервалов m;
- допускаемое различие w / мощность n;
- коэффициенты разброса R_x и R_ε генерируемых средних точек x_k и неопределенностей ε_k соответственно;
- количество V индивидуальных задач.

При задании параметров имеется возможность выбора между автоматическим и ручным режимом установки значения мощности n разбиения ДАЗ. Для автоматического расчета значения мощности n (см. главу 3) пользователь вводит значение допускаемого различия w, переключая тумблер в окне «Параметры генерации» в положение «w». Для ручной установки n пользователь переключает тумблер в положение «n» и вводит значение n в соответствующее поле.

После ввода необходимых параметров и нажатия кнопки «Генерация» запускается генерация псевдослучайных данных.

В программе также существует возможность получения исходных интервальных данных посредством загрузки файла с данными в формате Microsoft Excel. Для этого необходимо нажать на кнопку «Загрузка» и в открывшемся диалоговом окне указать путь к файлу для считывания.

В окне «Исходные данные» отображаются сгенерированные средние точки x_k интервалов и соответствующие значения неопределенностей ε_k в виде числовых массивов. График демонстрирует сгенерированные данные в виде интервалов $I_k = [x_k - \varepsilon_k, x_k + \varepsilon_k]$.

Окно «Результаты работы методов» разделено на три части, в каждой из которых отображаются результат комплексирования x^* и его неопределенность

ε*, полученные методами IF&PA, ОБ и АБ соответственно. Обработка данных начинается автоматически после завершения генерации. Результирующие интервалы также отражаются на графике в окне «Исходные данные» для каждого метода разным цветом. Границы результирующих интервалов обозначены пунктирной линией, а середины – сплошной.

Окно «Архивация данных» позволяет указать путь к файлу с протоколом исследований и записать в него экспериментальные данные после нажатия кнопки «Сохранить».

В последующих параграфах приведены алгоритмы генерации и обработки данных.

2.3.2 Модуль генерации данных

Модуль реализует функцию генерации интервальных данных в соответствии с параметрами, заданными пользователем (см. п. 2.3.1). Программа предоставляет возможность генерировать выборку, распределенную по нормальному или равномерному закону.

Формальная запись алгоритма генерации данных по ***нормальному закону*** распределения представлена ниже.

Алгоритм 1 Генерация нормально распределенных значений средних точек x_k и неопределенностей ε_k

---

Вход:

x_{nom}:	номинальное значение
m:	число интервалов
Δx:	максимально допустимая абсолютная неопределенность
R_x:	коэффициент разброса значений средних точек x_k
R_ε:	коэффициент разброса значений неопределенностей ε_k

Пусть:

x_k :	сгенерированное значение средней точки
$b_{\min}$:	минимальная правая граница генерирования x_k
$b_{\max}$:	максимальная правая граница генерирования x_k
a :	левая граница генерирования x_k
b :	правая граница генерирования x_k
μ :	математическое ожидание распределения
σ :	среднеквадратическое отклонение распределения
$x_{\min}$:	минимальное сгенерированное значение x_k
$x_{\max}$:	максимальное сгенерированное значение x_k
ε_k :	сгенерированное значение неопределенности
ε_{av} :	среднее значение неопределенности

d_{min}: минимальная правая граница генерирования ε_k
d_{max}: максимальная правая граница генерирования ε_k
c : левая граница генерирования ε_k
d : правая граница генерирования ε_k
r_{k1}, r_{k2}: равномерно распределенные случайные значения в интервале (0, 1)
z_k : нормально распределенные случайные значения в интервале (0, 1)
y_k : нормально распределенные случайные значения с параметрами μ и σ

[задание параметров генерации средних точек]
1: $b_{min} \leftarrow x_{nom}$; $b_{max} \leftarrow R_x(x_{nom} + \Delta x)$
2: $b \leftarrow$ rand (b_{min}, b_{max}); $a \leftarrow 2\, x_{nom} - b$
3: $\mu \leftarrow (a + b)/2$; $\sigma \leftarrow (\mu - a)/3$
[генерация средних точек]
4: NORMALRAND (μ; σ; x_k) [вызов процедуры генерации]
[задание параметров генерации неопределенностей]
5: $\varepsilon_{av} \leftarrow (x_{max} - x_{min})/\sqrt{18}$
6: $d_{min} \leftarrow \varepsilon_{av}$; $d_{max} \leftarrow R_\varepsilon \cdot (\varepsilon_{av} + \delta x \cdot \varepsilon_{av})$
7: $d \leftarrow$ rand(d_{min}, d_{max}) ; $c \leftarrow 2 \cdot \varepsilon_{av} - d$
8: $\mu \leftarrow (c + d)/2$; $\sigma \leftarrow (\mu - c)/3$
[генерация неопределенностей]
9: NORMALRAND (μ; σ; ε_k) [вызов процедуры генерации]

10: **procedure** NORMALRAND (μ; σ; y_k): [генерация нормально распределенных значений]
11: **for** k = 1 **to** m **do**
[генерация равномерно распределенных значений в интервале (0, 1)]
12: $r_{k1} \leftarrow$ rand (0, 1) ; $r_{k2} \leftarrow$ rand (0, 1)
[получение значений со стандартным нормальным распределением]
13: $z_k \leftarrow \sqrt{-2 \ln r_{k1}} \cos(2\pi r_{k2})$
[получение нормально распределенных значений с параметрами μ и σ]
14: $y_k \leftarrow \mu + \sigma z_k$ [конец процедуры генерации]
end for
Выход:
x_k, ε_k

Номинальное значение x_{nom} исследуемой величины, которое является входным параметром для Алгоритма 1, задается пользователем. Выбор значения x_{nom} осуществляется для конкретной величины в пределах диапазонов, которые варьируются в зависимости от объекта измерений (см. п. 4.2.1).

На шагах 1-2 Алгоритма 1 определяются границы интервала (a, b), в пределах которого будет происходить генерация средних точек x_k. Для расчета максимальной правой границы (b_{max}) генерирования x_k используется значение максимально допустимой абсолютной неопределенности Δx, задаваемое пользователем. Основанием для выбора значения Δx служат следующие

соображения. Если информация о том, с какой неопределенностью получено значение x_k, отсутствует, в качестве Δx принимается погрешность записи (округления) числа, равная половине единицы младшего разряда. В случае если предполагается, что значение x_k является результатом измерения, полученным с помощью некоторого средства измерений (СИ), то в качестве Δx используется значение неопределенности СИ, приведенное в соответствующей технической документации (паспорт, спецификация, сертификат о калибровке и т.д.).

С учетом предположения, что средняя точка x_k находиться в диапазоне ($x_{\text{nom}} \pm \Delta x$), значение параметра $b_{\max}$ определяется с помощью формулы:

$$b_{\max} = x_{\text{nom}} + \Delta x. \qquad (2.24)$$

Для учета возможного расширения или сужения диапазона разброса значений x_k относительно номинального в формулу (2.24) вводится ***коэффициент разброса*** R_x. В таблице 2.3 приведены некоторые значения коэффициента R_x в зависимости от степени (в процентах) предполагаемого изменения (расширения или сужения) диапазона генерирования значений x_k. При этом значения, записанные со знаком плюс, обозначают расширение диапазона, а со знаком минус, напротив, его сужение. Таким образом, параметр $b_{\max}$ определяется по формуле:

$$b_{\max} = R_x \cdot (x_{\text{nom}} + \Delta x). \qquad (2.25)$$

Таблица 2.3 – Значения коэффициента разброса в зависимости от степени предполагаемого изменения диапазона генерирования

Изменение диапазона, %	–70	–50	–20	–10	0	+10	+20	+50	+70	+100
R	0,35	0,75	0,9	0,95	1	1,05	1,1	1,25	1,35	1,5

Процедура определения границ генерирования средних точек x_k схематически показана на рисунке 2.7.

Шаг 3 описывает расчет математического ожидания μ и среднеквадратического отклонения σ распределения случайной величины.

На шаге 4 вызывается процедура NORMALRAND с исходными параметрами μ, σ и x_k для генерации значений, распределенных по нормальному

закону. Описание процедуры NORMALRAND представлено далее (см. шаги 10-14).

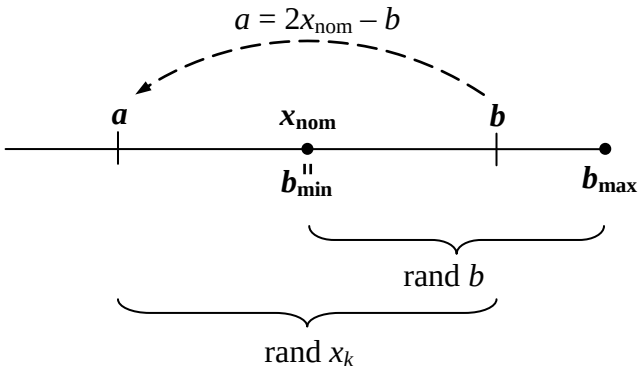


Рисунок 2.7 – Определение границ генерирования средних точек x_k

Шаг 5 служит для установления среднего значения неопределенности для полученных x_k. Известно, что для нормального распределения с математическим ожиданием μ и стандартным отклонением σ интервал $(\mu \pm 3\sigma)$ покрывает примерно 99,73 процента распределения. Таким образом, если верхняя и нижняя границы определяют 99,73 процента, а также если известно, что величина распределена по нормальному закону, то дисперсия может быть оценена по формуле, приведенной в Руководстве [2]:

$$\varepsilon_{av}{}^2 = \frac{x_s^2}{9}, \tag{2.26}$$

где $x_s = (x_{max} - x_{min})/2$ – полуширина интервала, в котором находятся значения x_k. Тогда среднее значение стандартной неопределенности рассчитывается следующим образом:

$$\varepsilon_{av} = \frac{x_{max} - x_{min}}{\sqrt{18}} \ . \tag{2.27}$$

На шагах 6-7 определяется интервал (c, d) генерирования неопределенностей ε_k. Для нахождения параметров c и d применимы те же соображения, что и для параметров a и b. При расчете максимальной правой границы (d_{max}) генерирования неопределенности ε_k используется способ, аналогичный описанному выше для расчета b_{max}. Значение параметра d_{max}, с

учетом предполагаемого изменения диапазона генерирования, рассчитывается по формуле:

$$d_{\max} = R_{\varepsilon} \cdot (\varepsilon_{av} + (\Delta x \cdot \varepsilon_{av}) / x_{nom}), \quad (2.28)$$

где R_{ε} – коэффициент разброса генерирования значений ε_k, определяемый в соответствии с таблицей 2.3.

На шаге 8 рассчитываются математическое ожидание μ и среднеквадратическое отклонение σ распределения случайной величины. На шаге 9 осуществляется генерация случайных значений неопределенностей ε_k с помощью вызова процедуры NORMALRAND с исходными параметрами μ, σ и ε_k.

Шаги 10-14 определяют процедуру NORMALRAND. Для генерации выборки заданного объема используется метод Монте-Карло – метод трансформирования распределений проведением случайной выборки из распределений вероятностей. Процедура генерации случайных значений по методу Монте-Карло представлена в Дополнении [1] к Руководству [2] и состоит из нескольких этапов. На первом этапе (шаг 12) независимо генерируются два случайных значения r_{k1} и r_{k2} со стандартным равномерным распределением на интервале (0, 1). Генерация осуществляется посредством улучшенного генератора Вихманна-Хилла, обладающего хорошими статистическими свойствами и рекомендованного к применению при генерации псевдослучайных чисел из равномерного распределения [1]. На втором этапе (шаг 13) вычисляются случайные значения z_k, имеющие стандартное нормальное распределение, с использованием преобразования Бокса-Мюллера [25]:

$$z_k = \sqrt{-2 \ln r_{k1}} \cos(2\pi r_{k2}). \quad (2.29)$$

На третьем этапе (шаг 14) формируются случайные значения y_k, имеющие нормальное распределение с математическим ожиданием μ и стандартным отклонением σ, по формуле

$$y_k = \mu + \sigma z_k. \quad (2.30)$$

Результатом работы Алгоритма 1 становятся случайные выборки средних точек x_k и неопределенностей ε_k, имеющие нормальное распределение с заданными параметрами.

Формальная запись алгоритма генерации данных по ***равномерному закону*** распределения представлена ниже.

Алгоритм 2 Генерация равномерно распределенных значений средних точек x_k и неопределенностей ε_k

---

Вход:

см. Алгоритм 1

Пусть:

x_k : сгенерированное значение средней точки
$b_{\min}$: минимальная правая граница генерирования x_k
$b_{\max}$: максимальная правая граница генерирования x_k
a : левая граница генерирования x_k
b : правая граница генерирования x_k
r_k : равномерно распределенные случайные значения в интервале (0, 1)
y_k : равномерно распределенные случайные значения в интервале (a, b)
$x_{\min}$: минимальное сгенерированное значение x_k
$x_{\max}$: максимальное сгенерированное значение x_k
ε_k : сгенерированное значение неопределенности
ε_{av} : среднее значение неопределенности
$d_{\min}$: минимальная правая граница генерирования ε_k
$d_{\max}$: максимальная правая граница генерирования ε_k
c : левая граница генерирования ε_k
d: правая граница генерирования ε_k

[задание параметров генерации средних точек]

1: $b_{\min} \leftarrow x_{\text{nom}}$; $b_{\max} \leftarrow R_x (x_{\text{nom}} + \Delta x)$

2: $b \leftarrow \text{rand}\ (b_{\min}, b_{\max})$; $a \leftarrow 2\ x_{\text{nom}} - b$

[генерация средних точек]

3: UNIFORMRAND (a; b; x_k) [вызов процедуры генерации]

[задание параметров генерации неопределенностей]

4: $\varepsilon_{av} \leftarrow (x_{\max} - x_{\min}) / \sqrt{12}$

5: $d_{\min} \leftarrow \varepsilon_{av}$; $d_{\max} \leftarrow R_\varepsilon\ (\varepsilon_{av} + \delta x \cdot \varepsilon_{av})$

6: $d \leftarrow \text{rand}(d_{\min}, d_{\max})$; $c \leftarrow 2 \cdot \varepsilon_{av} - d$

[генерация неопределенностей]

7: UNIFORMRAND(c; d; ε_k) [вызов процедуры генерации]

8: **procedure** UNIFORMRAND (a; b; y_k) [генерация равномерно распределенных значений]

9: **for** k = 1 **to** m **do**

10: $r_k \leftarrow \text{rand}\ (0, 1)$

11: $y_k \leftarrow a + (b - a) r_k$ [получение равномерно распределенных значений в интервале (a, b)]

end for

Выход:

x_k, ε_k

Шаги 1-2 Алгоритма 2 повторяют аналогичные шаги в Алгоритме 1. На шаге 3 происходит вызов процедуры UNIFORMRAND с исходными параметрами a, b и x_k для генерации значений средних точек x_k, распределенных по равномерному закону в заданном интервале (a, b). Описание процедуры UNIFORMRAND приведено далее (см. шаги 8-11).

Шаг 4 служат для установления среднего значения неопределенности для полученных x_k. Поскольку распределение равномерное, предполагается, что x_k с одинаковой вероятностью может принимать любые значения в интервале ($x_{\min}$, $x_{\max}$). Тогда дисперсия может быть оценена по формуле, приведенной в Руководстве [2]:

$$\varepsilon_{\text{av}}^2 = \frac{(x_{\max} - x_{\min})^2}{12}, \tag{2.31}$$

и соответственно, среднее значение стандартной неопределенности рассчитывается как

$$\varepsilon_{\text{av}} = \frac{x_{\max} - x_{\min}}{\sqrt{12}} \ . \tag{2.32}$$

Шаги 5-6 Алгоритма 2 аналогичны шагам 6-7 Алгоритма 1. На шаге 7 осуществляется генерация случайной выборки значений неопределенности ε_k с помощью вызова процедуры UNIFORMRAND.

Шаги 8-11 определяют процедуру UNIFORMRAND. Для генерации равномерно распределенных значений методом Монте-Карло [1] посредством улучшенного генератора Вихманна-Хилла (шаг 10) генерируются равномерно распределенные значения r_k в интервале (0, 1), с помощью которых по формуле (2.33) формируется выборка значений y_k (шаг 11), распределенных равномерно в интервале (a, b):

$$y_k \leftarrow a + (b - a) r_k \ . \tag{2.33}$$

Результатом работы Алгоритма 2 становятся случайные выборки средних точек x_k и неопределенности ε_k, имеющие равномерное распределение с заданными параметрами.

На основании полученных значений x_k и ε_k формируются исходные интервалы $I_k = [x_k - \varepsilon_k, x_k + \varepsilon_k]$.

2.3.3 Модуль обработки данных

Модуль реализует процедуру комплексирования сгенерированных интервальных данных методами IF&PA, ОБ и АБ.

Алгоритм метода IF&PA приведен в параграфе 2.2.3. Формальная запись алгоритма комплексирования методом ОБ представлена ниже.

Алгоритм 3 Комплексирование интервалов с помощью одобрительного голосования на основе правила относительного большинства

Вход:

m : число интервалов
x_k : средняя точка интервала I_k
l_k : нижняя граница интервала I_k
u_k : верхняя граница интервала I_k

Пусть:

L_k : число интервалов, включающих нижнюю границу k-го интервала
U_k : число интервалов, включающих верхнюю границу k-го интервала
$L_{\max}$: максимальное число интервалов, включающих нижнюю границу одного из интервалов
$U_{\max}$: максимальное число интервалов, включающих верхнюю границу одного из интервалов
$l_г$: нижняя граница результирующего интервала
$u_г$: верхняя граница результирующего интервала
x^*: результат комплексирования (средняя точка результирующего интервала)
ε^*: неопределенность результата комплексирования

[нахождение согласованных интервалов]

1: **for** $k = 1$ **to** m **do**
2: $L_k \leftarrow 0$; $U_k \leftarrow 0$
3: **for** $j = 1$ **to** m **do**
4: **if** $j \neq k$ **then**
5: **if** $l_k \in (l_j, u_j)$ **then** $L_k + 1$
6: **if** $u_k \in (l_j, u_j)$ **then** $U_k + 1$
end for

[определение границ результирующего интервала]

7: **if** $L_k > L_{\max}$ **then** $L_{\max} \leftarrow L_k$; $l_г \leftarrow l_k$
8: **if** $U_k > U_{\max}$ **then** $U_{\max} \leftarrow U_k$; $u_г \leftarrow u_k$
end for
9: $x^* \leftarrow (u_г + l_г)/ 2$
10: $\varepsilon^* = x^* - l_г$

Выход:

x^*, ε^*

Формальная запись алгоритма комплексирования методом АБ представлена ниже.

Алгоритм 4 Комплексирование интервалов с помощью одобрительного голосования на основе правила абсолютного большинства

Вход:

см. Алгоритм 3

Пусть:

q: значение уровня согласованности результирующего интервала
u_r: верхняя граница результирующего интервала
l_r: нижняя граница результирующего интервала
x^*: результат комплексирования (средняя точка результирующего интервала)
ε^*: неопределенность результата комплексирования

1: $q \leftarrow \lceil 0{,}5 \cdot m \rceil$
2: **for** $k = 1$ **to** m
3: **sort** u_k (max; min) [сортировка от максимального к минимальному]
4: **if** $u_q \le l_k$ **then** $q \leftarrow q + 1$
end for
5: $u_r \leftarrow u_q$
6: **for** $k = 1$ **to** m
7: **sort** l_k (min; max) [сортировка от минимального к максимальному]
8: **if** $l_q \ge u_k$ **then** $q \leftarrow q + 1$
end for
9: $l_r \leftarrow l_q$
10: $x^* \leftarrow (u_r + l_r)/2$
11: $\varepsilon^* = x^* - l_r$

Выход:

x^*, ε^*

Результатом работы всех трех реализованных в программе алгоритмов обработки данных является значение x^* результата комплексирования и соответствующая неопределенность ε^*.

2.3.4 Модуль визуализации и архивации данных

Функция модуля заключается в представлении информации на лицевой панели программы в числовом и графическом виде (рисунок 2.6), а также сохранении результатов работы программы в форме протокола исследований в формате Microsoft Excel. Протокол включает в себя следующую информацию:

- заданные параметры генерации (см. п. 2.3.2);
- исходные сгенерированные интервальные данные в виде массивов значений x_k , ε_k, l_k , u_k;
- значение мощности n разбиения ДАЗ;

- полученный результат комплексирования x^* с соответствующей неопределенностью ε^*.

2.4 Экспериментальные исследования свойств метода IF&PA

В ходе экспериментальных исследований генерировались случайные интервальные данные, которые подвергались обработке методом IF&PA и, для сравнения, методом ОБ и методом АБ. Выбор методов для сравнения обусловлен тем, что правила большинства признаны наиболее робастными из существующих правил голосования [34, 67, 97].

2.4.1 Выбор меры точности, робастности и достоверности исследуемых методов

Суждение о качестве работы методов основывалось на оценке точности, робастности и достоверности.

Под ***робастностью*** метода комплексирования будем понимать независимость результата комплексирования x^* от закона распределения интервальных данных.

Точностью метода комплексирования будем называть близость результата комплексирования x^* к номинальному значению x_{nom}.

Как о точности, так и о робастности метода можно судить по величине ***отклонения*** ξ результата комплексирования x^* от номинального значения x_{nom}:

$$\xi = |x_{\text{nom}} - x^*|. \qquad (2.34)$$

Критерий ξ допускает наглядную графическую интерпретацию, которая одновременно демонстрирует отличия в точности и робастности между исследуемыми методами. Результаты решения 100 задач некоторым методом можно представить ***кривой*** $\boldsymbol{\xi(\nu)}$, где значения ξ упорядочены по возрастанию, ν – номер индивидуальной задачи после упорядочения, $\nu = 1, 2, \ldots, 100$. Для примера, на рисунке 2.8 показаны графики $\xi(\nu)$, соответствующие двум различным методам. На этих графиках робастность метода демонстрируется шириной "лепестка", образованного кривыми $\xi(\nu)$ для нормального и равномерного

распределений: чем уже лепесток, тем большей робастностью характеризуется метод. Точность метода демонстрируется средним расстоянием между кривой $\xi(v)$ и осью абсцисс.

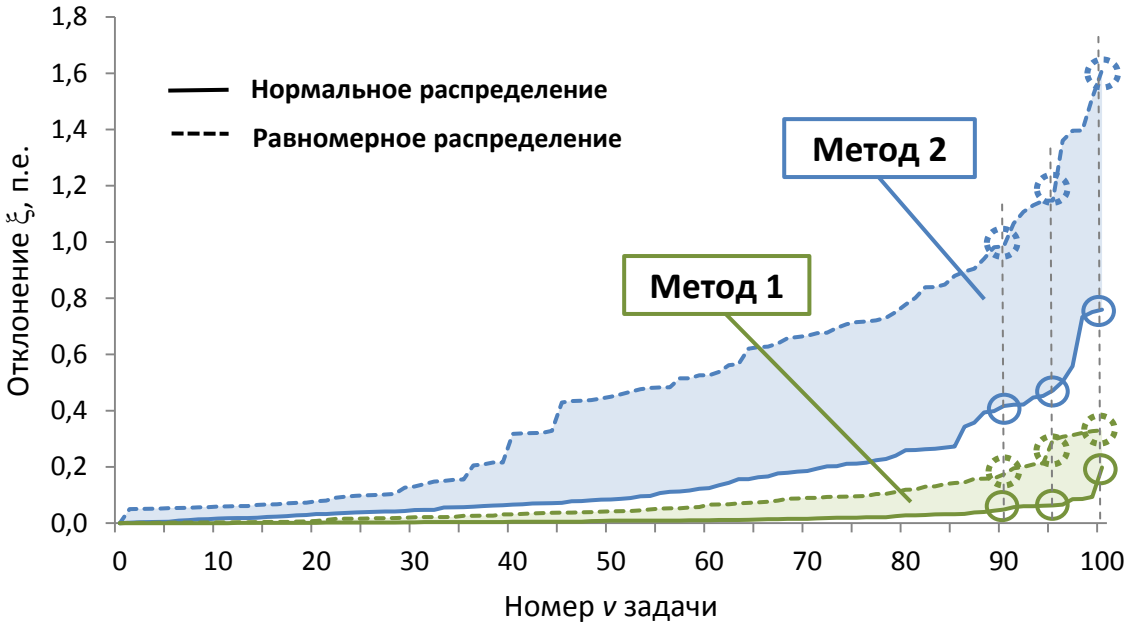


Рисунок 2.8 – Отклонения ξ, соответствующие двум произвольным методам

Введем дополнительный показатель, характеризующий достоверность результатов метода [7]. Под ***достоверностью*** метода комплексирования будем понимать степень доверия к полученным результатам x^*. В качестве численного показателя достоверности будем использовать оценки вероятностей $P(\xi \le \xi_{гр})$ того, что отклонение ξ не превышает некоторое фиксированное значение $\xi_{гр}$. Для определения оценок воспользуемся следующим алгоритмом.

1. Генерация равномерной шкалы значений $\xi_{гр}$ с начальным значением 0,00 и шагом 0,01.

2. Сравнение отклонений ξ, полученных в каждой индивидуальной задаче, со значениями шкалы $\xi_{гр}$.

3. Для каждого $\xi_{гр}$ определение количества V' задач, для которых ξ меньше фиксированного значения $\xi_{гр}$.

4. Нахождение оценок вероятностей по формуле

$$P(\xi \le \xi_{гр}) = \frac{V'}{V}, \qquad (2.35)$$

где V' – число индивидуальных задач, для которых выполняется условие $\xi \le \xi_{гр}$, а $V = 100$ – общее число задач. Так, если значение ξ меньше или равно значению 0,14 в 90 задачах из 100, то считаем, что вероятность $P(0{,}14) = 0{,}90$.

В ходе исследований были проведены пять экспериментальных прогонов по сто индивидуальных задач, в каждом из которых были сгенерированы данные, распределенные по нормальному и равномерному законам. Таким образом, объем исследуемой выборки составил 1000 задач, по 500 задач для нормального и равномерного распределения. Задаваемые параметры генерации данных (см. п. 2.3.2) приведены в таблице 2.4.

Таблица 2.4 – Параметры генерации данных

m	x_{nom}, п.е.	Δx, п.е.	R_x	R_ε
15	3	1	1	2

Значение мощности n разбиения ДАЗ для каждой индивидуальной задачи определялось с помощью процедуры, описанной в главе 3.

Пример сгенерированных данных для одной из 100 задач прогона 1 ($v = 7$, нормальное распределение) приведен в таблице 2.5.

Таблица 2.5 – Сгенерированные данные для индивидуальной задачи 7 (нормальное распределение)

k	x_k	ε_k
1	2,8990	0,2059
2	3,1957	0,3731
3	3,1392	0,2838
4	3,3188	0,2362
5	2,8857	0,3693
6	3,3602	0,2719
7	2,9513	0,2575
8	2,2839	0,1966
9	2,8003	0,2354
10	2,9990	0,2688
11	3,2272	0,2544
12	3,1280	0,1708
13	3,1261	0,1686
14	3,2338	0,1098
15	2,8365	0,2210

2.4.2 Результаты экспериментальных исследований

В ходе эксперимента каждым из исследуемых методов определялся результат комплексирования x^*, отклонение ξ значения x^* от номинального x_{nom} и неопределенность ε^*. Результаты обработки данных для одного экспериментального прогона из 100 задач для нормального и равномерного распределений показаны в таблице 2.6 и таблице 2.7 соответственно.

Таблица 2.6 – Результаты обработки данных для 100 индивидуальных задач (нормальное распределение)

v	IF&PA			Метод ОБ			Метод АБ		
	x^*	ε^*	ξ	x^*	ε^*	ξ	x^*	ε^*	ξ
1	3,102	0,019	0,102	3,120	0,037	0,120	3,082	0,091	0,082
2	2,977	0,009	0,023	2,969	0,002	0,031	2,967	0,059	0,033
3	3,023	0,001	0,023	3,037	0,015	0,037	3,025	0,101	0,025
4	2,767	0,035	0,233	2,666	0,009	0,334	2,666	0,009	0,334
5	2,969	0,004	0,031	2,973	0,042	0,027	2,965	0,053	0,035
6	3,069	0,001	0,069	3,207	0,007	0,207	3,157	0,166	0,157
7	2,970	0,013	0,030	3,097	0,008	0,097	3,075	0,220	0,075
8	2,923	0,010	0,077	2,889	0,014	0,111	2,918	0,091	0,082
9	3,005	0,003	0,005	3,002	0,008	0,002	3,003	0,018	0,003
10	2,970	0,001	0,030	3,020	0,002	0,020	2,986	0,075	0,014
11	3,134	0,008	0,134	3,130	0,004	0,130	3,135	0,111	0,135
12	2,998	0,001	0,002	2,998	0,001	0,002	2,997	0,004	0,003
13	2,837	0,042	0,163	2,800	0,005	0,200	3,005	0,209	0,005
14	2,943	0,017	0,057	2,958	0,002	0,042	2,935	0,040	0,065
15	3,042	0,001	0,042	3,038	0,005	0,038	3,042	0,026	0,042
16	2,946	0,011	0,054	3,274	0,005	0,274	3,104	0,175	0,104
17	2,994	0,001	0,006	3,005	0,003	0,005	2,986	0,048	0,014
18	3,011	0,001	0,011	3,012	0,001	0,012	3,006	0,025	0,006
19	2,893	0,002	0,107	2,961	0,009	0,039	2,931	0,039	0,069
20	2,982	0,028	0,018	3,016	0,004	0,016	2,989	0,060	0,011
21	3,105	0,018	0,105	2,992	0,001	0,008	3,078	0,140	0,078
22	3,132	0,029	0,132	3,157	0,004	0,157	2,956	0,256	0,044
23	2,970	0,006	0,030	2,969	0,007	0,031	2,986	0,028	0,014
24	3,070	0,001	0,070	3,093	0,005	0,093	3,151	0,083	0,151
25	3,021	0,003	0,021	3,021	0,004	0,021	2,997	0,119	0,004
26	3,006	0,002	0,006	2,999	0,001	0,001	3,004	0,010	0,004
27	3,080	0,006	0,080	3,076	0,009	0,076	3,042	0,161	0,042
28	2,943	0,025	0,057	2,926	0,013	0,074	2,908	0,114	0,092
29	3,135	0,010	0,135	3,203	0,011	0,203	3,153	0,325	0,153

Продолжение таблицы 2.6

v	IF&PA			Метод ОБ			Метод АБ		
	x*	**ε***	**ξ**	***x****	**ε***	**ξ**	***x****	**ε***	**ξ**
30	2,905	0,005	0,095	2,904	0,006	0,096	3,003	0,142	0,003
31	3,065	0,006	0,065	3,051	0,019	0,051	3,046	0,084	0,046
32	3,090	0,005	0,090	3,103	0,008	0,103	3,060	0,158	0,060
33	3,001	0,001	0,001	2,999	0,007	0,001	3,026	0,058	0,026
34	2,885	0,001	0,115	2,896	0,011	0,104	2,930	0,078	0,070
35	3,215	0,032	0,215	3,285	0,004	0,285	3,205	0,098	0,205
36	2,953	0,005	0,047	2,943	0,001	0,057	3,008	0,088	0,008
37	3,010	0,021	0,010	3,035	0,046	0,035	3,026	0,068	0,026
38	2,968	0,040	0,032	2,928	0,001	0,072	2,955	0,239	0,045
39	3,024	0,002	0,024	3,043	0,021	0,043	2,898	0,187	0,102
40	2,989	0,001	0,011	2,998	0,004	0,002	2,960	0,098	0,040
41	2,978	0,002	0,022	2,934	0,019	0,066	2,960	0,182	0,040
42	2,730	0,011	0,270	2,743	0,024	0,257	2,710	0,127	0,290
43	3,003	0,001	0,003	3,004	0,001	0,004	3,003	0,005	0,003
44	2,933	0,005	0,067	2,997	0,001	0,003	3,009	0,082	0,009
45	2,880	0,025	0,120	2,906	0,001	0,095	2,804	0,119	0,196
46	2,950	0,001	0,050	2,951	0,001	0,049	2,951	0,016	0,049
47	2,926	0,006	0,074	2,934	0,013	0,066	2,934	0,013	0,066
48	3,008	0,019	0,008	2,930	0,001	0,071	2,993	0,096	0,007
49	2,953	0,004	0,047	2,957	0,009	0,043	3,064	0,237	0,064
50	3,076	0,015	0,076	3,119	0,005	0,119	3,128	0,182	0,128
51	2,921	0,073	0,079	2,978	0,017	0,022	2,980	0,332	0,020
52	2,978	0,016	0,022	2,984	0,022	0,016	2,984	0,022	0,016
53	3,096	0,012	0,096	3,004	0,086	0,004	3,054	0,164	0,054
54	2,912	0,015	0,088	2,933	0,005	0,067	2,933	0,005	0,067
55	2,894	0,018	0,106	2,910	0,012	0,090	2,891	0,063	0,109
56	2,896	0,001	0,104	3,087	0,067	0,087	3,101	0,202	0,101
57	3,131	0,017	0,131	3,145	0,002	0,145	3,136	0,125	0,136
58	2,936	0,013	0,064	3,026	0,128	0,026	2,990	0,153	0,010
59	2,987	0,013	0,013	2,790	0,004	0,211	2,790	0,004	0,211
60	3,027	0,008	0,027	3,022	0,002	0,022	3,042	0,093	0,042
61	2,947	0,020	0,053	2,974	0,019	0,026	2,991	0,150	0,009
62	2,999	0,001	0,001	2,999	0,001	0,001	2,999	0,001	0,001
63	3,014	0,012	0,014	2,996	0,002	0,004	2,998	0,035	0,003
64	2,989	0,031	0,011	3,041	0,024	0,041	2,989	0,108	0,011
65	3,144	0,019	0,144	3,222	0,056	0,222	3,159	0,172	0,159
66	3,043	0,051	0,043	2,857	0,010	0,143	3,024	0,177	0,024
67	2,909	0,003	0,091	2,912	0,007	0,088	2,986	0,118	0,014
68	3,016	0,018	0,016	3,003	0,006	0,003	3,021	0,060	0,021
69	3,042	0,001	0,042	3,038	0,005	0,038	3,038	0,005	0,038

Продолжение таблицы 2.6

v	IF&PA			Метод ОБ			Метод АБ		
	x*	**ε***	**ξ**	***x****	**ε***	**ξ**	***x****	**ε***	**ξ**
70	3,087	0,003	0,087	3,080	0,010	0,080	3,056	0,052	0,056
71	2,991	0,001	0,009	3,094	0,002	0,094	3,118	0,129	0,118
72	2,986	0,002	0,014	2,994	0,006	0,006	2,997	0,028	0,003
73	2,979	0,005	0,021	3,017	0,002	0,017	3,014	0,056	0,014
74	3,042	0,006	0,042	3,018	0,014	0,018	3,016	0,053	0,016
75	2,895	0,027	0,105	2,910	0,011	0,090	2,930	0,107	0,070
76	3,030	0,009	0,030	3,016	0,023	0,016	3,005	0,107	0,005
77	2,925	0,004	0,075	3,055	0,001	0,055	3,053	0,054	0,053
78	3,018	0,026	0,018	3,026	0,144	0,026	3,022	0,185	0,022
79	3,020	0,003	0,020	3,036	0,019	0,036	3,032	0,025	0,032
80	3,096	0,008	0,096	3,131	0,043	0,131	3,137	0,155	0,137
81	3,034	0,005	0,034	3,001	0,004	0,001	3,003	0,052	0,003
82	3,007	0,002	0,007	3,006	0,000	0,006	3,002	0,013	0,002
83	2,998	0,006	0,002	2,982	0,021	0,018	2,975	0,028	0,025
84	2,987	0,006	0,013	3,003	0,022	0,003	3,026	0,100	0,026
85	2,914	0,079	0,086	2,842	0,015	0,158	2,925	0,117	0,075
86	3,193	0,008	0,193	3,195	0,011	0,195	3,178	0,064	0,178
87	2,998	0,001	0,002	2,997	0,001	0,004	2,997	0,001	0,004
88	3,040	0,001	0,040	2,996	0,038	0,004	2,994	0,047	0,006
89	3,022	0,007	0,022	3,033	0,018	0,033	3,050	0,079	0,050
90	2,935	0,009	0,065	2,962	0,138	0,038	2,963	0,197	0,037
91	2,933	0,002	0,067	2,978	0,047	0,022	2,931	0,103	0,069
92	3,045	0,011	0,045	3,021	0,001	0,021	2,964	0,092	0,036
93	2,913	0,050	0,087	2,955	0,009	0,045	2,946	0,102	0,054
94	2,919	0,009	0,081	2,947	0,007	0,053	2,952	0,106	0,048
95	3,169	0,060	0,169	3,159	0,070	0,159	3,096	0,182	0,096
96	2,933	0,005	0,067	2,947	0,020	0,053	2,942	0,057	0,059
97	2,962	0,018	0,038	3,015	0,021	0,015	2,895	0,290	0,105
98	2,764	0,010	0,236	2,826	0,002	0,174	2,756	0,103	0,244
99	3,002	0,004	0,002	2,999	0,001	0,001	2,999	0,001	0,001
100	3,077	0,019	0,077	3,139	0,043	0,139	3,094	0,137	0,094

Таблица 2.7 – Результаты обработки данных для 100 индивидуальных задач (равномерное распределение)

v	IF&PA			Метод ОБ			Метод АБ		
	x*	**ε***	**ξ**	***x****	**ε***	**ξ**	***x****	**ε***	**ξ**
1	3,016	0,024	0,016	3,011	0,030	0,011	3,007	0,108	0,007
2	2,912	0,018	0,088	2,938	0,020	0,062	2,856	0,196	0,144
3	3,013	0,050	0,013	3,061	0,214	0,061	3,078	0,413	0,078

Продолжение таблицы 2.7

v	IF&PA			Метод ОБ			Метод АБ		
	x*	**ε***	**ξ**	***x****	**ε***	**ξ**	***x****	**ε***	**ξ**
4	2,866	0,074	0,134	2,643	0,009	0,358	2,865	0,367	0,135
5	3,043	0,048	0,043	3,211	0,311	0,211	3,287	0,523	0,287
6	2,898	0,035	0,102	3,039	0,007	0,039	3,039	0,007	0,039
7	3,043	0,032	0,043	3,071	0,061	0,071	3,085	0,107	0,085
8	3,038	0,090	0,038	3,558	0,020	0,558	3,340	0,393	0,340
9	2,933	0,002	0,067	2,934	0,003	0,066	2,962	0,055	0,038
10	3,054	0,002	0,054	3,038	0,018	0,038	3,011	0,224	0,011
11	3,019	0,048	0,019	3,759	0,006	0,759	3,759	0,006	0,759
12	2,712	0,005	0,288	2,914	0,207	0,086	2,915	0,235	0,085
13	2,814	0,055	0,186	2,760	0,004	0,240	2,784	0,085	0,216
14	3,054	0,002	0,054	3,056	0,001	0,056	3,010	0,138	0,010
15	2,919	0,009	0,081	2,964	0,009	0,036	3,017	0,066	0,017
16	3,000	0,001	0,000	3,004	0,001	0,004	3,002	0,005	0,002
17	2,880	0,025	0,120	2,888	0,034	0,112	3,090	0,327	0,090
18	3,083	0,049	0,083	3,057	0,023	0,057	3,105	0,450	0,105
19	2,958	0,005	0,042	2,966	0,013	0,035	2,991	0,100	0,009
20	2,587	0,253	0,413	2,267	0,016	0,733	2,573	0,431	0,427
21	2,979	0,020	0,021	2,977	0,019	0,023	2,962	0,052	0,038
22	2,973	0,100	0,027	3,183	0,002	0,183	3,238	0,449	0,238
23	3,137	0,001	0,137	3,177	0,040	0,177	2,916	0,480	0,084
24	3,201	0,014	0,201	3,156	0,059	0,156	3,200	0,401	0,200
25	3,195	0,029	0,195	3,268	0,006	0,268	3,078	0,405	0,078
26	2,868	0,099	0,132	2,971	0,053	0,029	3,017	0,475	0,017
27	2,871	0,012	0,129	2,888	0,028	0,113	2,826	0,415	0,174
28	2,989	0,074	0,011	2,967	0,189	0,033	2,959	0,262	0,041
29	2,525	0,004	0,475	2,528	0,007	0,472	2,624	0,150	0,376
30	3,181	0,014	0,181	3,166	0,015	0,166	2,953	0,480	0,047
31	2,956	0,013	0,044	2,936	0,006	0,064	2,955	0,024	0,045
32	2,604	0,018	0,396	2,497	0,060	0,504	2,506	0,413	0,494
33	2,900	0,007	0,100	2,933	0,032	0,067	2,915	0,134	0,085
34	3,006	0,042	0,006	3,047	0,002	0,047	3,067	0,043	0,067
35	3,032	0,030	0,032	3,047	0,026	0,047	3,053	0,130	0,053
36	2,975	0,006	0,025	3,107	0,023	0,107	3,114	0,094	0,114
37	2,892	0,055	0,108	2,982	0,004	0,018	2,917	0,159	0,083
38	2,972	0,011	0,028	2,982	0,001	0,018	2,983	0,029	0,017
39	2,747	0,017	0,253	2,922	0,021	0,078	2,975	0,244	0,025
40	3,239	0,353	0,239	3,751	0,021	0,751	3,751	0,021	0,751
41	3,129	0,021	0,129	3,272	0,120	0,272	3,231	0,165	0,231
42	2,968	0,006	0,032	2,978	0,001	0,022	2,971	0,031	0,029
43	2,835	0,031	0,165	2,772	0,004	0,228	2,772	0,004	0,228

Продолжение таблицы 2.7

v	IF&PA			Метод ОБ			Метод АБ		
	x*	**ε***	**ξ**	***x****	**ε***	**ξ**	***x****	**ε***	**ξ**
44	2,850	0,007	0,150	2,911	0,018	0,089	2,819	0,274	0,182
45	2,999	0,001	0,001	3,001	0,002	0,001	2,991	0,019	0,009
46	3,104	0,033	0,104	3,097	0,251	0,097	3,037	0,620	0,037
47	3,017	0,014	0,017	3,042	0,008	0,042	3,024	0,241	0,024
48	2,851	0,026	0,149	2,957	0,004	0,043	3,000	0,261	0,000
49	2,856	0,030	0,144	2,878	0,052	0,122	2,878	0,052	0,122
50	2,992	0,050	0,008	3,260	0,085	0,260	3,129	0,263	0,129
51	2,826	0,005	0,174	3,452	0,033	0,452	3,421	0,100	0,421
52	3,153	0,009	0,153	3,163	0,020	0,163	3,160	0,129	0,160
53	3,079	0,042	0,079	3,047	0,010	0,047	3,019	0,205	0,019
54	2,942	0,007	0,058	2,930	0,002	0,070	2,940	0,014	0,060
55	2,789	0,013	0,211	2,798	0,004	0,202	2,798	0,004	0,202
56	3,082	0,030	0,082	3,083	0,032	0,083	3,046	0,077	0,046
57	3,168	0,020	0,168	3,143	0,013	0,143	2,987	0,282	0,013
58	3,012	0,030	0,012	2,981	0,001	0,019	3,013	0,298	0,013
59	2,944	0,024	0,056	2,958	0,009	0,042	2,923	0,112	0,077
60	2,911	0,057	0,089	3,009	0,006	0,009	3,033	0,194	0,033
61	2,756	0,017	0,244	2,737	0,036	0,263	2,941	0,314	0,059
62	2,952	0,033	0,048	2,984	0,002	0,016	2,907	0,269	0,093
63	2,822	0,038	0,178	3,260	0,023	0,260	3,260	0,023	0,260
64	3,065	0,011	0,065	3,032	0,016	0,032	3,057	0,068	0,057
65	2,851	0,010	0,149	2,806	0,056	0,194	2,963	0,231	0,037
66	3,087	0,018	0,087	3,070	0,001	0,070	3,126	0,257	0,126
67	3,035	0,010	0,035	3,072	0,008	0,072	3,072	0,083	0,072
68	2,914	0,019	0,086	2,876	0,058	0,124	2,965	0,244	0,035
69	2,691	0,006	0,309	2,579	0,015	0,421	2,924	0,360	0,076
70	2,965	0,005	0,035	2,865	0,025	0,135	2,842	0,049	0,158
71	3,010	0,001	0,010	3,017	0,003	0,017	3,018	0,011	0,018
72	2,931	0,348	0,069	3,265	0,015	0,265	3,216	0,683	0,216
73	3,091	0,001	0,091	3,223	0,004	0,223	3,216	0,088	0,216
74	3,000	0,001	0,000	3,004	0,001	0,004	3,004	0,001	0,004
75	2,802	0,252	0,198	2,606	0,109	0,394	2,824	0,338	0,176
76	3,006	0,039	0,006	3,059	0,019	0,059	2,961	0,139	0,039
77	3,451	0,038	0,451	3,416	0,004	0,416	3,451	0,042	0,451
78	3,256	0,037	0,256	3,202	0,091	0,202	3,263	0,208	0,263
79	3,108	0,045	0,108	3,186	0,020	0,186	3,077	0,163	0,077
80	3,219	0,002	0,219	3,211	0,010	0,211	3,256	0,309	0,256
81	2,966	0,018	0,034	2,944	0,004	0,056	2,958	0,118	0,042
82	3,000	0,003	0,000	2,975	0,002	0,025	3,000	0,054	0,000
83	2,916	0,003	0,084	2,916	0,003	0,084	2,929	0,222	0,071

Продолжение таблицы 2.7

v	IF&PA			Метод ОБ			Метод АБ		
	x*	**ε***	**ξ**	***x****	**ε***	**ξ**	***x****	**ε***	**ξ**
84	2,950	0,139	0,050	3,079	0,010	0,079	2,937	0,443	0,063
85	3,043	0,022	0,043	3,040	0,025	0,040	3,038	0,065	0,038
86	2,994	0,016	0,006	2,972	0,038	0,028	2,997	0,269	0,003
87	2,966	0,007	0,034	2,844	0,129	0,156	2,657	0,335	0,343
88	3,096	0,059	0,096	3,095	0,059	0,095	2,960	0,331	0,040
89	2,997	0,002	0,003	3,005	0,001	0,005	3,001	0,007	0,001
90	2,980	0,004	0,020	2,986	0,001	0,014	2,977	0,097	0,023
91	3,055	0,015	0,055	3,116	0,080	0,116	3,107	0,164	0,107
92	2,687	0,021	0,313	2,821	0,004	0,179	3,000	0,457	0,000
93	2,593	0,097	0,407	2,578	0,112	0,422	2,566	0,232	0,434
94	3,093	0,052	0,093	3,398	0,029	0,398	3,096	0,413	0,096
95	2,559	0,075	0,441	2,553	0,080	0,447	2,425	0,287	0,575
96	2,971	0,008	0,029	2,987	0,023	0,013	2,986	0,142	0,014
97	2,997	0,002	0,003	3,004	0,001	0,004	3,004	0,001	0,004
98	2,978	0,008	0,022	3,078	0,013	0,078	2,972	0,120	0,028
99	2,734	0,007	0,266	2,785	0,057	0,215	2,783	0,161	0,217
100	2,988	0,002	0,012	3,343	0,002	0,343	3,245	0,197	0,245

На рисунке 2.9 представлены кривые $\xi(v)$, полученные методами IF&PA и ОБ. Из рисунка 2.9 видно, что средняя ширина лепестка для метода IF&PA (кривые 3 и 4) составляет 0,09, а для метода ОБ (кривые 1 и 2) – 0,16. На рисунке 2.10 показаны графики $\xi(v)$, полученные методами IF&PA и АБ. Средняя ширина лепестка для метода АБ (кривые 1 и 2) составляет 0,13. Таким образом, из полученных экспериментальных данных следует, что метод IF&PA почти в два раза превосходит по степени ***робастности*** методы ОБ и АБ, основанные на самых робастных правилах голосования.

Из рисунков 2.9 и 2.10 видно, что средние расстояния, полученные методом IF&PA, составляют 0,10 для нормального и 0,19 для равномерного распределений. Средние расстояния, полученные для нормального и равномерного распределений, для метода ОБ составляют 0,12 и 0,28 соответственно, а для метода АБ – 0,11 и 0,24 соответственно. Полученные результаты свидетельствуют о большей (в 1,1-1,5 раза) точности метода IF&PA по сравнению с методами ОБ и АБ.

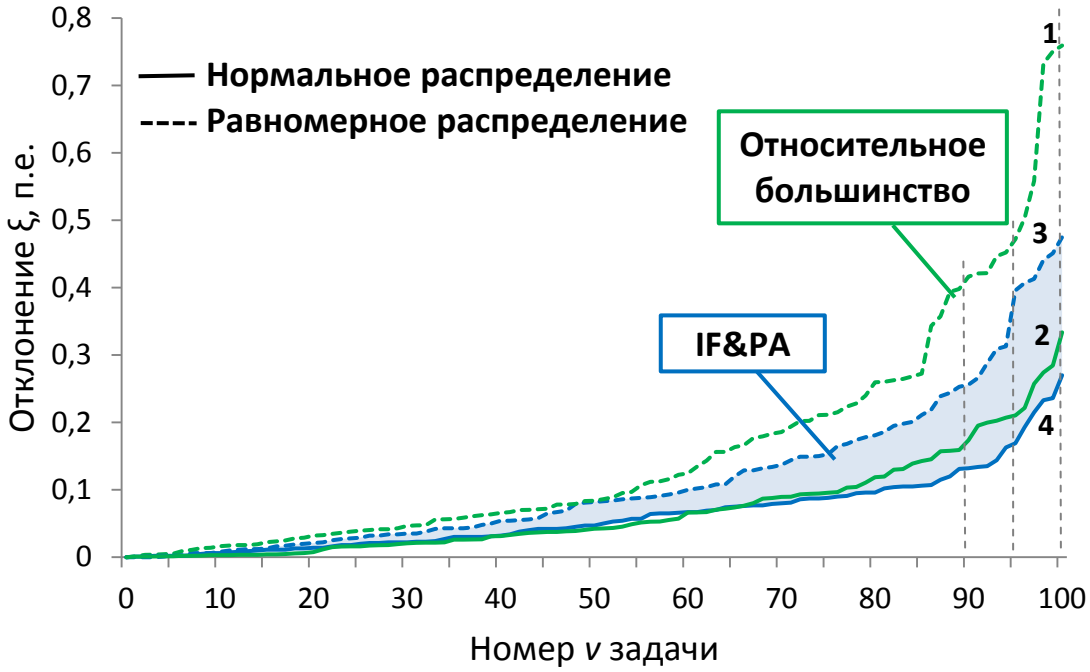


Рисунок 2.9 – Отклонения ξ, полученные методом IF&PA и методом ОБ для нормального (сплошная линия) и равномерного (пунктирная линия) распределений

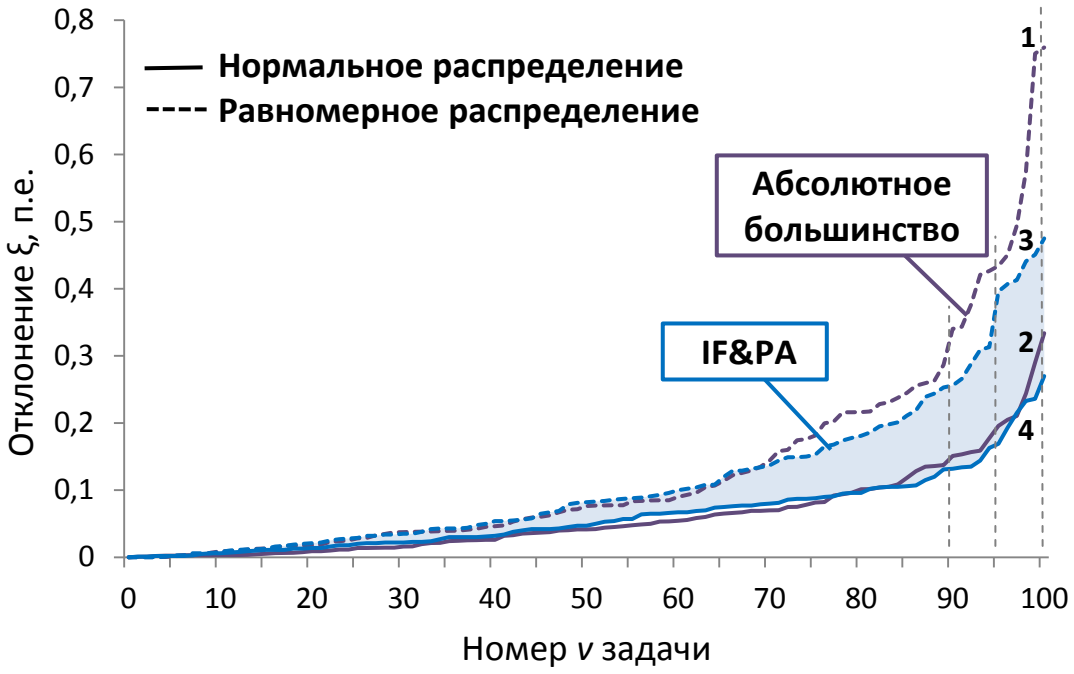


Рисунок 2.10 – Отклонения ξ, полученные методом IF&PA и методом АБ для нормального (сплошная линия) и равномерного (пунктирная линия) распределений

Достоверность оценивалась на основании данных, приведенных в таблицах 2.6 и 2.7, путем расчета оценок вероятностей $P(\xi \le \xi_{гр})$ по формуле

(2.35) для трех исследуемых методов. Значения $\xi_{гр}$ для вероятностей P = 0,90, P = 0,95 и P = 1 приведены в таблице 2.8.

Таблица 2.8 – Значения $\xi_{гр}$ для трех методов при значениях вероятностей P = 0,90, P = 0,95 и P = 1

Метод	**$\xi_{гр}$**					
	Нормальное распределение			**Равномерное распределение**		
	***P* = 0,90**	***P* = 0,95**	***P* = 1,00**	***P* = 0,90**	***P* = 0,95**	***P* = 1,00**
IF&PA	0,14	0,17	0,27	0,26	0,40	0,48
ОБ	0,18	0,22	0,34	0,42	0,48	0,76
АБ	0,16	0,20	0,34	0,35	0,44	0,76

Из данных таблицы 2.8 следует, что для всех рассмотренных вероятностей P наименьшие граничные отклонения $\xi_{гр}$ были получены методом IF&PA. В случае нормального распределения значения $\xi_{гр}$ для методов ОБ и АБ превышают $\xi_{гр}$ для метода IF&PA в среднем на 14-25 %, а в случае равномерного распределения – на 10-58 %.

Проанализируем значения ***неопределенностей*** ε* результатов комплексирования x*, полученные методом IF&PA по (2.23) и методами ОБ и АБ (см. шаги 10 и 11 Алгоритмов 3 и 4 соответственно) для нормального и равномерного распределений (рисунки 2.11 и 2.12 соответственно). Можно заметить, что значения ε* для методов IF&PA (кривая 3) и ОБ (кривая 2) находятся на одном уровне и не превышают 0,08 и 0,14 соответственно для нормального распределения. При равномерном распределении максимальные значения ε* равны 0,35 и 0,31 для IF&PA и ОБ соответственно. Неопределенность ε*, полученная методом АБ (кривая 1), характеризуется существенно бóльшими значениями, достигающими 0,33 и 0,68 для нормального и равномерного распределений соответственно. Таким образом, в ходе экспериментов результирующий интервал $I_г$ = [x* – ε*, x* + ε*] в среднем в 70 % случаев оказывался шире, чем некоторые исходные интервалы, т.е. использование метода АБ приводило к увеличению исходной неопределенности измерений.

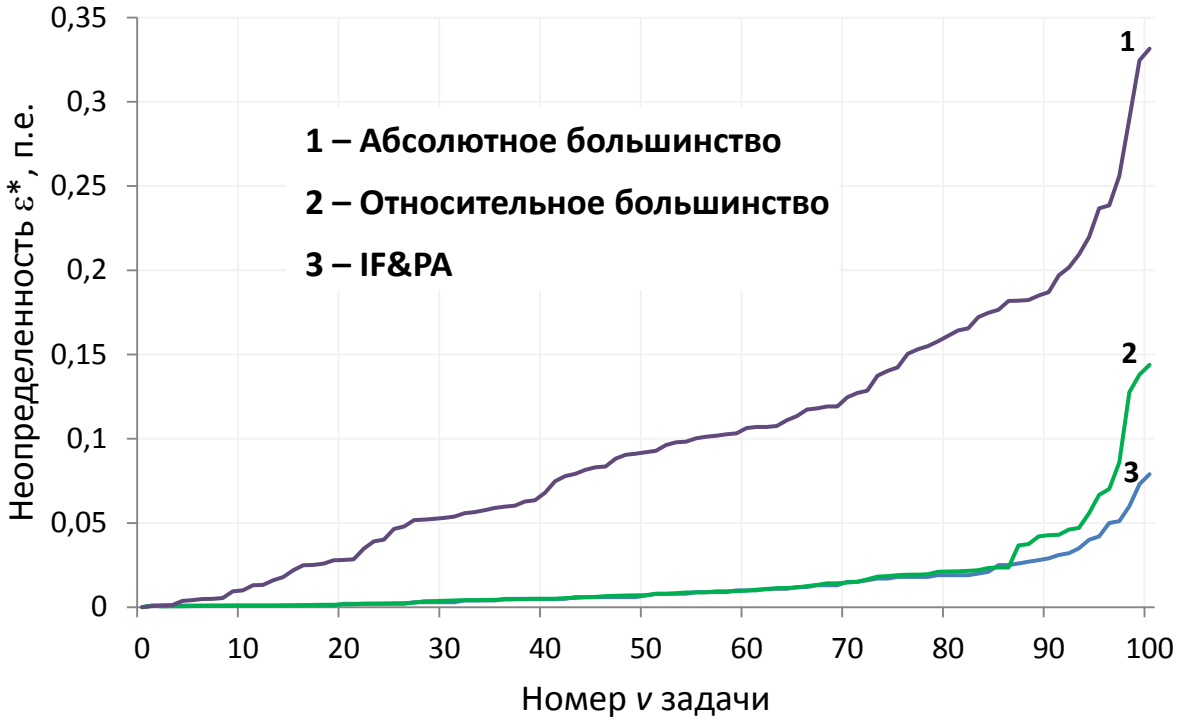


Рисунок 2.11 – Неопределенности ε*, полученные методами IF&PA, ОБ и АБ (нормальное распределение)

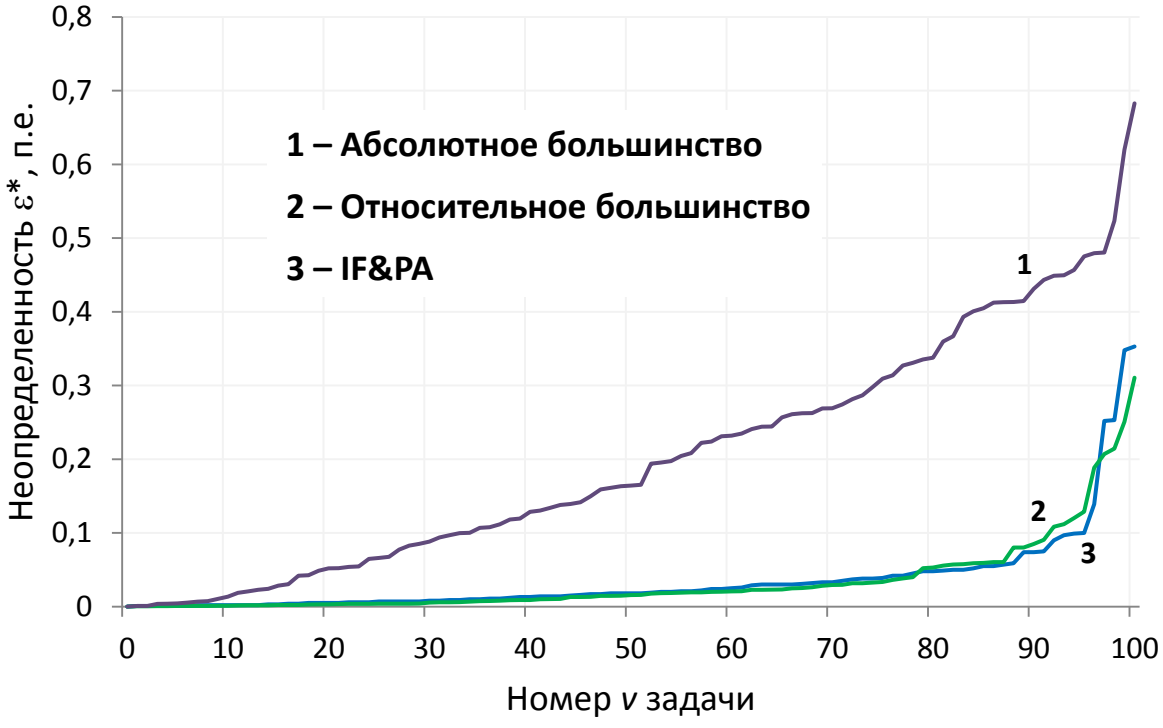


Рисунок 2.12 – Неопределенности ε*, полученные методами IF&PA, ОБ и АБ (равномерное распределение)

Следует отметить, что метод ОБ характеризуется высокой вероятностью появления парадокса (см. п. 1.4.5), заключающегося в получении одновременно

нескольких несогласованных результирующих интервалов. В ходе экспериментальных исследований это явление также имело место, и относительная частота появления парадокса в каждом прогоне составила в среднем 0,45. Использование метода ОБ, таким образом, было связано с проблемой выбора результирующего интервала из нескольких, причем не существует никаких объективных признаков, позволяющих считать один из интервалов предпочтительнее других. В данной работе в случае возникновения парадокса среди нескольких подходящих результирующих интервалов в качестве единственного результата I_r выбирался первый найденный интервал.

Таким образом, на основании результатов экспериментальных исследований можно утверждать, что разработанный метод IF&PA характеризуется повышенной точностью, робастностью, достоверностью и низкой результирующей неопределенностью.

Выводы к главе 2

1. Введено понятие инранжирования для определения ранжирований, наведенных интервалами. Показано, что мощности множеств инранжирований в зависимости от мощности n разбиений ДАЗ имеют вид последовательности, элементы которой являются треугольными числами.

2. Предложен метод комплексирования интервальных данных IF&PA, в котором результатом комплексирования x^* является наилучшее дискретное значение в ранжировании консенсуса, найденном для набора наведенных интервалами ранжирований дискретных значений.

3. Разработано программное обеспечение IntFusion для проведения численных экспериментальных исследований качества работы предложенного метода. Метод IF&PA сравнивался с методами одобрительного голосования, основанными на правилах относительного большинства (ОБ) и абсолютного большинства (АБ), которые признаны наиболее робастными из существующих правил голосования.

4. Суждение о качестве работы метода основывалось на оценке показателей точности, робастности и достоверности. Мерой точности и робастности служило отклонение ξ результата комплексирования от номинального значения. В качестве численного показателя достоверности использовались оценки вероятностей $P(\xi \le \xi_{гр})$ того, что отклонение ξ не превышает некоторое фиксированное значение $\xi_{гр}$. Результаты решения 100 задач каждым из исследуемых методов были представлены в виде кривой $\xi(v)$, где значения ξ упорядочены по возрастанию, v – номер индивидуальной задачи после упорядочения, $v = 1, \ldots, 100$.

5. Робастность метода определяется шириной "лепестка", образованного на графике кривыми $\xi(v)$ для нормального и равномерного распределений: чем уже лепесток, тем большей робастностью характеризуется метод. Ширина лепестка для метода IF&PA не превышает 0,23, а для методов ОБ и АБ – 0,42. Таким образом, IF&PA почти в два раза превосходит по степени робастности методы ОБ и АБ.

6. Точность метода характеризуется максимальным расстоянием между кривой $\xi(v)$ и осью абсцисс. Максимальные расстояния, полученные IF&PA, составляют 0,27 для нормального и 0,48 для равномерного распределений. Максимальные расстояния для методов ОБ и АБ составляют 0,34 и 0,76 соответственно. Полученные результаты свидетельствуют о большей (в 1,3-1,6 раза) точности метода IF&PA по сравнению с методами ОБ и АБ.

7. При оценке достоверности методов наименьшие граничные отклонения $\xi_{гр}$ для вероятностей $P = 0{,}90$, $P = 0{,}95$ и $P = 1$ были получены методом IF&PA. В случае нормального распределения значения $\xi_{гр}$ для методов ОБ и АБ превышают $\xi_{гр}$ для метода IF&PA на 14-25 %, а в случае равномерного распределения – на 10-58 % .

8. Метод IF&PA свободен от недостатков методов ОБ и АБ, а именно исключает возможность возникновения парадоксов и не приводит к увеличению исходной неопределенности измерений.

ГЛАВА 3
РАЗБИЕНИЕ ДИАПАЗОНА АКТУАЛЬНЫХ ЗНАЧЕНИЙ

В этой главе исследована проблема нахождения подходящего разбиения диапазона актуальных значений (ДАЗ) для метода IF&PA. Показаны свойства этого разбиения и представлен способ определения мощности разбиения, основанный на поправке Шеппарда для дисперсии дискретизированных значений ДАЗ.

3.1 Необходимость выбора мощности *n* разбиения ДАЗ

Представленный в главе 2 метод комплексирования IF&PA работает на множестве ранжирований элементов конечного множества, поэтому необходимо представлять исходные интервалы $\{I_k\}$ набором дискретных элементов a_i. Этап 1 метода предполагает переход от пространства непрерывных функций к дискретному пространству посредством формирования ДАЗ из исходных интервалов на вещественной оси, его разбиения на $n - 1$ подынтервалов и последующего представления элементами a_i дискретного множества A, $i = 1, \ldots, n$. Выбор числа n дискретных значений a_i, которое будем называть ***мощностью*** разбиения ДАЗ, оказывает существенное влияние на точность определения результата комплексирования x^*. Поскольку результатом x^* становится получивший наивысший ранг элемент a_i множества A, выбор подходящего числа n должен гарантировать необходимую и достаточную точность представления дискретных значений a_i множества A.

Разбиение ДАЗ, изначально содержащего бесконечное число вещественных значений, на конечное число подынтервалов представляет собой процесс ***дискретизации*** данных, который сопровождается неизбежной потерей информации. Очевидно, что точность представления дискретных значений a_i напрямую связана с нормой h разбиения (см. п. 2.2.1) и, как следует из формулы (2.14), с мощностью n разбиения.

Ранее проведенные численные экспериментальные исследования метода IF&PA [7] показали, что удовлетворительный результат комплексирования x^* в ряде случаев может быть получен уже при мощности разбиения $n = 4, 5$. В ходе подобных экспериментов суждение о близости полученного результата x^* к номинальному значению x_{nom} может быть легко сформировано, т.к. x_{nom} заранее известно (задано). Благодаря этому можно, перебирая различные значения n для некоторой индивидуальной задачи, найти такое n, при котором результат комплексирования x^* наиболее близок к x_{nom}.

При применении IF&PA в реальных условиях значение x_{nom} не может быть заранее известно, поэтому возникает необходимость разработки такого способа определения значения мощности n разбиения ДАЗ, который позволит с высокой степенью надежности получать результат комплексирования, близкий к номинальному значению.

Прежде, чем переходить к разработке способа выбора мощности n, рассмотрим некоторые свойства разбиения ДАЗ.

3.2 Свойства разбиения диапазона актуальных значений

3.2.1 Разрешающая способность

Процесс представления ДАЗ элементами дискретного множества $A = \{a_1, a_2, \ldots, a_n\}$ подробно описан в п. 2.2.1.

Будем использовать мощность n множества A в качестве нижнего индекса в его обозначении, т.е. $A_n = \{a_1, a_2, \ldots, a_n\}$.

Воспользуемся простым модельным примером: $a_1 = 2{,}00$; $a_n = 5{,}00$, т.е. ДАЗ имеет вид [2,00; 5,00]. Тогда, например, при мощности $n = 4$ множество A_4 будет иметь следующий вид:

$$A_4 = \{2{,}00; 3{,}00; 4{,}00; 5{,}00\}.$$

Разбиение ДАЗ имеет общие черты с группированием данных при построении гистограмм или округлением и, вообще говоря, представляет собой процесс дискретизации [114, 115]. При этом действительное число x на вещественной оси заменяется дискретным значением из ограниченного строго

упорядоченного множества $\{a_1, a_2, \ldots, a_n\}$. Тогда все значения x, лежащие в интервале $(a_i \pm 0{,}5h)$, соотносятся со значением a_i.

На рисунке 3.1 показан пример функции преобразования значений непрерывной величины x в дискретные значения a_i для $n = 7$. При разбиении ДАЗ на шесть подынтервалов значения x, находящиеся слева от числа 2,25 (не включая само значение 2,25), заменяют значением $a_1 = 2$, а любое x в диапазоне [2,25; 2,75) – значением $a_2 = 2{,}5$. Таким образом, норма h определяет ***разрешающую способность*** метода, т.е. минимальное возможное изменение значения x, имеющее место при переходе от ДАЗ к множеству A.

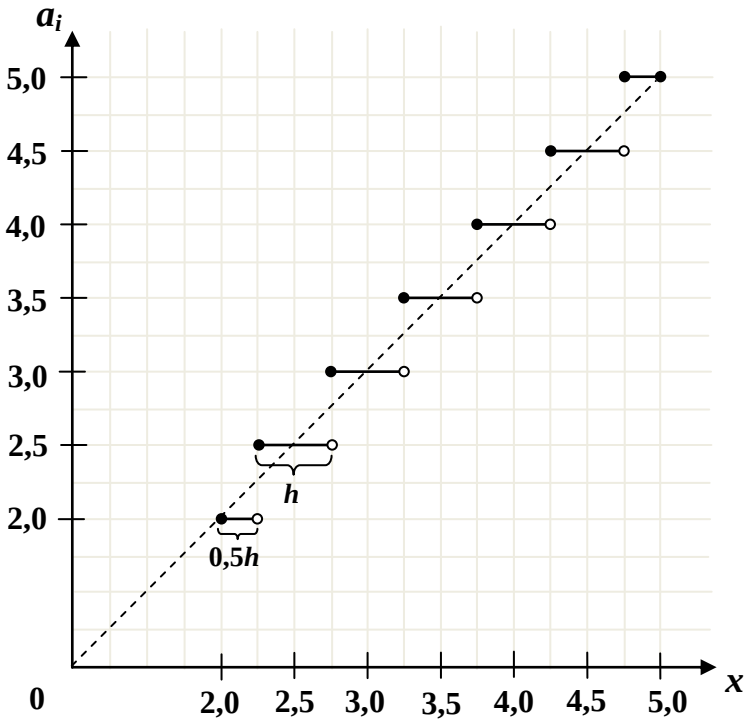


Рисунок 3.1 – Функция преобразования значений непрерывной величины x в дискретные значения a_i для $n = 7$

Для каждого конкретного разбиения ДАЗ можно определить границы, в которых лежит отношение $\frac{a_i}{x}$, характеризующее отклонение дискретного значения a_i от значения x непрерывной величины:

$$\frac{a_1}{a_1 + 0{,}5h} < \frac{a_i}{x} \le \frac{a_1 + h}{a_1 + 0{,}5h}. \qquad (3.1)$$

Например, для приведенного на рисунке 3.1 разбиения при $n = 7$ справедливо неравенство:

$$\frac{8}{9} < \frac{a_i}{x} \le \frac{10}{9}. \qquad (3.2)$$

Из выражения (3.1) следует, что максимально возможное отклонение a_i от x имеет вид:

$$\max \frac{a_i}{x} = \frac{h}{2(a_1 + 0{,}5h)}. \qquad (3.3)$$

Очевидно, что при улучшении разрешающей способности значение функции $\max(a_i / x)$ будет уменьшаться.

Из выражений (2.34) и (3.1) следует также взаимосвязь разрешающей способности h и точности ξ результата комплексирования (см. п. 2.4.1). После разбиения ДАЗ на $n - 1$ подынтервалов значение $x_{\text{ном}}$ будет находиться внутри одного из подынтервалов. При этом отклонение дискретного значения a_i, выбранного в качестве результата комплексирования, от значения $x_{\text{ном}}$ не превышает половины нормы разбиения h, т.е.

$$|a_i - x_{\text{ном}}| \le 0{,}5h. \qquad (3.4)$$

Тогда норма h является мерой точности при разбиении ДАЗ. При уменьшении h увеличивается разрешающая способность метода, что ведет к снижению возможной погрешности комплексирования.

3.2.2 Вложенность множеств A_n

Разобьем ДАЗ [2,00; 5,00] на число подынтервалов $n - 1$, где n последовательно принимает значения из ряда {2, 3, …, 20}. Множества A_n, полученные в результате разбиений при различных значениях мощности n, приведены в таблице 3.1. Из таблицы 3.1 видно, что дискретные значения a_i периодически повторяются при разных n.

На рисунке 3.2 показаны разбиения [A; Z], где значения A и Z могут быть любыми вещественными числами, для мощностей n = {3, 4, 5, …, 11}. Вместо конкретных числовых значений использованы символьные обозначения, что дает

более наглядное представление о периодическом характере появления значений a_i для разбиений при различных n.

Таблица 3.1 – Результаты разбиения ДАЗ [2,00; 5,00] для значений n={2, 3, …, 20}

a_i	A_n																		
	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}	A_{16}	A_{17}	A_{18}	A_{19}	A_{20}
a_1	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00
a_2	5,00	3,50	3,00	2,75	2,60	2,50	2,43	2,38	2,33	2,30	2,27	2,25	2,23	2,21	2,20	2,19	2,18	2,17	2,16
a_3	-	5,00	4,00	3,50	3,20	3,00	2,86	2,75	2,67	2,60	2,55	2,50	2,46	2,43	2,40	2,38	2,35	2,33	2,32
a_4	-	-	5,00	4,25	3,80	3,50	3,29	3,13	3,00	2,90	2,82	2,75	2,69	2,64	2,60	2,56	2,53	2,50	2,47
a_5	-	-	-	5,00	4,40	4,00	3,71	3,50	3,33	3,20	3,09	3,00	2,92	2,86	2,80	2,75	2,71	2,67	2,63
a_6	-	-	-	-	5,00	4,50	4,14	3,88	3,67	3,50	3,36	3,25	3,15	3,07	3,00	2,94	2,88	2,83	2,79
a_7	-	-	-	-	-	5,00	4,57	4,25	4,00	3,80	3,64	3,50	3,38	3,29	3,20	3,13	3,06	3,00	2,95
a_8	-	-	-	-	-	-	5,00	4,63	4,33	4,10	3,91	3,75	3,62	3,50	3,40	3,31	3,24	3,17	3,11
a_9	-	-	-	-	-	-	-	5,00	4,67	4,40	4,18	4,00	3,85	3,71	3,60	3,50	3,41	3,33	3,26
a_{10}	-	-	-	-	-	-	-	-	5,00	4,70	4,45	4,25	4,08	3,93	3,80	3,69	3,59	3,50	3,42
a_{11}	-	-	-	-	-	-	-	-	-	5,00	4,73	4,50	4,31	4,14	4,00	3,88	3,76	3,67	3,58
a_{12}	-	-	-	-	-	-	-	-	-	-	5,00	4,75	4,54	4,36	4,20	4,06	3,94	3,83	3,74
a_{13}	-	-	-	-	-	-	-	-	-	-	-	5,00	4,77	4,57	4,40	4,25	4,12	4,00	3,89
a_{14}	-	-	-	-	-	-	-	-	-	-	-	-	5,00	4,79	4,60	4,44	4,29	4,17	4,05
a_{15}	-	-	-	-	-	-	-	-	-	-	-	-	-	5,00	4,80	4,63	4,47	4,33	4,21
a_{16}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5,00	4,81	4,65	4,50	4,37
a_{17}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5,00	4,82	4,67	4,53
a_{18}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5,00	4,83	4,68
a_{19}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5,00	4,84
a_{20}	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5,00

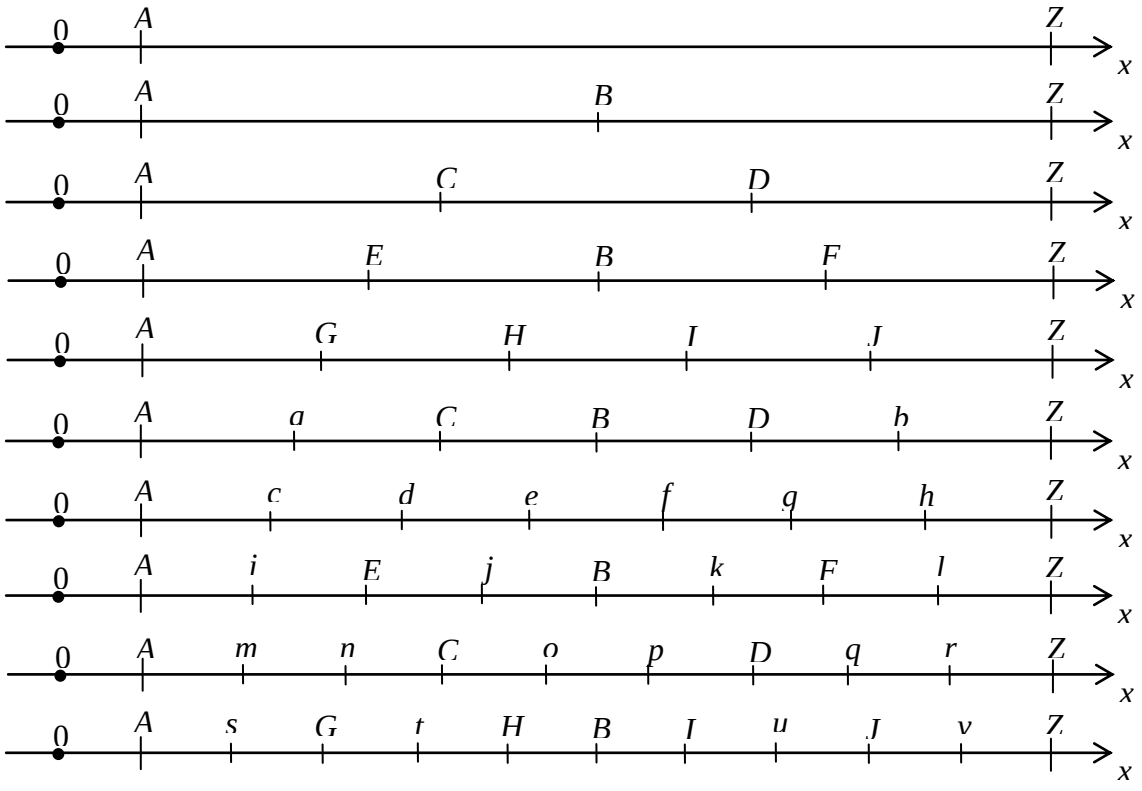


Рисунок 3.2 – Разбиение ДАЗ на n–1 подынтервалов при n = {3, 4, 5, …, 11}

В таблице 3.2 представлены данные о частоте появления значений a_i для разбиений мощности n = {2, 3, …, 20}. Заметим, что в таблицы не приведены значения, впервые появляющиеся после разбиения ДАЗ на 11 и более подынтервалов, т.к. частота появления этих значений равна единице.

Таблица 3.2 – Частота появления значений a_i для разбиений мощности n = {2, 3, …, 20}

Символьное обозначение значения a_i	Число появлений значения a_i	Число n, при котором появляется значение a_i	Значение a_i для [2,00; 5,00]
A	20	2,3,4, …19,20	2,00
Z	20	2,3,4, …19,20	5,00
B	9	3,5,7,9,11,13,15,17,19	3,50
C	6	4,7,10,13,16,19	3,00
D	6	4,7,10,13,16,19	4,00
E	4	5,9,13,17	2,75
F	4	5,9,13,17	4,25
G	3	6,11,16	2,60
H	3	6,11,16	3,20
I	3	6,11,16	3,80
J	3	6,11,16	4,40
a	3	7,13,19	2,50
b	3	7,13,19	4,50
c	2	8,15	2,43
d	2	8,15	2,86
e	2	8,15	3,29
f	2	8,15	3,71
g	2	8,15	4,14
h	2	8,15	4,57
i	2	9,17	2,38
j	2	9,17	3,13
k	2	9,17	3,88
l	2	9,17	4,63
m	2	10,19	2,33
n	2	10,19	2,67
o	2	10,19	3,33
p	2	10,19	3,67
q	2	10,19	4,33
r	2	10,19	4,67
s	1	11	2,30
t	1	11	2,90

Продолжение таблицы 3.2

Символьное обозначение значения a_i	Число появлений значения a_i	Число n, при котором появляется значение a_i	Значение a_i для [2,00; 5,00]
u	1	11	4,10
v	1	11	4,70
…	1	…	…

Отметим, что значения A и Z представляют границы ДАЗ и являются элементами всех множеств A_n. Из таблицы 3.2 видно, что значение B, полученное при первом разбиении на 2 подынтервала, встречается в последующих множествах наиболее часто. Следующими по частоте появления являются значения C и D, полученные при разбиении на 3 подынтервала. Таким образом, значения, появляющиеся в первых разбиениях при малых n, характеризуются наибольшей частотой появления в последующих разбиениях.

Из данных таблиц 3.1, 3.2 и рисунка 3.2 следует, что некоторые множества полностью включают в себя другие множества с меньшим индексом. Будем называть это свойство ***вложенностью множеств*** $\boldsymbol{A_n}$.

В таблице 3.3 представлены множества, вложенные в A_n для значений $n = \{2, 3, \ldots, 20\}$.

Запишем последовательности множеств, включающих A_n для различных значений мощности $n = \{2, 3, \ldots, 20\}$:

$$A_3 \subset A_3, A_5, A_7, A_9, A_{11}, A_{13}, A_{15}, A_{17}, A_{19}$$
$$A_4 \subset A_4, A_7, A_{10}, A_{13}, A_{16}, A_{19}$$
$$A_5 \subset A_5, A_9, A_{13}, A_{17}$$
$$A_6 \subset A_6, A_{11}, A_{16} \qquad (3.5)$$
$$A_7 \subset A_7, A_{13}, A_{19}$$
$$A_8 \subset A_8, A_{15}$$
$$A_9 \subset A_9, A_{17}$$
$$A_{10} \subset A_{10}, A_{19}$$

Ближайшее множество, которому принадлежит множество A_n, определяется с помощью формулы:

$$A_n \subset A_{2n-1}. \qquad (3.6)$$

Таблица 3.3 – Множества, вложенные в A_n для n = {2, 3, …, 20}

n	**Множества, вложенные в A_n**
2	A_2
3	A_2
4	A_2
5	A_2, A_3
6	A_2
7	A_2, A_3, A_4
8	A_2
9	A_2, A_3, A_5
10	A_2, A_4
11	A_2, A_3, A_6
12	A_2
13	A_2, A_3, A_4, A_5, A_7
14	A_2
15	A_2, A_3, A_8
17	A_2, A_3, A_5, A_9
18	A_2
19	A_2, A_3, A_4, A_7, A_{10}
20	A_2

Из формулы (3.6) видно, что ближайшее множество, включающее в себя A_n, всегда имеет нечетный порядковый номер. Тогда для всех множеств с нечетным порядковым номером справедливо обратное: для каждого множества A_n можно найти ближайшее принадлежащее ему множество:

$$A_{(n+1)/2} \subset A_n. \tag{3.7}$$

Пусть $(n+1)/2 = q$. В соответствии с формулой (3.7), множество $A_n = \{ a_1^n, a_2^n, \ldots, a_n^n \}$ включает в себя множество $A_q = \{ a_1^q, a_2^q, \ldots, a_q^q \}$, т.е. состоит из q элементов множества A_q и $q - 1$ элементов дополнения множества A_q до множества A_n:

$$A_n = A_q \cup \bar{A}_q. \tag{3.8}$$

Тогда легко определить элементы множества A_n, воспользовавшись выражением:

$$A_n = \{ a_1^q,\ a_2^q,\ \ldots,\ a_q^q,\ \frac{a_1^q + a_2^q}{2}, \frac{a_2^q + a_3^q}{2},\ \ldots,\ \frac{a_{q-1}^q + a_q^q}{2} \}. \qquad (3.9)$$

Таким образом, множество A_n характеризуется в два раза меньшей нормой h разбиения, чем вложенное в него множество A_q, т.к. включает в себя не только элементы A_q, но также и $q - 1$ значений, занимающих промежуточное положение между элементами A_q. Например, множество A_7 помимо дискретных значений {2,00; 3,00; 4,00; 5,00} множества A_4 включает также значения {2,50; 3,50; 4,50}, не принадлежащие A_4.

3.3 Определение мощности разбиения ДАЗ

Непрерывная величина x может быть интерпретирована как реализация случайной переменной X и описана оценками математического ожидания μ и среднеквадратического отклонения (СКО) σ. Для определения мощности n разбиения ДАЗ будем использовать так называемую поправку Шеппарда, вводимую для оценки дисперсии σ_d^2 дискретизированных данных [93, 113]. В нашем случае, поправка Шеппарда определяется формулой:

$$\sigma_d^2 = \sigma^2 + \frac{h^2}{12}. \qquad (3.10)$$

где σ и σ_d – среднеквадратические отклонения (СКО) непрерывных (до разбиения) и дискретных (после разбиения ДАЗ) значений соответственно.

В соответствии с подходом, представленным в [113], из выражения (3.10) получаем относительное СКО:

$$\frac{\sigma_d}{\sigma} = \sqrt{1 + \frac{h^2}{12\sigma^2}}. \qquad (3.11)$$

Пусть w – допускаемое различие (разность) между значениями σ_d и σ. Значение w может задаваться или выбираться на основании требований, предъявляемых к качеству (точности) комплексирования интервалов.

Задавая w в относительных единицах, имеем:

$$\sigma_d \le (1 + w)\sigma. \qquad (3.12)$$

В соответствии с (3.11) и (3.12) можно записать:

$$\frac{\sigma_d}{\sigma} = \sqrt{1 + \frac{h^2}{12\sigma^2}} \le (1 + w), \tag{3.13}$$

откуда после ряда простых преобразований получаем формулу для вычисления нормы разбиения *h*:

$$h \le \sigma\sqrt{24w + 12w^2}\ . \tag{3.14}$$

С учетом формулы (2.14) мощность *n* разбиения ДАЗ определяется выражением:

$$n = \left\lceil \frac{a_n - a_1}{\sigma\sqrt{24w + 12w^2}} \right\rceil + 1. \tag{3.15}$$

Для оценки параметра σ могут быть использованы рекомендации, приведенные в ГОСТ Р 54500.3-2011/Руководство ИСО/МЭК 98-3:2008 Неопределенность измерения – Часть 3: Руководство по выражению неопределенности измерения [2].

3.4 Численные экспериментальные исследования разбиения ДАЗ

Предложенный подход для определения мощности разбиения ДАЗ был экспериментально апробирован с помощью специально разработанного программного обеспечения IntFusion, описанного в п. 2.3. Основными целями проведения численных экспериментальных исследований являлись:

- выбор значения допускаемого различия *w*;
- проверка практической применимости предложенного способа для расчета мощности *n* разбиения ДАЗ.

Было проведено 8 экспериментальных прогонов по сто индивидуальных задач, в каждом из которых были сгенерированы данные, распределенные по нормальному закону. Во всех индивидуальных задачах данные генерировались с различным СКО σ, определяемым в соответствии с параметрами генерации (см. шаги 1-3 Алгоритма 1), которые варьировались для каждого прогона в различных комбинациях. В качестве демонстрации в этом параграфе приведены результаты

трех, произвольно выбранных из восьми, экспериментов по 100 задач. Задаваемые параметры генерации данных для трех экспериментов отражены в таблице 3.4.

Таблица 3.4 – Параметры генерации данных

Эксперимент	x_{nom}**, п.е.**	Δx**, п.е.**	m	R_x	R_ε	σ**, п.е.**
1	3	1,5	15	1	1	0,1-0,5
2	3	1,5	15	1	2	0,1-0,7
3	3	1	15	1	2	0,1-0,3

3.4.1 Выбор допускаемого различия *w*

Для осуществления обоснованного выбора допускаемого различия w сгенерированные данные были обработаны методом IF&PA, при этом мощность n определялась по формуле (3.15). В качестве параметра метода были использованы значения из ряда $w = \{0{,}004; 0{,}005; \ldots 0{,}04\}$. Для всех w определялись результаты комплексирования x^* и отклонения ξ значения x^* от номинального x_{nom}, которые были использованы для получения по формуле (2.35) оценок вероятностей $P(\xi \le \xi_{гр})$ того, что отклонение ξ не превышает некоторое граничное значение $\xi_{гр}$. Значения $\xi_{гр}$ для вероятностей $P = 0{,}90$, $P = 0{,}95$ и $P = 1$ представлены в таблице 3.5. Заметим, что полученные значения n не показаны в таблице 3.5, т.к. при фиксированном w они принимали различные значения в зависимости от конкретного σ соответствующей индивидуальной задачи.

На рисунке 3.3 представлен график отклонений ξ, определенных при различных w для Эксперимента 1. Поскольку кривые отклонений для всех экспериментов демонстрируют схожее поведение, приведен график лишь для одного эксперимента. Для лучшей наглядности на рисунке показаны кривые только для значений из ряда $w = \{0{,}004; 0{,}006; 0{,}01; 0{,}015; 0{,}03\}$.

Из рисунка 3.3 и таблицы 3.5 видно, что кривые отклонений при разных значениях w отличаются друг от друга не более чем на (12–26) %. Это значит, что минимально возможное для данного метода допускаемое различие $w = 0{,}004 = 0{,}4$ % может быть рекомендовано к использованию при расчете мощности n разбиения ДАЗ. Тогда формулы (3.14) и (3.15) могут быть упрощены до следующих эмпирических выражений:

$$h \le 0{,}31\sigma, \tag{3.16}$$

$$n = \left\lceil \frac{a_n - a_1}{0{,}31\sigma} \right\rceil + 1. \tag{3.17}$$

Таблица 3.5 – Значения граничных отклонений $\xi_{гр}$ для вероятностей P = 0,90, 0,95 и 1,00 при различных значениях w

w	$\xi_{гр}$								
	Эксперимент 1			**Эксперимент 2**			**Эксперимент 3**		
	***P*=0,90**	***P*=0,95**	***P*=1,00**	***P*=0,90**	***P*=0,95**	***P*=1,00**	***P*=0,90**	***P*=0,95**	***P*=1,00**
0,004	0,24	**0,28**	0,40	0,25	**0,28**	**0,32**	**0,16**	**0,20**	**0,34**
0,005	0,23	0,34	0,46	**0,22**	0,34	0,45	0,17	**0,20**	0,35
0,006	0,23	0,32	0,41	0,28	0,34	0,47	**0,16**	**0,20**	0,35
0,008	**0,21**[1]	0,34	0,52	0,24	0,36	0,54	0,20	0,24	0,42
0,10	0,23	**0,28**	**0,36**	0,27	0,38	0,48	0,17	0,24	0,42
0,13	0,24	0,32	0,46	0,25	0,33	0,64	0,18	0,26	0,37
0,15	0,22	0,35	0,51	0,29	0,33	0,64	0,20	0,26	**0,34**
0,2	0,26	0,33	0,52	0,29	0,38	0,51	0,17	**0,20**	0,48
0,3	0,23	0,29	0,45	0,28	0,35	0,49	0,19	0,30	0,48
0,4	0,29	0,36	0,64	0,29	0,37	0,69	0,22	0,28	0,45

[1]Минимальные значения $\xi_{гр}$ для заданной вероятности обозначены полужирным шрифтом

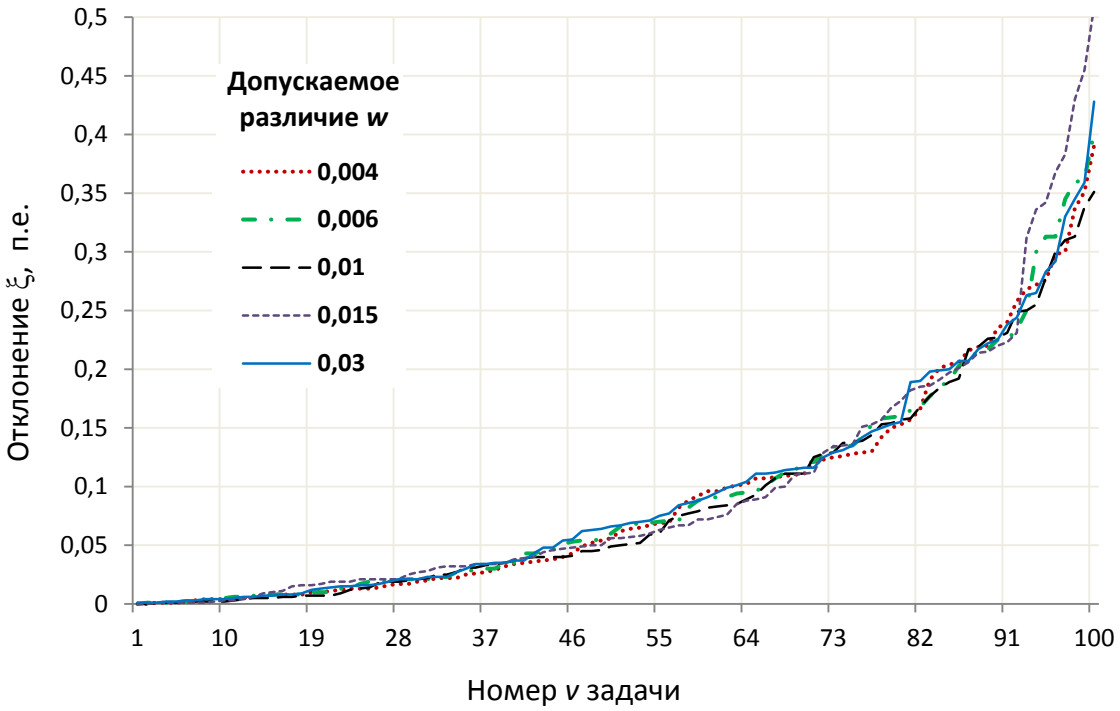


Рисунок 3.3 – График отклонений ξ при различных значениях допускаемого различия w (Эксперимент 1)

3.4.2 Проверка практической применимости способа расчета мощности разбиения ДАЗ

Те же сгенерированные данные были использованы для проверки применения предложенного способа расчета мощности *n* для разбиения ДАЗ при применении IF&PA в реальных практических задачах, когда номинальное значение величины неизвестно. Вообще говоря, наилучший способ определения *n* состоит в *последовательном выборе* значения *n* из ограниченного ряда значений, нахождении *x** и фиксации *n*, при котором $\xi = |x_{\text{nom}} - x^*|$ минимально. Он не применим при неизвестном x_{nom}, но позволяет проверить применимость формулы (3.17). Проверка состояла в определении значений $\xi_{\text{гр}}$ (для вероятностей $P = 0{,}90$, $P = 0{,}95$ и $P = 1$) как при последовательном выборе, так и расчете *n* для 8 экспериментальных прогонов.

При последовательном выборе мощность *n* принимала значения из ряда {4, 5, 6,…15}. Поскольку лежащая в основе метода IF&PA задача о ранжировании Кемени относится к классу NP-полных и требует экспоненциального времени для решения (см. п. 2.1.2), ряд для выбора *n* был ограничен значением 15, т.к. оно позволяет находить результат комплексирования *x** за приемлемое (до 2 с) время. Минимальное значение ряда *n* было принято равным 4 по причине нерациональности разбиения ДАЗ менее чем на три подынтервала. Значения $\xi_{\text{гр}}$, полученные при различных *n*, представлены в таблице 3.6.

Как видно из данных таблицы 3.6, наименьшие значения $\xi_{\text{гр}}$ для различных случаев наблюдаются при разном значении *n*. Ясно, что случайный выбор *n* может привести к непредсказуемому и неточному результату комплексирования *x**.

Те же данные для трех экспериментов были обработаны методом IF&PA, при этом для расчета мощности *n* разбиения использовалась формула (3.17). ***Целью*** применения предложенного способа расчета мощности *n* являлось определение такого числа *n*, при котором с заданной вероятностью *P* достигается наименьшее граничное отклонение $\xi_{\text{гр}}$ результата комплексирования *x** от номинального значения x_{nom}. Под наименьшим граничным отклонением $\xi_{\text{гр}}$

понимается минимальное значение, получаемое с вероятностью P для разбиений ДАЗ мощности n, где $n = \{4, 5, 6,\ldots 15\}$.

Таблица 3.6 – Значения граничных отклонений $\xi_{гр}$ для вероятностей P = 0,90, 0,95 и 1,00 при различных значениях n

n	**$\xi_{гр}$**								
	Эксперимент 1			**Эксперимент 2**			**Эксперимент 3**		
	***P*=0,90**	***P*=0,95**	***P*=1,00**	***P*=0,90**	***P*=0,95**	***P*=1,00**	***P*=0,90**	***P*=0,95**	***P*=1,00**
4	0,39	0,48	0,52	0,39	0,48	0,67	0,25	0,31	0,41
5	0,24	0,41	0,74	0,24	0,41	0,74	0,24	0,28	0,44
6	0,31	0,37	0,43	0,31	0,37	0,49	0,20	0,28	0,48
7	0,28	0,40	0,52	0,28	0,40	0,50	0,18	0,24	0,41
8	0,30	0,35	0,51	0,30	0,35	0,64	0,20	0,26	0,37
9	0,23	0,29	**0,34**	0,23	0,29	0,54	0,17	0,24	0,42
10	**0,21**	**0,28**	0,52	0,28	0,38	0,51	0,17	0,24	0,35
11	0,26	0,35	0,47	0,26	0,35	0,47	0,19	0,23	0,39
12	0,25	0,40	0,45	0,25	0,40	0,59	**0,16**	0,22	**0,33**
13	0,22	0,32	0,54	**0,22**	0,32	0,51	0,17	0,24	0,36
14	0,25	0,29	0,44	0,25	**0,28**	0,56	**0,16**	**0,19**	0,34
15	0,23	0,29	0,43	0,23	0,29	**0,37**	0,18	0,26	**0,33**

Пример рассчитанных в соответствии с (3.17) значений мощности n и соответствующих отклонений ξ при допускаемом различии w = 0,004 для трех экспериментов из 100 задач приведены в таблице 3.7.

Таблица 3.7 – Значения мощности n и отклонения ξ при w = 0,004 для 100 задач в трех экспериментах

v	**Эксп. 1**		**Эксп. 2**		**Эксп. 3**		**v**	**Эксп. 1**		**Эксп. 2**		**Эксп. 3**	
	n	**ξ**	***n***	**ξ**	***n***	**ξ**		***n***	**ξ**	***n***	**ξ**	***n***	**ξ**
1	14	0,128	15	0,016	13	0,000	51	15	0,012	14	0,005	15	0,027
2	14	0,056	14	0,033	13	0,014	52	14	0,013	14	0,271	14	0,056
3	13	0,126	15	0,116	13	0,136	53	14	0,022	16	0,302	15	0,021
4	16	0,001	13	0,045	15	0,021	54	16	0,037	15	0,034	14	0,046
5	15	0,034	13	0,039	15	0,023	55	14	0,276	14	0,012	14	0,047
6	13	0,108	15	0,137	14	0,026	56	14	0,130	16	0,140	15	0,130
7	14	0,022	12	0,224	14	0,041	57	15	0,272	13	0,074	15	0,062
8	15	0,082	16	0,080	15	0,040	58	16	0,017	13	0,070	14	0,005
9	14	0,069	13	0,158	14	0,071	59	14	0,204	14	0,121	15	0,037
10	14	0,026	14	0,003	14	0,024	60	13	0,191	15	0,282	15	0,035
11	15	0,040	14	0,261	14	0,016	61	15	0,096	13	0,075	14	0,003

Продолжение таблицы 3.7

v	Эксп. 1		Эксп. 2		Эксп. 3		*v*	Эксп. 1		Эксп. 2		Эксп. 3	
	n	ξ	***n***	ξ	***n***	ξ		***n***	ξ	***n***	ξ	***n***	ξ
12	14	0,109	15	0,015	15	0,031	62	14	0,107	16	0,022	13	0,043
13	14	0,005	13	0,109	14	0,004	63	15	0,004	15	0,077	16	0,078
14	14	0,142	14	0,036	14	0,173	64	14	0,004	14	0,228	13	0,027
15	14	0,110	14	0,067	15	0,010	65	15	0,016	16	0,039	14	0,014
16	14	0,013	16	0,000	15	0,006	66	14	0,216	13	0,039	15	0,020
17	15	0,129	13	0,058	15	0,003	67	15	0,043	14	0,071	13	0,007
18	13	0,092	14	0,053	14	0,151	68	15	0,000	15	0,214	14	0,070
19	16	0,235	15	0,057	14	0,002	69	15	0,101	14	0,047	15	0,035
20	14	0,391	14	0,042	14	0,137	70	14	0,111	13	0,156	13	0,009
21	15	0,011	16	0,131	16	0,028	71	14	0,010	14	0,248	16	0,150
22	13	0,219	13	0,081	14	0,030	72	16	0,052	14	0,243	15	0,081
23	14	0,125	13	0,067	14	0,041	73	15	0,062	14	0,257	16	0,014
24	14	0,201	14	0,010	15	0,061	74	14	0,351	13	0,008	16	0,077
25	14	0,069	14	0,002	15	0,029	75	16	0,008	14	0,307	16	0,101
26	13	0,220	15	0,170	14	0,186	76	14	0,064	13	0,071	15	0,006
27	16	0,166	14	0,022	16	0,063	77	15	0,022	15	0,001	15	0,077
28	14	0,298	15	0,114	15	0,029	78	15	0,008	13	0,009	15	0,272
29	14	0,002	13	0,102	13	0,065	79	14	0,099	15	0,002	15	0,131
30	15	0,000	15	0,015	14	0,081	80	14	0,096	13	0,015	14	0,338
31	15	0,032	15	0,007	15	0,016	81	15	0,014	13	0,128	14	0,121
32	14	0,102	14	0,285	14	0,011	82	13	0,258	14	0,104	15	0,001
33	14	0,240	13	0,060	13	0,237	83	13	0,121	14	0,016	14	0,003
34	14	0,013	14	0,234	15	0,123	84	14	0,019	15	0,101	15	0,014
35	14	0,004	14	0,001	15	0,080	85	15	0,300	13	0,316	13	0,117
36	14	0,157	15	0,105	14	0,006	86	14	0,006	14	0,025	15	0,018
37	14	0,067	16	0,045	14	0,042	87	13	0,007	15	0,033	15	0,063
38	14	0,010	15	0,007	13	0,128	88	15	0,008	14	0,086	13	0,156
39	14	0,001	15	0,091	15	0,074	89	13	0,029	16	0,000	15	0,032
40	13	0,004	14	0,130	15	0,098	90	14	0,027	13	0,009	14	0,113
41	13	0,001	15	0,117	15	0,111	91	15	0,206	15	0,099	16	0,062
42	14	0,054	15	0,046	15	0,199	92	14	0,337	15	0,016	14	0,060
43	15	0,150	14	0,018	13	0,206	93	13	0,036	14	0,071	13	0,022
44	15	0,107	15	0,196	13	0,091	94	14	0,049	16	0,008	16	0,141
45	14	0,008	13	0,018	13	0,039	95	14	0,088	13	0,102	13	0,010
46	14	0,123	15	0,206	13	0,258	96	14	0,268	14	0,258	13	0,019
47	15	0,017	14	0,008	16	0,054	97	14	0,020	16	0,093	14	0,034
48	15	0,035	15	0,014	14	0,045	98	15	0,065	14	0,024	14	0,012
49	15	0,038	16	0,038	14	0,041	99	15	0,003	15	0,129	15	0,178
50	13	0,025	15	0,010	16	0,067	100	14	0,153	15	0,032	15	0,090

Для принятия решения о возможности использования предложенного способа сравним наименьшие значения граничных отклонений $\xi_{гр}$, полученные с помощью последовательного выбора n (таблица 3.5) и посредством расчета n по формуле (3.17) при $w = 0{,}004$ (таблица 3.6). Минимальные значения $\xi_{гр}$, взятые из таблиц 3.5 и 3.6, сведены в таблицу 3.8 для большего удобства анализа и наглядности.

Таблица 3.8 – Минимальные значения $\xi_{гр}$, полученные при последовательном выборе n и при расчете n по формуле (3.17)

Эксперимент	**$\xi_{гр}$**					
	Последовательный выбор *n*			**Расчет *n* по формуле (3.17)**		
	***P* = 0,90**	***P*=0,95**	***P* = 1,00**	***P* = 0,90**	***P* = 0,95**	***P* = 1,00**
1	0,21	**0,28** (n = 10)	0,34	0,24	**0,28** (w = 0,004)	0,40
2	0,22	**0,28** (n = 14)	0,37	0,25	**0,28** (w = 0,004)	0,32
3	0,16	**0,19** (n = 14)	0,33	0,16	**0,20** (w = 0,004)	0,34

Из данных таблицы 3.8 следует, что разбиение ДАЗ в соответствии с предложенным способом расчета мощности n позволило получить наименьшие граничные значения отклонений $\xi_{гр}$ при значениях параметров $P = 0{,}95$ и $w = 0{,}004$. Расхождение между минимальными значениями $\xi_{гр}$, полученными для двух способов определения n не превышает 0,01. Таким образом, предложенный способ расчета мощности n разбиения ДАЗ в соответствии с выражениями (3.15) и (3.17) может быть рекомендован к использованию при комплексировании интервальных данных методом IF&PA.

Выводы к главе 3

1. Предложен способ расчета мощности n разбиения ДАЗ, основанный на поправке Шеппарда для дисперсии дискретизированных данных.

2. Проведены численные экспериментальные исследования, включающие 8 экспериментальных прогонов по сто индивидуальных задач, в каждом из которых

были сгенерированы данные, распределенные по нормальному закону. Во всех индивидуальных задачах данные генерировались с различным СКО, определяемым в соответствии с параметрами генерации.

3. Для обоснованного выбора допускаемого различия *w* сгенерированные данные были обработаны методом IF&PA, при этом мощность *n* рассчитывалась предложенным способом. Для значений *w* = {0,004; 0,005; …; 0,04} определялись результаты комплексирования *x** и значения $\xi_{гр}$. Наименьшие значения $\xi_{гр}$ были получены для *w* = 0,004, которое на этом основании было рекомендовано к использованию при расчете мощности *n* разбиения ДАЗ.

4. Сгенерированные данные были использованы для проверки применимости способа расчета *n* при применении IF&PA в реальных практических задачах. Проверка состояла в определении значений $\xi_{гр}$ (для вероятностей *P* = 0,90, *P* = 0,95 и *P* = 1) как при последовательном выборе значения *n* из ряда {4, 5, 6,…15}, так и расчете *n* предложенным способом. Доказано, что разбиение ДАЗ в соответствии с предложенным способом расчета мощности *n* позволило получить наименьшие граничные значения отклонений $\xi_{гр}$ при *P* = 0,95 и *w* = 0,004. Разница между минимальными значениями $\xi_{гр}$, полученными для двух способов определения *n*, не превышает 0,01.

ГЛАВА 4
КОМПЛЕКСИРОВАНИЕ ДАННЫХ В БЕСПРОВОДНЫХ СЕНСОРНЫХ СЕТЯХ ДЛЯ ЭКОЛОГИЧЕСКОГО МОНИТОРИНГА

В этой главе рассмотрена возможность повышения точности мультисенсоров беспроводной сенсорной сети (БСС) для экологического мониторинга на основе метода комплексирования интервальных данных IF&PA. Предложен робастный алгоритм повышения точности результата измерения мультисенсоров БСС и представлены результаты верификации предложенного алгоритма на входных синтетических данных. Приведены результаты обработки данных реальных БСС алгоритмом повышения точности. Предложен алгоритм выбора подмножества активных узлов в кластере БСС для снижения энергопотребления и проведены численные экспериментальные исследования его работоспособности.

4.1 Беспроводные сенсорные сети

Беспроводная сенсорная сеть представляет собой распределенную, самоорганизующуюся систему сбора, обработки и передачи информации, состоящую из автономных, не требующих специальной установки и обслуживания, устройств. Такие устройства, называемые ***узлами***, снабжены сенсорами – первичными измерительными преобразователями, которые измеряют параметры различных физических полей, сред и объектов в подлежащих мониторингу точках исследуемой области. Разнородные измерительные данные передаются по радиоканалу от узла к узлу на центральный узел (ЦУ) сети для обработки и принятия решений.

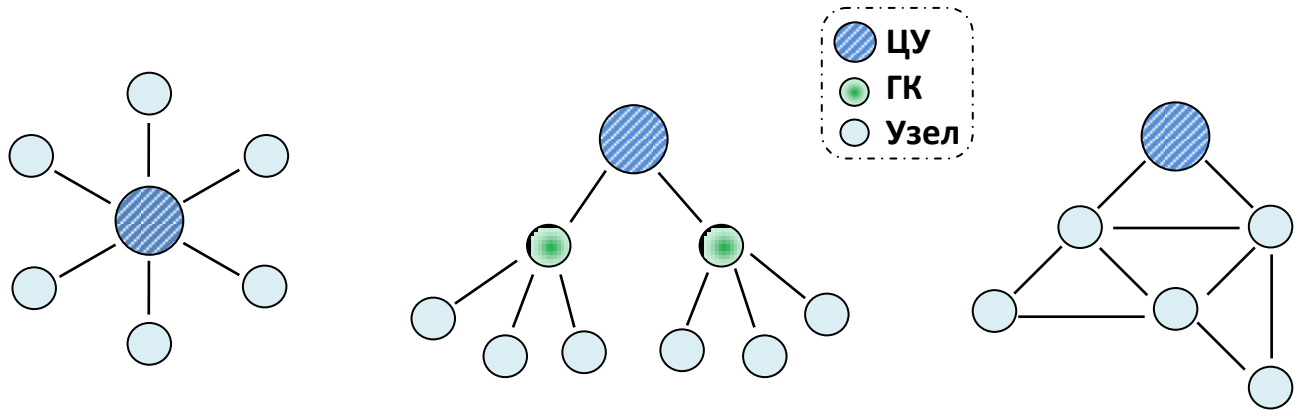
Узел БСС включает четыре основных модуля [66]. Модуль питания обеспечивает автономное питание узла в течение длительного времени. Модуль сбора данных содержит ***мультисенсор*** – набор сенсоров, преобразующих физические величины (температуру, влажность, освещенность и т.п.) в электрические сигналы. Вычислительный модуль (микроконтроллер и память)

управляет функционированием узла, в частности, принимает сигналы от сенсоров, обрабатывает и хранит измерительные данные. Модуль беспроводной передачи данных реализует обмен информацией между узлами сети по радиоканалу. Дополнительно в состав узлов также часто включается модуль определения местоположения (GPS).

В общем случае БСС состоит из определенного количества узлов, которые имеют автономное питание, и ЦУ с питанием от сети. Маршрут данных при передаче от узлов к ЦУ зависит от используемой топологии сети. Наиболее широкое применение получили следующие топологии, показанные на рисунке 4.1:

- звезда;
- кластерное дерево;
- ячеистая сеть.

В сети с топологией «звезда» узлы обычно находятся на расстоянии одного «шага» (single-hop) от ЦУ, в то время как в сетях с топологиями «кластерное дерево» и «ячеистая сеть» чаще реализована многошаговая (multi-hop) передача данных. Достоинства и недостатки трех представленных топологий приведены в таблице 4.1.



а) б) в)

Рисунок 4.1 – Виды топологий БСС
а) звезда; б) кластерное дерево; в) ячеистая сеть

Среди стандартов беспроводной связи, используемых в БСС, наибольшее распространение получили стандарты IEEE 802.15.4 (ZigBee) [121] и IEEE 802.11 (Wi-Fi) [48].

Таблица 4.1 – Достоинства и недостатки топологий БСС

Топология	Достоинства	Недостатки
Звезда	• простота настройки; • низкое энергопотребление; • лёгкий поиск неисправностей и обрывов в сети.	• низкая надежность (выход из строя ЦУ приводит к неработоспособности сети или ее сегмента); • низкая масштабируемость; • небольшая зона покрытия.
Кластерное дерево	• невысокое энергопотребление; • лёгкий поиск неисправностей и обрывов в сети; • масштабируемость сети и расширение зоны покрытия без дополнительных затрат на инфраструктуру.	• выход из строя главы кластера (ГК) приводит к неработоспособности кластера; • требует сложной схемы балансировки нагрузки между кластерами и узлами в кластере.
Ячеистая сеть	• высокая отказоустойчивость; • хорошая масштабируемость; • автоматический поиск маршрутов передачи данных.	• высокое энергопотребление и меньший срок службы; • сложность настройки.

К основным достоинствам БСС можно отнести возможность расположения в труднодоступных местах, оперативность и удобство развертывания, масштабируемость и надежность сети в целом – в случае выхода из строя одного из узлов данные могут быть переданы через соседние узлы. Благодаря этому БСС широко используются для экологического, промышленного, структурного, медицинского мониторинга [83, 85, 116, 119], в системах безопасности [37], в системах отслеживания цели [90], и др.

4.2 Повышение точности сенсоров БСС для экологического мониторинга

Экологический мониторинг представляет собой комплексную систему наблюдений за состоянием окружающей среды, оценки и прогноза изменений состояния окружающей среды под воздействием природных и антропогенных факторов [14]. Благодаря тому, что БСС позволяют реализовывать сбор данных с больших территорий, контроль условий окружающей среды и оперативное

принятие необходимых решений, они успешно применяются для целей экологического мониторинга [70, 85]. Объектами исследования в БСС для экологического мониторинга являются вода, почва, воздух, биоресурсы и т.п., а измеряемые сенсорами параметры могут иметь химическую, физическую, биологическую, радиологическую и т.д. природу.

Обычно узлы БСС имеют автономное питание и долгое время функционируют в сложных и экстремальных условиях окружающей среды. В результате климатические факторы и разрядка элемента питания влияют на качество предоставляемой мультисенсором информации. Кроме того, поскольку множество узлов БСС обычно обладает большой мощностью, для снижения стоимости развертывания сети часто используются недорогие сенсоры, которые характеризуются низкой точностью измерений. Ручная калибровка мультисенсоров требует значительных материальных и временных ресурсов, а автокалибровка сопряжена с дополнительными энергетическими затратами [29, 101, 102]. Все это приводит к ситуации, когда данные, полученные мультисенсорами сети, не могут быть признаны надежными.

Для решения проблемы ненадежных данных в БСС может быть использован метод комплексирования IF&PA, представленный в главе 2.

Рассмотрим БСС для экологического мониторинга, имеющую топологию «кластерное дерево», где все множество узлов разделено на непересекающиеся кластеры (рисунок 4.2) [77]. Каждый кластер состоит из m сенсорных узлов $S = \{s_1, s_2, \ldots, s_m\}$, распределенных по некоторой площади. Узлы одного кластера, расположенные на некоторой исследуемой области в пределах диапазона передачи друг друга, будем называть соседними узлами. Благодаря высокой плотности распределения узлов в сети, показания сенсоров соседних узлов в кластере коррелированны друг с другом. В рассматриваемой сети соседние узлы измеряют одно и то же значение физической величины, которое затем может быть обоснованно приписано всей исследуемой области [41].

Каждый узел оснащен мультисенсором, включающим сенсоры, которые измеряют p параметров окружающей среды [8]. Обладая информацией о

неопределенности ε_k^i i-ого сенсора, каждый k-й узел предоставляет измерительные данные об i-й величине в форме интервала $d_k^i = [x_k^i \pm \varepsilon_k^i]$, где x_k^i – измеренное сенсором значение. Узлом формируется набор данных $D_k = \{d_k^1, d_k^2, \ldots, d_k^p\}$ обо всех измеряемых мультисенсором величинах и передается главе кластера (ГК). Функции ГК в каждом j-ом кластере заключаются в сборе данных $D^j = \{D_1, D_2, \ldots, D_m\}$ с m узлов в кластере (включая собственные данные) и их передаче на ЦУ. ЦУ обладает измерительной информацией от всех узлов в сети, которую может использовать для дальнейшей обработки.

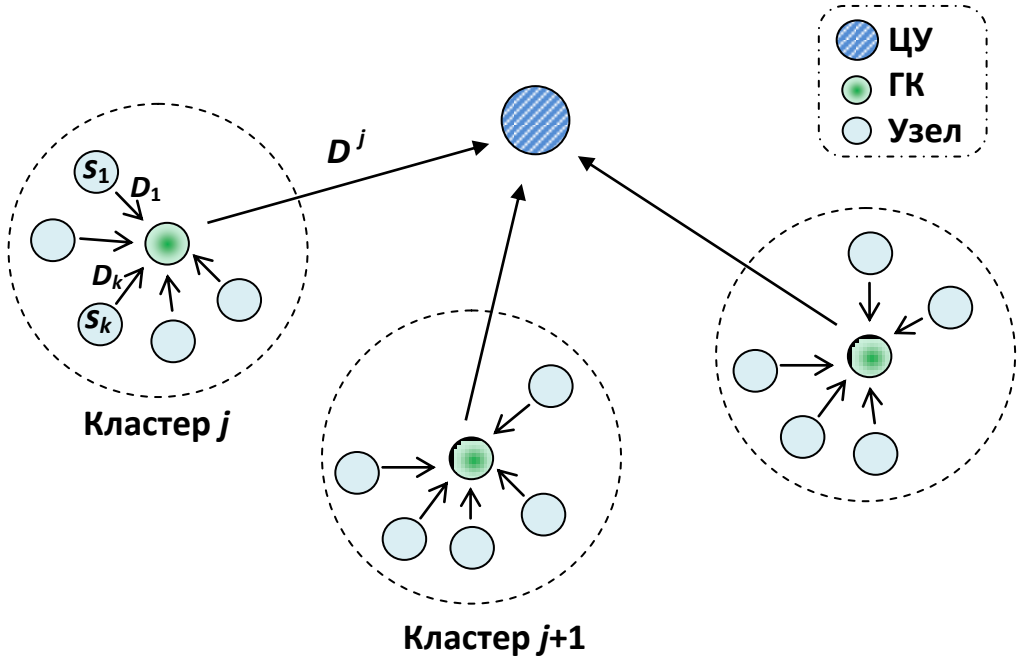


Рисунок 4.2 – Модель БСС

Таким образом, для каждого кластера имеются данные измерений каждой i-ой величины, полученные от m сенсоров. На основании этих данных, которые с большой вероятностью могут быть неточными, неполными или противоречивыми, необходимо определить максимально возможно точное значение измеряемой величины. Поскольку измерительные данные представлены в виде интервалов, они могут быть обработаны методом комплексирования IF&PA. При этом метод IF&PA применим к числовым интервалам для всех p величин в независимости от природы измеряемой величины.

Алгоритм повышения точности SensAcc результата измерений сенсоров каждой *i*-ой величины в *j*-ом кластере в БСС состоит из следующих этапов.

***Этап* 1.** Получение значения измеряемой величины x_k сенсором *k*-го узла.

***Этап* 2.** Формирование интервала $d_k = [x_k - \varepsilon_k, x_k + \varepsilon_k]$ на основании известной неопределенности ε_k сенсора *k*-го узла.

***Этап* 3.** Формирование набора исходных интервалов $\{I_k\}$ для *m* узлов в кластере, представляющих измерительные данные d_k, $k = 1, \ldots, m$ (рисунок 4.3).

***Этап* 4.** Нахождение для интервалов $\{I_k\}$ результата комплексирования x^* и неопределенности ε^* методом IF&PA (см. п. 2.2.3).

***Этап* 5.** Формирование значения измеряемой величины в *j*-м кластере в виде результата комплексирования x^* с неопределенностью результата измерения ε^*.

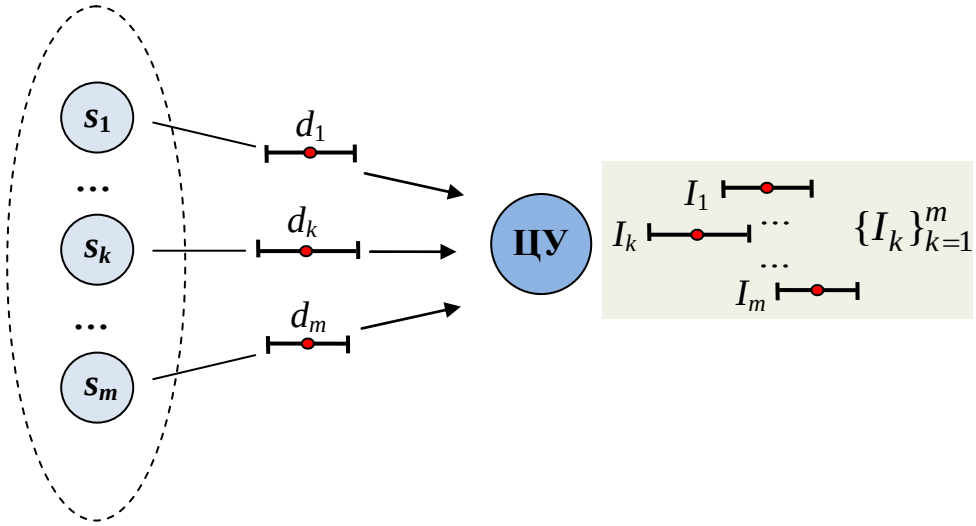


Рисунок 4.3 – Формирование набора исходных интервалов $\{I_k\}$ из измерительных данных *m* узлов кластера

Заметим, что алгоритм повышения точности сенсоров на основе IF&PA реализуется в ЦУ, не ограниченном емкостью элемента питания, и таким образом не влияет на энергопотребление в сети.

4.2.1 Верификация алгоритма повышения точности на синтетических входных данных

В целях проверки работоспособности предложенного алгоритма повышения точности мультисенсоров были проведены численные

экспериментальные исследования с помощью ПО IntFusion, описанного в п. 2.3. Получено свидетельство о государственной регистрации программы для ЭВМ № 2016663686 «Повышение точности сенсоров беспроводной сети методом агрегирования предпочтений» [9].

В качестве объекта экспериментальных исследований рассматривалась БСС для экологического мониторинга, предназначенная для контроля параметров почвы в некотором географически ограниченном регионе. При этом предполагалось, что мультисенсоры m узлов сети измеряют три величины:

- температуру t,
- объемную влажность h,
- электропроводность G.

Моделируемая БСС состояла из 151 узла и была разделена на 10 кластеров, по 15 узлов каждый. Для m мультисенсоров в каждом кластере генерировались синтетические интервальные данные измерений трех величин. Задаваемые параметры генерации (см. п. 2.3.3) синтетических (т.е. сгенерированных по методу Монте-Карло) измеренных значений, распределенных по нормальному закону, для трех величин приведены в таблице 4.2. В качестве максимальной неопределенности Δx использовалась паспортная неопределенность сенсора величины. Используемые при генерации значения Δx согласованы с неопределенностями реальных сенсоров.

Таблица 4.2 – Параметры генерации данных

Величина	m	x_{nom}	Δx	R_x	R_ε
t, °C	15	15	1 °C	1,5	2
h, % ОВС	15	22	3 % ОВС	1,5	2
G, мСм/м	15	3,5	3 %	1,5	2

Сгенерированные измерительные данные для m мультисенсоров в одном кластере приведены в таблице 4.3. После обработки данных методом IF&PA были получены значения x^*, ε^* и ξ для кластера и усредненные значения $\overline{\varepsilon}_k$ и $\overline{\xi}_k$ по узлам кластера, представленные в таблице 4.4. На рисунке 4.4 показаны исходные

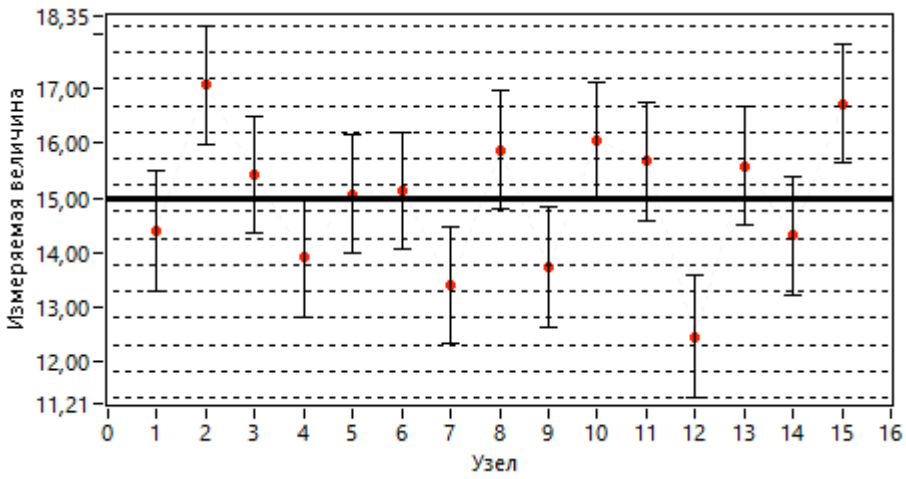
интервалы m узлов и результирующий интервал $I_r = [x^* - \varepsilon^*, x^* + \varepsilon^*]$ для трех измеряемых величин.

Таблица 4.3 – Сгенерированные измерительные данные

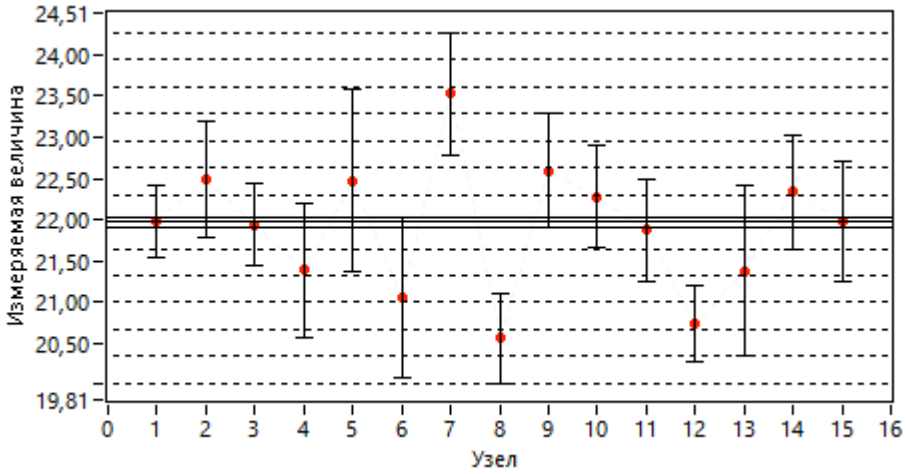
k	t, °C		h, %		G, мСм/м	
	x_k	ε_k	x_k	ε_k	x_k	ε_k
1	14,391	1,107	21,966	0,439	3,565	0,502
2	17,082	1,085	22,480	0,707	2,708	0,479
3	15,421	1,084	21,935	0,504	2,659	0,409
4	13,911	1,111	21,383	0,808	4,243	0,382
5	15,068	1,098	22,467	1,107	3,996	0,428
6	15,135	1,063	21,061	0,976	3,915	0,468
7	13,386	1,085	23,529	0,738	3,672	0,436
8	15,887	1,085	20,553	0,543	4,020	0,404
9	13,737	1,108	22,598	0,685	2,682	0,544
10	16,043	1,094	22,280	0,614	3,843	0,477
11	15,672	1,092	21,873	0,615	2,311	0,377
12	12,441	1,119	20,730	0,464	3,073	0,459
13	15,592	1,078	21,373	1,031	4,058	0,398
14	14,308	1,094	22,332	0,688	2,388	0,517
15	16,737	1,082	21,985	0,729	3,396	0,525

Таблица 4.4 – Результаты комплексирования x^*, ε^* и ξ для кластера; усредненные по узлам кластера неопределенности $\overline{\varepsilon}_k$ и отклонения $\overline{\xi}_k$

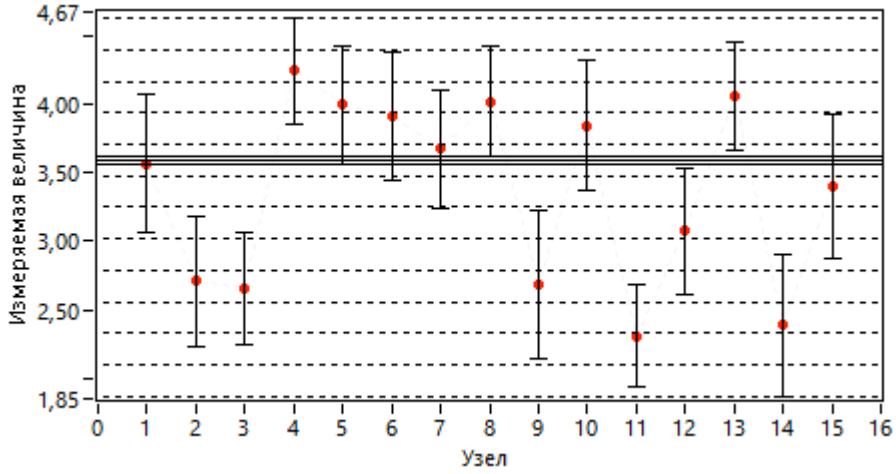
t, °C					h, %					G, мСм/м				
Кластер			Узлы		Кластер			Узлы		Кластер			Узлы	
x^*	ε^*	ξ	$\overline{\varepsilon}_k$	$\overline{\xi}_k$	x^*	ε^*	ξ	$\overline{\varepsilon}_k$	$\overline{\xi}_k$	x^*	ε^*	ξ	$\overline{\varepsilon}_k$	$\overline{\xi}_k$
14,989	0,033	0,011	1,092	1,031	21,975	0,062	0,025	0,710	0,588	3,592	0,025	0,092	0,454	0,573



а)



б)



в)

Рисунок 4.4 – Исходные интервалы и результаты комплексирования для величин а) температура, б) влажность, в) электропроводность

Из таблицы 4.4 видно, что отклонения ξ результатов комплексирования x^* на 1-2 порядка меньше, чем усредненные отклонения $\overline{\xi}_k$ значений x_k, измеренных сенсорами. При этом неопределенности ε* результатов x^* в 10-30 меньше, чем усредненные неопределенности $\overline{\varepsilon}_k$ измеренных значений x_k. Заметим, что неопределенность ε* результата комплексирования x^* зависит от количества m узлов в кластере: чем больше m, тем меньше значение ε* (см. п. 4.3.2). Результаты позволяют сделать вывод о том, что предложенный алгоритм SensAcc обеспечивает существенное повышение точности результата измерений сенсоров в БСС.

4.2.2 Результаты обработки данных реальной БСС

Алгоритм повышения точности SensAcc на основе IF&PA был применен для обработки данных, полученных в реальной БСС.

БСС, состоящая из 54 сенсорных узлов, была развернута в исследовательской лаборатории Intel Berkeley Research lab для контроля параметров окружающей среды в период с 28 февраля по 5 апреля 2004 г. [49].

В сети использовались беспроводные сенсорные узлы Mica2Dot [68] с платой Weather board, оснащенной сенсорами температуры t и влажности h SensirionSHT11[35] с соответствующими погрешностями ε_t = 0,4 °C и ε_h = 3 %. Функционирование узла осуществлялось на основе операционной системы TinyOS 1.0.

Схема расположения сенсорных узлов в лаборатории показана на рисунке 4.5. Все узлы сети разделены на кластеры, при этом ГК на схеме обозначены круглыми фигурами, а остальные узлы кластера – шестиугольными.

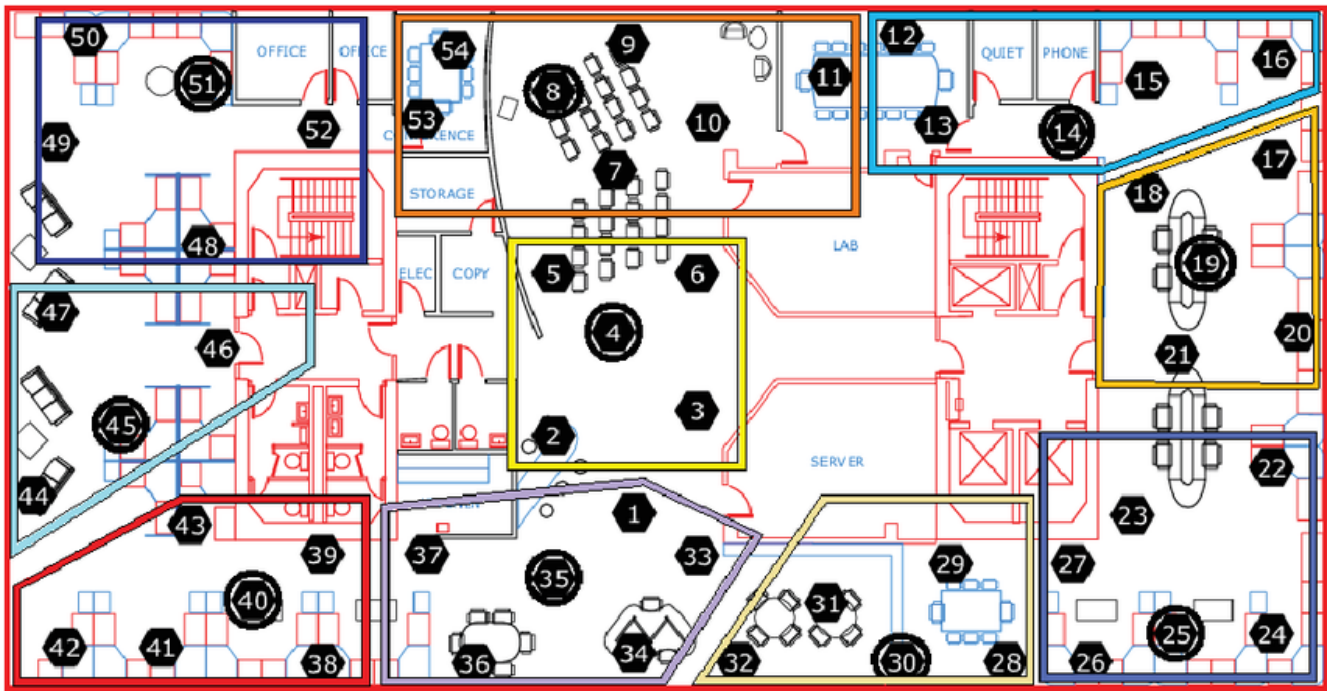


Рисунок 4.5 – Схема БСС в Intel Berkeley Research Lab

Сенсоры узлов собирали данные о температуре и влажности с интервалом 31 с. Показания сенсоров для каждого кластера, собранные в одном цикле опроса (для разных кластеров использовались разные циклы), представлены в таблице 4.5. Заметим, что в каждом цикле опроса существовали узлы, данные с которых

отсутствовали. Кроме того, в некоторых случаях узлы предоставляли показания, сильно отличающиеся от показаний соседних узлов (например, кластер 4, узел 21), т.е. имели место выбросы. В таблице 4.5 также приведены нижние l_k и верхние u_k границы исходных интервалов, рассчитанные на основании погрешностей сенсоров. В качестве примера на рисунке 4.6 показаны исходные интервалы и результирующий интервал $I_г = [x^* - \varepsilon^*, x^* + \varepsilon^*]$, полученные для значений температуры, измеренных 5 узлами кластера 10.

Таблица 4.5 – Показания сенсоров (средние точки), нижние и верхние границы исходных интервалов

j	***k***	***t*, °C**			***h*, %**		
		x_k	l_k	u_k	x_k	l_k	u_k
1	2	22,6148	22,2148	23,0148	45,4402	42,4402	48,4402
	3	22,6246	22,2246	23,0246	44,3494	41,3494	47,3494
	4	21,4976	21,0976	21,8976	48,0918	45,0918	51,0918
	5	-	-	-	-	-	-
	6	22,0954	21,6954	22,4954	45,8353	42,8353	48,8353
2	7	21,4388	21,0388	21,8388	35,1609	32,1609	38,1609
	8	20,8312	20,4312	21,2312	36,4052	33,4052	39,4052
	9	20,9292	20,5292	21,3292	37,1620	34,1620	40,1620
	10	20,7528	20,3528	21,1528	37,7107	34,7107	40,7107
	11	19,1260	18,726	19,526	41,4444	38,4444	44,4444
	53	-	-	-	-	-	-
	54	21,5172	21,1172	21,9172	47,8966	44,8966	50,8966
3	12	25,0648	24,6648	25,4648	36,1983	33,1983	39,1983
	13	24,7316	24,3316	25,1316	38,1213	35,1213	41,1213
	14	24,1142	23,7142	24,5142	38,7016	35,7016	41,7016
	15	-	-	-	-	-	-
	16	22,9480	22,548	23,348	41,7133	38,7133	44,7133
4	17	22,8500	22,45	23,25	40,9392	37,9392	43,9392
	18	22,9578	22,5578	23,3578	40,7368	37,7368	43,7368
	19	20,3314	19,9314	20,7314	38,0871	35,0871	41,0871
	20	19,9394	19,5394	20,3394	40,3652	37,3652	43,3652
	21	32,6402	32,2402	33,0402	24,6548	21,6548	27,6548
5	22	30,6410	30,241	31,041	26,9679	23,9679	29,9679
	23	29,2396	28,8396	29,6396	29,1157	26,1157	32,1157
	24	30,3078	29,9078	30,7078	27,2914	24,2914	30,2914
	25	-	-	-	-	-	-
	26	27,0444	26,6444	27,4444	33,8396	30,8396	36,8396
	27	25,7802	25,3802	26,1802	35,4726	32,4726	38,4726
6	28	27,6618	27,2618	28,0618	33,4206	30,4206	36,4206

Продолжение таблицы 4.5

j	*k*	*t*, °C			*h*, %		
		x_k	l_k	u_k	x_k	l_k	u_k
6	29	26,1232	25,7232	26,5232	36,0603	33,0603	39,0603
	30	27,7304	27,3304	28,1304	33,3507	30,3507	36,3507
	31	26,8876	26,4876	27,2876	34,9875	31,9875	37,9875
	32	-	-	-	-	-	-
7	1	23,1342	22,7342	23,5342	39,7557	36,7557	42,7557
	33	25,2412	24,8412	25,6412	36,2673	33,2673	39,2673
	34	20,1158	19,7158	20,5158	49,1615	46,1615	52,1615
	35	21,5368	21,1368	21,9368	46,4265	43,4265	49,4265
	36	20,2824	19,8824	20,6824	49,0969	46,0969	52,0969
	37	23,3596	22,9596	23,7596	42,9189	39,9189	45,9189
8	38	25,3098	24,9098	25,7098	36,0257	33,0257	39,0257
	39	-	-	-	-	-	-
	40	29,4552	29,0552	29,8552	29,3293	26,3293	32,3293
	41	28,6124	28,2124	29,0124	30,3238	27,3238	33,3238
	42	29,7492	29,3492	30,1492	28,6163	25,6163	31,6163
	43	27,8382	27,4382	28,2382	31,4902	28,4902	34,4902
9	44	19,5670	19,167	19,967	40,6693	37,6693	43,6693
	45	19,9786	19,5786	20,3786	38,9061	35,9061	41,9061
	46	19,8120	19,412	20,212	39,7896	36,7896	42,7896
	47	19,6356	19,2356	20,0356	41,1414	38,1414	44,1414
10	48	19,6356	19,2356	20,0356	40,3314	37,3314	43,3314
	49	19,0280	18,628	19,428	40,9055	37,9055	43,9055
	50	18,3224	17,9224	18,7224	43,4857	40,4857	46,4857
	51	19,5180	19,118	19,918	40,3652	37,3652	43,3652
	52	19,2436	18,8436	19,6436	41,6125	38,6125	44,6125

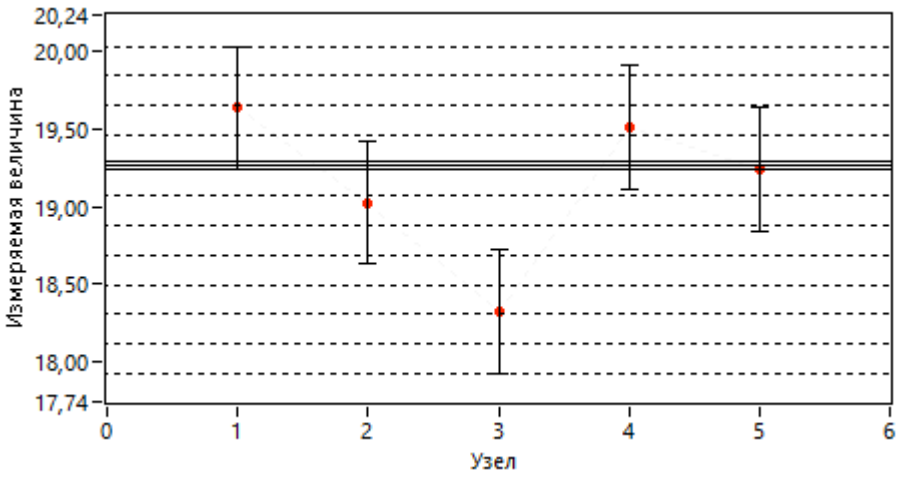


Рисунок 4.6 – Исходные интервалы и результаты комплексирования результатов измерения температуры (кластер 10)

Результаты комплексирования данных методом IF&PA представлены в таблице 4.6.

Таблица 4.6 – Результаты комплексирования данных БСС в Intel Berkeley Research Lab

j	***t*, °C**		***h*, %**	
	x*	**ε***	***x****	**ε***
1	22,350	0,126	46,221	1,129
2	20,853	0,299	36,064	1,353
3	24,881	0,217	38,956	0,242
4	21,039	0,308	39,482	1,543
5	30,412	0,171	28,319	1,649
6	26,927	0,361	34,706	1,645
7	21,790	0,147	45,195	0,724
8	27,529	0,091	29,883	1,393
9	19,773	0,194	40,024	1,882
10	19,267	0,032	41,909	1,423

Как видно из таблицы 4.6, результаты измерений t^* и h^*, полученные с помощью обработки методом IF&PA, характеризуются гораздо меньшими значениями неопределенностей $\varepsilon_t^* < \varepsilon_t$ и $\varepsilon_h^* < \varepsilon_h$, чем результаты измерений, полученные каждым сенсором в отдельности: при измерении температуры в среднем в 2 раза, а влажности – в 2,3 раза.

4.3 Энергосбережение в БСС

Проблема энергосбережения является одной из актуальных проблем для БСС. Обычно сенсорные узлы имеют автономное питание и долгое время функционируют в сложных и экстремальных условиях окружающей среды или в областях, труднодоступных для технического обслуживания [15, 75]. Таким образом, модуль питания узла не может быть заменен или подзаряжен, и его разрядка автоматически приводит к выводу узла из состава сети и снижению общего времени жизни БСС. Под ***временем жизни*** сети будем понимать время от запуска БСС до выхода из строя более 50 % сенсорных узлов [38].

Многие приложения БСС требуют использования большого количества узлов, с высокой плотностью распределенных по исследуемой области, для

обеспечения зоны покрытия, надежности сети и гарантии получения данных. Это неизбежно приводит к увеличению числа передач в сети и, как следствие, повышению энергопотребления. Стоит отметить, что расход энергии на передачу данных намного превышает расход на измерительные и вычислительные операции и может достигать 70 % от общих энергозатрат БСС [27]. Следовательно, значительно снизить энергопотребление можно за счет способа, позволяющего уменьшить число передач без сокращения количества узлов в сети.

Для решения этой задачи в БСС применяется ***циклический режим работы***, обеспечивающий чередование активного и спящего режимов функционирования узла [117]. В каждом цикле опроса часть узлов, называемая ***подмножеством активных узлов*** (ПАУ), переходит в активный режим, т.е. измеряет, передает и получает данные, в то время как оставшиеся узлы находятся в спящем режиме.

Схема выбора ПАУ предполагает нахождение компромисса между точностью измерений и энергопотреблением. Ясно, что чем больше узлов находятся в активном режиме, тем выше точность получаемого значения измеряемой величины, однако тем больше энергии требуется для передачи данных. Для реализации алгоритма выбора ПАУ прежде всего необходимо определить, ***сколько*** узлов должны быть активны в каждом цикле опроса и на основании ***каких критериев*** будут выбираться такие узлы.

В последующих параграфах представлен разработанный алгоритм выбора подмножества активных узлов (ActiveNode), основанный на агрегировании предпочтений. ActiveNode предназначен для совместного использования с методом IF&PA с целью минимизации энергопотребления при поддержании необходимого уровня точности измерений.

4.3.1 Алгоритм выбора активных узлов в кластере

Пусть БСС имеет структуру, показанную на рисунке 4.2. Результаты измерений соседних t узлов в каждом кластере комплексируются на ЦУ методом IF&PA для получения значения измеряемой величины (см. п 4.1.2). ЦУ хранит

таблицу с местоположением всех узлов в сети. Все t узлов в кластере находятся в пределах диапазона передачи R_t друг друга. Предположим, что в сети используются однотипные узлы, обладающие одинаковым исходным запасом энергии E_{in}, хотя это условие не влияет на работу ActiveNode и может быть изменено по необходимости. Сенсорный узел регистрирует количество оставшейся энергии E_{res} своего элемента питания и передает E_{res} на ГК вместе с измерительной информацией в каждом цикле опроса.

Целью применения ActiveNode является выбор ПАУ, которые потребляют минимально возможное количество энергии в кластере, при этом обеспечивая достаточный объем данных для получения значения измеряемой величины с требуемой точностью.

Для выбора ПАУ прежде всего необходимо определить узел, выполняющий функции ГК [56]. В сетях с кластерной топологией существует так называемая проблема «бутылочного горлышка», т.е. ситуация, когда ГК быстро выходит из строя вследствие разрядки элемента питания, поскольку в каждом цикле он затрачивает энергию на обмен сообщениями как с ЦУ, так и с узлами в кластере. Одним из способов решения этой проблемы является, например, использование в качестве ГК узлов с заведомо бóльшим запасом энергии, однако данный подход сопряжен с дополнительными затратами на аппаратное обеспечение и является лишь временным решением проблемы.

В алгоритме ActiveNode для снижения нагрузки на один узел при передаче сообщений предлагается применить схему ***динамического выбора*** ГК, при которой в каждом цикле опроса роль ГК выполняет новый узел.

Определим показатели, оказывающие влияние на выбор ГК. Очевидно, что затраты энергии на обмен сообщениями между ГК и ЦУ снижаются при уменьшении расстояния, т.к. энергопотребление при передаче данных прямо пропорционально расстоянию передачи [92]. Запас энергии узла также значим, поскольку выполнение функций ГК требует дополнительного расхода энергии на передачу сообщений и вычислительные операции. Кроме того, необходимо обеспечивать покрытие исследуемой области, для чего можно использовать

информацию о предыдущей активности узла. Если узел был активен в предыдущем цикле, он становится менее предпочтительным кандидатом на роль ГК. При этом местоположение ПАУ в кластере будут изменяться в каждом цикле опроса.

Таким образом, выбор ГК основывается на трех критериях:

- количество оставшейся энергии E_{res},
- расстояние до ЦУ r_{s} и
- предыдущая активность узла.

При выборе оставшихся узлов ПАУ важную роль, помимо перечисленных выше показателей, играет также точность сенсоров узла. Часто вследствие влияния внешних факторов или разрядки элемента питания сенсоры предоставляют данные, измеренные с большой неопределенностью. Такие данные, переданные на ЦУ, могут повлиять на процесс комплексирования и привести к получению неверного результата. Следовательно, важно оценить точность сенсоров узла и исключить его из измерительного процесса в случае низкой надежности предоставляемых данных. В качестве показателя качества полученной измерительной информации будем использовать обобщенную точность мультисенсора узла. При этом, поскольку мультисенсор измеряет одновременно p величин, оценка должна проводиться по каждой из них.

Для оценивания обобщенной точности вводится функционал $\Delta = \Phi(\Delta_h^k)$, где Δ_h^k – отклонение, рассчитанное для некоторого узла k по формуле:

$$\Delta_h^k = | x_h^* - x_h^k |, \; h = 1, \ldots, p; \; k = 1, \ldots, t. \qquad (4.1)$$

В формуле (4.1) x_h^* – результат комплексирования измеряемой величины h, полученный в результате обработки измерительных данных методом IF&PA (см. п. 2.2.3), а x_h^k – значение величины h, полученное k-ым узлом. Для каждого узла число отклонений Δ равно количеству измеряемых величин p, при этом узел с меньшим значением Δ_h^k является более предпочтительным для выбора, т.к. имеет бóльшую точность.

Таким образом, выбор оставшихся узлов ПАУ в каждом цикле осуществляется на основании следующих критериев:

- количество оставшейся энергии E_{res},
- расстояние до ГК r_{ch} и
- обобщенная точность мультисенсора узла.

Алгоритм ActiveNode является итеративным процессом, который запускается на ЦУ в начале каждого цикла опроса и состоит из 4 основных этапов, представленных далее.

Этап 1. Выбор главы кластера.

На этом этапе осуществляется формирование профиля предпочтения Λ, включающего в себя ранжирования t узлов по трем критериям: количество оставшейся энергии E_{res}, расстояние до ЦУ r_{s} и предыдущая активность узла. На основании профиля Λ по правилу Кемени (п. 2.1) определяется ранжирование консенсуса. Узел, стоящий на первом месте в итоговом ранжировании консенсуса, выбирается на роль ГК.

Этап 2. Выбор активных узлов в кластере.

Формируется профиль предпочтения Λ, включающий ранжирования $t - 1$ узлов по трем критериям: количество оставшейся энергии E_{res}, расстояние до ГК r_{ch} и обобщенная точность мультисенсора узла. Определяется ранжирование консенсуса. Элементами ПАУ $S_a = \{s_1, s_2, \ldots, s_g\}$, где g – число активных узлов в кластере (см. п. 4.2.2), становятся $g - 1$ первых узлов в ранжировании консенсуса и узел, выбранный в качестве ГК.

Этап 3. Активация узлов.

ЦУ рассылает всем выбранным ГК сообщение об активации, после чего ГК переходят в активный режим и передают такое же сообщение узлам своего кластера, входящим в S_a. Измерительная информация ПАУ каждого кластера вместе с информацией об E_{res} передается на ЦУ. По завершению передачи узлы S_a переходят в спящий режим.

Этап 4. Расчет значений Δ_h^k.

После получения измерительной информации на ЦУ запускается алгоритм IF&PA, который определяет результаты комплексирования x^* для каждой из p измеряемых величин. На основании x^* рассчитываются значения Δ_h^k . Формируется профиль предпочтения Λ, включающий ранжирования узлов по значениям Δ_h^k для каждой из p величин, определяется ранжирование консенсуса, которое передается в следующий цикл опроса.

Рассмотрим пример работы алгоритма ActiveNode на следующем примере.

Пример 4.1. Пусть число узлов в кластере $t = 7$, и каждый i-ый узел после нескольких циклов опроса характеризуется набором значений показателей, представленным в таблице 4.7.

Таблица 4.7 – Оставшаяся энергия, расстояние до ЦУ и предыдущая активность для семи узлов в кластере

Номер узла *i*	**Энергия E_{res}, мДж**	**Расстояние r_s до ЦУ, м**	**Предыдущая активность**
1	5	32	+
2	6	40	–
3	8	48	–
4	6	54	+
5	4	28	+
6	5	50	+
7	3	52	–

По данным таблицы 4.7 получены следующие ранжирования (альтернативы a_i обозначены их индексами i):

$$\begin{gathered}\lambda_1: 3 \succ 2 \sim 4 \succ 1 \sim 6 \succ 5 \succ 7,\\ \lambda_2: 5 \succ 1 \succ 2 \succ 3 \succ 4 \succ 7 \succ 6,\\ \lambda_3: 2 \sim 3 \succ 1 \sim 4 \sim 5 \sim 6 \sim 7,\end{gathered} \tag{4.2}$$

Тогда ранжирование консенсуса, определенное по правилу Кемени, имеет вид:

$$\beta = 2 \sim 3 \succ 1 \sim 4 \succ 5 \succ 6 \succ 7. \tag{4.3}$$

Если в ранжирование консенсуса β первое место делят между собой несколько узлов, это значит, что они одинаково предпочтительны и любой может

с равным правом быть выбран на роль ГК. Так, для приведенного примера функции ГК будет выполнять узел s_2.

Пусть мультисенсоры узлов БСС измеряют четыре параметра почвы: температуру, влажность, электропроводность и концентрацию ионов меди. Результаты измерений параметров и соответствующие значения Δ_h^k, полученные по формуле (4.3) при $p = 4$ и $t = 6$, а также результаты комплексирования x^* для каждого параметра h представлены в таблице 4.8. Заметим, что узел s_2, выбранный ранее на роль ГК, не участвует в процедуре выбора ПАУ.

Таблица 4.8 – Результаты измерений параметров почвы при $p = 4$ и $t = 6$, результаты комплексирования x^*

k	**Температура, °C**	Δ_1^k	**Влажность, %**	Δ_2^k	**Электропроводность, мСм/м**	Δ_3^k	**Концентрация ионов меди, мг/кг**	Δ_4^k
1	16,7	2	19,2	0,2	2,9	1,5	31,3	1,1
3	17,1	2,4	20,2	0,8	4,9	0,5	33,5	1,1
4	14,0	0,7	16,0	3,4	3,8	0,6	31,1	1,3
5	14,7	0,0	19,3	0,1	3,4	1,0	34,5	2,1
6	13,7	1,0	19,7	0,3	4,3	0,1	32,6	0,2
7	14,3	0,4	18,6	0,8	3,0	1,4	32,0	0,4
x^*	14,7		19,4		4,4		32,4	

Узлы были проранжированы по полученным отклонениям Δ_1^k, ..., Δ_4^k:

$$\begin{gathered} \lambda_1\colon 5 \succ 7 \succ 4 \succ 6 \succ 1 \succ 3, \\ \lambda_2\colon 5 \succ 1 \succ 6 \succ 3 \sim 7 \succ 4, \\ \lambda_3\colon 6 \succ 3 \succ 4 \succ 5 \succ 7 \succ 1, \\ \lambda_4\colon 6 \succ 7 \succ 1 \sim 3 \succ 4 \succ 5. \end{gathered} \tag{4.4}$$

В данном профиле предпочтения λ_1, λ_2, λ_3 и λ_4 являются ранжированиями для температуры, влажности, электропроводности и концентрации ионов меди соответственно.

На основании профиля было определено следующее ранжирование консенсуса:

$$\beta = \{5 \sim 6 \succ 7 \succ 1 \succ 3 \succ 4\}. \tag{4.5}$$

Из (4.5) следует, что наилучшей обобщенной точностью мультисенсора характеризуются узлы s_5 и s_6.

Таблица 4.9 содержит данные об оставшейся энергии и расстоянии до ГК для шести узлов в кластере.

Таблица 4.9 – Оставшаяся энергия и расстояние до ГК шести узлов в кластере

Номер узла i	Энергия E_{res}, мДж	Расстояние r_{ch} до ГК, м
1	5	22
3	8	24
4	6	26
5	4	16
6	5	20
7	3	18

Профиль предпочтения, включающий ранжирования узлов по энергии и расстоянию до ГК (таблица 4.9), а также по обобщенной точности мультисенсора узла имеет вид:

$$\begin{aligned} &\lambda_1 : 3 \succ 4 \succ 1 \sim 6 \succ 5 \succ 7, \\ &\lambda_2 : 5 \succ 7 \succ 6 \succ 1 \succ 3 \succ 4, \\ &\lambda_3 : 5 \sim 6 \succ 7 \succ 1 \succ 3 \succ 4. \end{aligned} \qquad (4.6)$$

По полученному ранжированию консенсуса

$$\beta = \{5 \sim 6 \succ 7 \succ 1 \succ 3 \succ 4\}. \qquad (4.7)$$

в зависимости от числа g определяется ПАУ. Например, для $g = 4$ получаем ПАУ вида $S_a = \{s_2, s_5, s_6, s_7\}$.♦

В следующем параграфе рассмотрен вопрос выбора количества g активных узлов в кластере.

4.3.2 Выбор количества активных узлов в кластере

Для выбора количества активных узлов в кластере был исследован характер зависимости неопределенности ε^* результата комплексирования x^* методом IF&PA от количества t узлов в кластере. С помощью ПО IntFusion были сгенерированы и обработаны данные для различных значений t = 5, 7, ..., 19 (по 100 индивидуальных задач для каждого t) при фиксированных значениях СКО $\sigma = 0{,}2$ и $\sigma = 1$. В качестве значений t использовались только нечетные числа, т.к. для четных t число получаемых ранжирований Кемени существенно возрастает

[74]. Полученные кривые неопределенностей ε*, упорядоченные по возрастанию, показаны на рисунках 4.7 и 4.8 для значений σ = 0,2 и σ = 1 соответственно. Для лучшей наглядности на графиках представлены кривые только для значений t = 5, 7, 9, 13, 15, 19.

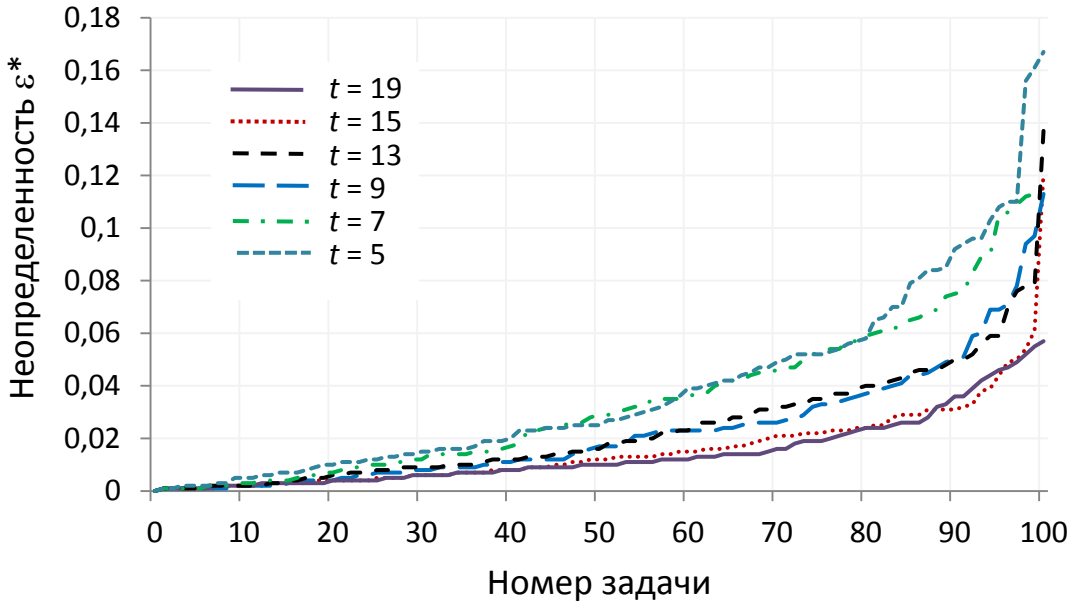


Рисунок 4.7 – Значения неопределенностей ε* при σ = 0,2 для 100 задач

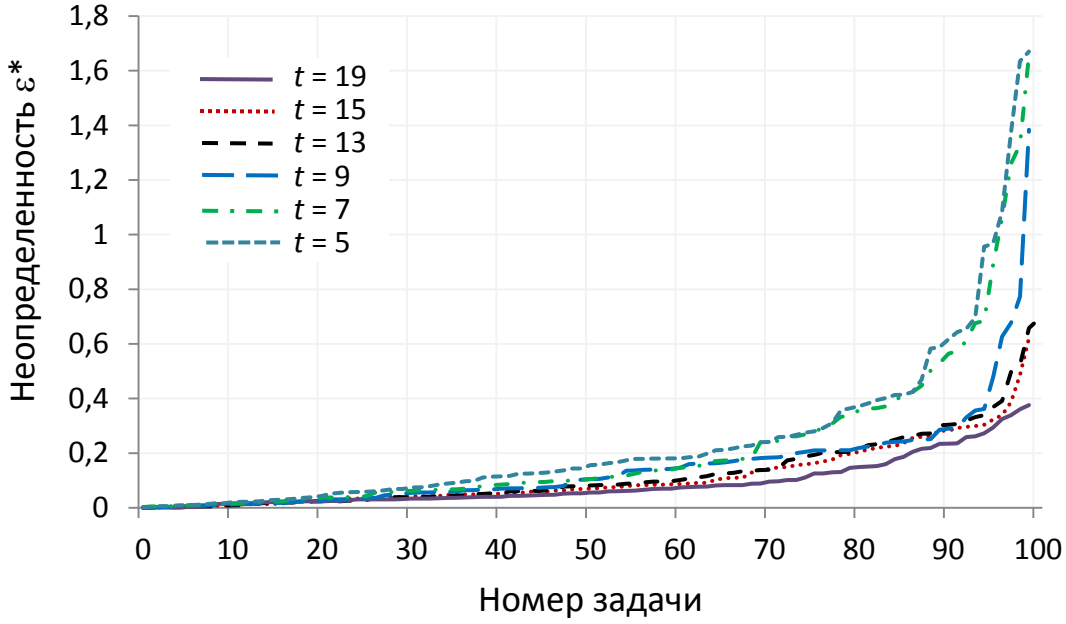


Рисунок 4.8 – Значения неопределенностей ε* при σ = 1 для 100 задач

Как видно из приведенных графиков, с увеличением числа t значение полученной неопределенности ε* результата комплексирования x* методом

IF&PA ожидаемо уменьшается. Значения ε* для наименьшего t = 5 более чем в два раза превышают ε* для наибольшего t = 19.

В таблице 4.10 представлены усредненные значения $\bar{\varepsilon}^*$ для 100 задач, полученные при различных t = 5, 7, ..., 19 для σ = 0,2 и σ = 1.

Таблица 4.10 – Усредненные значения $\bar{\varepsilon}^*$ неопределенностей для t = 5, 7, ..., 19 при σ = 0,2 и σ = 1

t	$\bar{\varepsilon}^*$	
	σ = 0,2	**σ = 1**
5	0,039	0,246
7	0,035	0,217
9	0,023	0,150
11	0,025	0,132
13	0,024	0,128
15	0,016	0,111
17	0,017	0,096
19	0,015	0,088

Из рисунков 4.7, 4.8 и данных таблицы 4.10 можно заметить, что наборы неопределенностей ε* для t = 9, 11, 13 и для t = 15, 17, 19 образуют группы, такие, что разница между средними значениями $\bar{\varepsilon}^*$ в группе не превышает 0,002 (1 % от σ) и 0,023 (2,3 % от σ) для σ = 0,2 и σ = 1 соответственно. Т.е при сокращении числа узлов с 13 до 9 (группа 1) и с 19 до 15 (группа 2) неопределенность ε* результата комплексирования x* существенно не изменяется. Следовательно, с целью экономии энергии без значительной потери точности в качестве числа узлов кластера можно выбирать наименьшее в своей группе значение t, т.е. 9 для группы 1 и 15 для группы 2.

Максимальная разница средних значений $\bar{\varepsilon}^*$ между группами 1 и 2 составляет 0,01 (0,05σ) и 0,062 (0,062σ) для σ = 0,2 и σ = 1 соответственно. Т.е число узлов t = 9 обеспечивает получение неопределенности ε*, отличающейся от ε* для t = 19 не более чем на 0,07σ.

Таким образом, при максимальном допускаемом увеличении неопределенности $\bar{\varepsilon}^*$ на 0,07σ можно сократить число узлов на 55 % (с 19 до 9).

Число g активных узлов в кластере рекомендуется принимать равным значению $0{,}45t$, округленному в бóльшую сторону до ближайшего нечетного числа.

4.3.3 Численные экспериментальные исследования алгоритма выбора активных узлов

Для экспериментального исследования предложенного алгоритма ActiveNode было разработано расширение ПО IntFusion в среде программирования NI LabVIEW. Получено свидетельство о государственной регистрации программы для ЭВМ № 2016663692 «Выбор активного подмножества узлов в кластере беспроводной сенсорной сети для снижения энергопотребления» [10]. ПО предназначено для моделирования работы кластера БСС и оценки возможности использования алгоритма ActiveNode для снижения энергопотребления сети.

Программа обеспечивает выполнение следующих функций:

- формирование ранжирований узлов по следующим критериям: расстояние между узлами, количество энергии узла, обобщенная точность мультисенсора, предыдущая активность узла;
- определение ранжирований консенсуса по правилу Кемени;
- выбор главы кластера и активных узлов на основании ранжирований консенсуса в каждом цикле опроса;
- мониторинг энергопотребления в кластере в каждом цикле опроса.

Для генерации синтетических измеренных значений сенсоров в БСС, а также комплексирования интервальных измерительных данных методом IF&PA программа вызывает модули генерации и обработки данных, описанные в п. 2.3.

Внешний вид экранного интерфейса пользователя (лицевой панели) разработанного расширения программного обеспечения IntFusion представлен на рисунке 4.9.

Лицевая панель состоит из трех основных функциональных окон.

Рисунок 4.9 – Внешний вид интерфейса пользователя расширения ПО IntFusion

В окне «Исходные параметры» осуществляется ввод пользователем следующих параметров:

- число p измеряемых величин;
- номинальные значения x_{nom} измеряемых величин;
- число t узлов в кластере;
- число g активных узлов в кластере;
- начальная энергия узлов E_{in}.

При нажатии кнопки «По умолчанию» в качестве исходных параметров будут выбраны значения по умолчанию (таблица 4.11). После ввода всех параметров программа запускается нажатием кнопки «Старт».

Таблица 4.11 – Значения исходных параметров по умолчанию

p	x_{nom}	t	g	E_{in}, мДж
5	5; 20; 2,8; 4321; 3,5	15	7	10

В окне «Текущие параметры» отображаются значения оставшейся энергии, расстояния до ЦУ, информация о предыдущей активности (Y – активен в предыдущем цикле, N – неактивен), измеренные значения величин и результаты комплексирования для всех узлов кластера, а также номер цикла опроса.

В окне «Результаты» отображаются номера ПАУ и номер узла, выбранного в качестве ГК в текущем цикле опроса.

В ходе исследований было проведено 10 экспериментальных прогонов. Предполагалось, что экспериментальная БСС состоит из 150 узлов, разделенных на 10 кластеров, по 15 узлов каждый. Каждый узел имеет в своем составе сенсоры температуры, влажности, электропроводности, освещенности и кислотности (pH). Опрос мультисенсоров осуществляется с периодом 6 с. В качестве параметров генерации использовались значения по умолчанию (таблица 4.11). Расстояния от узлов до ЦУ, использовавшиеся в исследуемой модели БСС, указаны в таблице 4.12.

Таблица 4.12 – Значения расстояний между узлами кластера и ЦУ

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
r_s	11	7	5	7	11	14	11	10	12	14	18	16	15	16	18

В каждом цикле опроса активные узлы кластера передавали синтетические измерительные данные на ЦУ, при этом расход энергии узлов был пропорционален расстоянию передачи. В моделируемом кластере БСС были реализованы три варианта функционирования:

- с использованием алгоритма ActiveNode;
- с использованием алгоритма случайного выбора узлов RandSel;
- без использования алгоритмов выбора узлов.

Следует отметить, что в варианте без использования алгоритмов выбора роль ГК играл первый узел, а при его выходе из строя – следующий по порядку второй, затем третий, и т.д.

В каждом эксперименте определялось энергопотребление отдельных узлов, общее энергопотребление в кластере и общее время жизни кластера – время до выхода из строя более 50 % узлов кластера.

На рисунке 4.10 представлена зависимость числа отказавших узлов от количества прошедших циклов опроса при использовании трех исследуемых вариантов функционирования. Алгоритм ActiveNode позволил увеличить время жизни узлов примерно в 2,5-3 раза по сравнению со значениями, полученными без использования алгоритма, и в 1,5 раза по сравнению с RandSel. Общее время жизни кластера для трех вариантов, равное времени выхода из строя восьмого по счету узла, приведено в таблице 4.13.

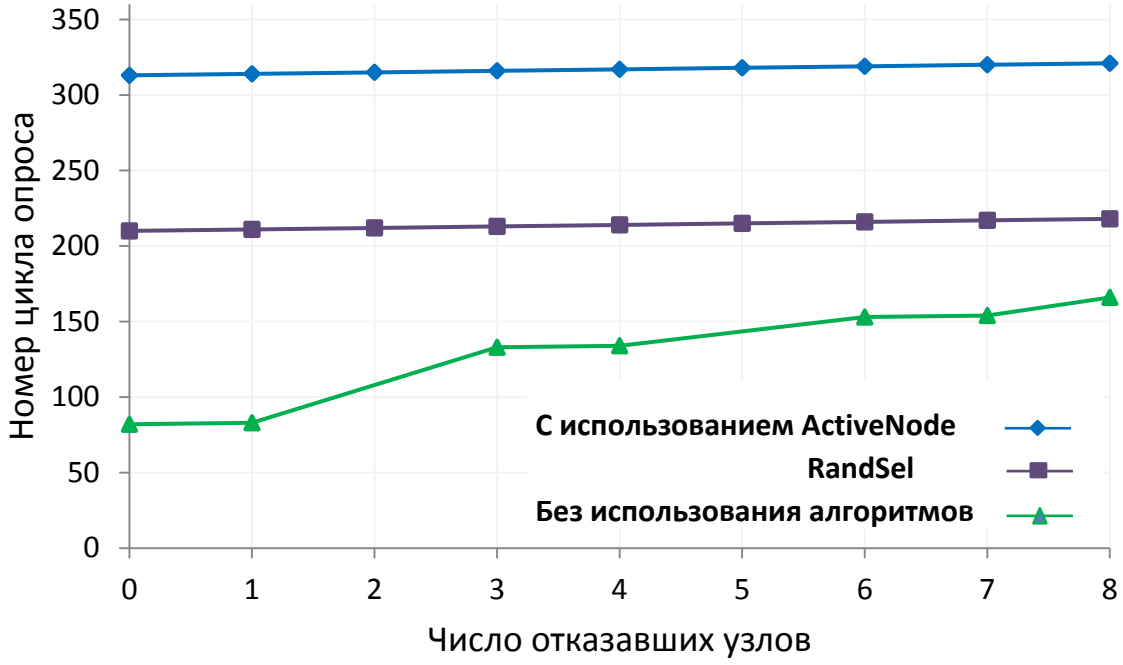


Рисунок 4.10 – Зависимость числа отказавших узлов от количества прошедших циклов опроса

На рисунке 4.11 представлены результаты сравнения трех вариантов функционирования в терминах общего энергопотребления в кластере за 166 циклов опроса. Результаты показывают, что энергопотребление в кластере при использовании алгоритма ActiveNode на протяжении всего времени жизни кластера значительно ниже, чем при использовании RandSel и без использования алгоритмов. Как видно из таблицы 4.13, в которой приведены данные об общем

энергопотреблении кластера за все время жизни, алгоритм ActiveNode показал ***наименьший расход энергии*** при ***наибольшем времени жизни***.

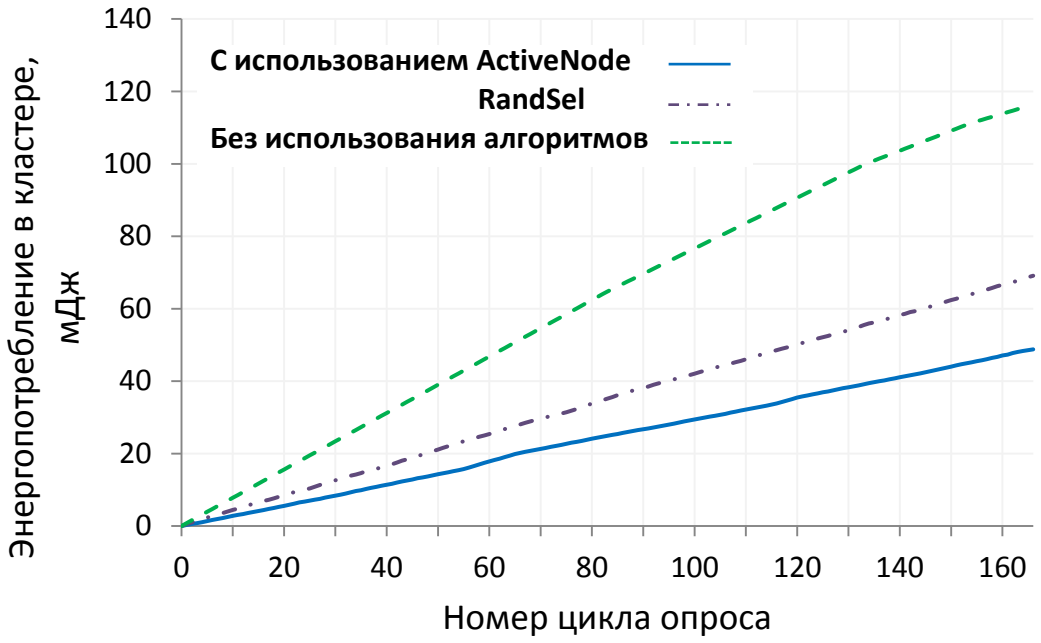


Рисунок 4.11 – Общее энергопотребление в кластере при использовании и без использования ActiveNode и RandSel

Таблица 4.13 – Сравнение вариантов функционирования в терминах времени жизни и общего энергопотребления в кластере

Вариант функционирования	**Время жизни, циклов**	**Энергопотребление, мДж**
ActiveNode	321	97,09
RandSel	218	114,37
Без использования алгоритмов	166	116,36

Таким образом, использование алгоритма ActiveNode позволяет сократить энергопотребление сети, при этом давая возможность получить значение измеряемой величины с требуемой точностью за счет объединения с методом комплексирования IF&PA.

Выводы к главе 4

1. Предложен алгоритм SensAcc повышения точности результата измерения мультисенсоров БСС для экологического мониторинга на основе метода комплексирования интервальных данных IF&PA. Верификация алгоритма

на входных синтетических данных показала, что отклонения ξ результатов комплексирования x^* на 1-2 порядка меньше, чем усредненные отклонения $\overline{\xi}_k$ значений x_k, измеренных сенсорами, при этом неопределенность ε^* полученных результатов x^* снижается в 10-30 раз.

2. Предложенный алгоритм SensAcc был применен для обработки данных, полученных в реальной БСС. Показано, что результаты комплексирования x^* характеризуются в среднем в 2-2,3 раза меньшими значениями неопределенностей ε^*, чем результаты измерений x_k, полученные каждым сенсором в отдельности.

3. Разработан алгоритм ActiveNode выбора подмножества активных узлов в кластере БСС, предназначенный для снижения энергопотребления (продление времени жизни) узлов БСС.

4. Экспериментальные исследования алгоритма ActiveNode показали, что он позволяет снизить энергопотребление (увеличить время жизни) узлов в кластере БСС в 2-3 раза, при этом обеспечивая объем данных, достаточный для достижения заданной точности измерений.

Заключение

1. Предложен и исследован метод комплексирования интервалов IF&PA, где результатом комплексирования является наилучшее дискретное значение в ранжировании консенсуса, найденном для набора наведенных интервалами ранжирований дискретных значений; метод характеризуется повышенными точностью, робастностью (независимостью от закона распределения входных данных) и достоверностью получаемого результата.
2. Для формирования ранжируемых дискретных значений предложен и экспериментально обоснован способ расчета мощности разбиения диапазона актуальных значений, полученного в результате объединения исходных интервалов, на основе поправки Шеппарда для дисперсии дискретизированных данных; способ позволяет определить значение n, при котором с вероятностью 0,95 обеспечивается получение результата комплексирования, наиболее близкого к номинальному значению для всех n от 4 до 15.
3. Разработан и исследован робастный алгоритм повышения точности результата измерения мультисенсоров в беспроводной сети на основе метода IF&PA, позволяющий снизить неопределенность результата измерения не менее чем в 2-2,3 раза по сравнению с неопределенностью показаний мультисенсоров беспроводной сенсорной сети при возможном непустом подмножестве неисправных сенсоров.
4. Разработан и исследован алгоритм выбора подмножества активных узлов в кластере беспроводной сенсорной сети на основе метода IF&PA, обеспечивающий снижение энергопотребления (продление времени жизни) узлов в кластере в 2-3 раза.
5. Результаты диссертационной работы используются в лаборатории мониторинга окружающей среды ТГУ и на кафедре систем управления и мехатроники Института кибернетики ТПУ.
6. Результаты диссертационной работы использованы при выполнении двух

НИР: гранта РНФ и базовой части государственного задания "Наука" Министерства образования и науки РФ.

Список сокращений и обозначений

1. АБ – абсолютное большинство
2. БСС – беспроводная сенсорная сеть
3. ГК – глава кластера
4. ДАЗ – диапазон актуальных значений
5. ОБ – относительное большинство
6. ПАУ – подмножество активных узлов
7. ЦУ – центральный узел
8. DFD – Dasarathy's functional model
9. JDL – Joint Directors of Laboratories
10. OEIS – On-Line Encyclopedia of Integer Sequences (Онлайн-энциклопедия целочисленных последовательностей)
11. IF&PA (interval fusion with preference aggregation) – метод комплексирования интервальных данных агрегированием предпочтений
12. RECURSALL – реализующий метод ветвей и границ рекурсивный алгоритм нахождения медианы Кемени
13. SensAcc – алгоритм повышения точности результата измерений сенсоров
14. ActiveNode – алгоритм выбора подмножества активных узлов в кластере
15. RandSel – алгоритм случайного выбора подмножества активных узлов в кластере
16. I – интервал
17. l – нижняя граница интервала
18. u – верхняя граница интервала
19. x – средняя точка интервала
20. ε – неопределенность измеренного значения
21. $\{I_k\}$ – набор исходных интервалов
22. I_r – результирующий интервал
23. x^* – результат комплексирования

24. ε*– неопределенность результата комплексирования
25. I' – интервал пересечения
26. $q(I')$ – уровень согласованности интервала пересечения
27. $A = \{a_1, a_2, ..., a_n\}$ – множество альтернатив; множество дискретных значений ДАЗ
28. n – мощность разбиения ДАЗ
29. λ – ранжирование (отношение предпочтения)
30. $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ – профиль предпочтения
31. m – число исходных интервалов
32. β – ранжирование консенсуса
33. β_{fin} – единственное итоговое ранжирование консенсуса
34. h – норма разбиения ДАЗ
35. T_n – мощность множеств инранжирований
36. F_n – мощность запрещенных ранжирований
37. m_{con} – мощность множества согласованных интервалов
38. σ_d – среднеквадратическое отклонение дискретных (после разбиения ДАЗ) значений
39. w – допускаемое различие между СКО непрерывных (до разбиения) и дискретных (после разбиения ДАЗ) значений
40. x_{nom} – номинальное значение
41. Δx – максимальная неопределенность
42. R_x – коэффициент разброса генерируемых средних точек x
43. R_ε – коэффициент разброса генерируемых неопределенностей ε
44. ξ – отклонение результата комплексирования от номинального значения
45. $P(\xi \le \xi_{гр})$ – оценки вероятностей того, что отклонение ξ не превышает некоторое фиксированное значение $\xi_{гр}$
46. $S = \{s_1, s_2, \ldots, s_m\}$ – множество узлов в кластере БСС
47. p – число измеряемых мультисенсором величин
48. d_k – измерительные интервальные данные от сенсора k-го узла

49. D_k – набор измерительных интервальных данных от k-го узла
50. D^j – набор измерительных интервальных данных от m узлов j-го кластера
51. $\overline{\varepsilon}_k$ – усредненная по узлам кластера неопределенность измеренных значений
52. $\overline{\xi}_k$ – усредненное по узлам кластера отклонение измеренных значений от номинального
53. E_{in} – исходное количество энергии узла
54. E_{res} – количество оставшейся энергии узла
55. r_{s} – расстояние от узла кластера до ЦУ
56. r_{ch} – расстояние от узла кластера до ГК
57. Δ_h^k – отклонение измеренного значения величины h от результата комплексирования
58. $S_a = \{s_1, s_2, \ldots, s_g\}$ – подмножество активных узлов в кластере
59. g – число активных узлов в кластере

Список используемой литературы

1. ГОСТ Р 54500.3.1-2011/Руководство ИСО/МЭК 98-3:2008/Дополнение 1:2008 Неопределенность измерения. Часть 3. Руководство по выражению неопределенности измерения. Дополнение 1. Трансформирование распределений с использованием метода Монте-Карло. – М.: Стандартинформ, 2012. – 82 с.

2. ГОСТ Р 54500.3-2011/Руководство ИСО/МЭК 98-3:2008 "Неопределенность измерения. Часть 3. Руководство по выражению неопределенности измерения", идентичный международному документу Руководство ИСО/МЭК 98-3:2008 "Неопределенность измерения. Часть 3. Руководство по выражению неопределенности измерения"– М.: Стандартинформ, 2012. – 101 с.

3. Гэри, М. Вычислительные машины и труднорешаемые задачи / М. Гэри, Д. Джонсон // М.: Мир, 1982. – 419 с.

4. Кемени, Дж. Кибернетическое моделирование. / Дж., Кемени, Дж. Снелл – М.: Сов. радио, 1972. – 192 с.

5. Литвак, Б.Г. Экспертная информация: Методы получения и анализа. / Б.Г. Литвак – М.: Радио и связь, 1982. – 184 с.

6. Муравьев, С.В. Агрегирование предпочтений как метод решения задач в метрологии и измерительной технике / С.В. Муравьев // Измерительная техника. – 2014. – № 2. – С. 19-23.

7. Муравьев, С.В. Обработка данных межлабораторных сличений методом агрегирования предпочтений / С.В. Муравьев, И.А. Маринушкина // Измерительная техника. – 2015. – № 12. – С. 3-7.

8. Муравьев, С.В. Цифровой цветометрический анализатор состава веществ на основе полимерных оптодов / Н.А. Гавриленко, А.С. Спиридонова, П.Ф. Баранов, Л.И. Худоногова // Приборы и техника эксперимента. – 2016. – № 4. – С. 115-123.

9. Свидетельство о государственной регистрации программы для ЭВМ № 2016663686 (RU); заявка № 2016661662 от 31.10.2016, дата рег. 13.12.2016; Бюл.

№ 1 от 10.01.2017 // Муравьев С.В., Худоногова Л.И. Повышение точности сенсоров беспроводной сети методом агрегирования предпочтений.

10. Свидетельство о государственной регистрации программы для ЭВМ № 2016663692 (RU); заявка № 2016661664 от 31.10.2016, дата рег. 13.12.2016; Бюл. № 1 от 10.01.2017 // Муравьев С.В., Худоногова Л.И. Выбор активного подмножества узлов в кластере беспроводной сенсорной сети для снижения энергопотребления.

11. Соболь, И.М. Численные методы Монте-Карло. / И.М. Соболь – М.: Наука, 1973. – 412 с.

12. Среда разработки приложений LabVIEW – National Instruments. http://russia.ni.com/labview (дата обращения: 15.01.17).

13. Тараканов, Е.В. Агрегирование данных мультисенсоров в беспроводных сенсорных сетях: диссертация на соискание ученой степени кандидата технических наук: спец. 05.11.13 / Е.В. Тараканов // ТПУ; науч. рук. С. В. Муравьев. – Томск, 2012.

14. Федеральный закон от 10.01.2002 N 7-ФЗ (ред. от 01.03.2017) «Об охране окружающей среды» (принят ГД ФС РФ 20.12.2001).

15. Худоногова, Л.И. Обеспечение отказоустойчивости алгоритмов передачи данных в беспроводных сенсорных сетях / Л.И. Худоногова, С.В. Муравьев // Ползуновский вестник. – 2015. – № 4. – С.44-46.

16. Abdelgawad, A. Resource-Aware Data Fusion Algorithms for Wireless Sensor Networks / A. Abdelgawad, M. Bayoumi // Lecture Notes in Electrical Engineering. – 2012. – Vol. 118. – P. 17-34.

17. Aguero, J.R. Inference of operative configuration of distribution networks using fuzzy logic techniques. Part II: extended real-time model / J.R. Aguero, A. Vargas // IEEE Transactions on Power Systems. – 2005. – Vol. 20. – № 3. – P. 1562-1569.

18. Alós-Ferrer, C. A simple characterization of approval voting / C. Alós-Ferrer // Social Choice and Welfare. – 2006. – Vol. 27. – Iss. 3. – P. 621-625.

19. Ayari, I. A framework for multi-sensor data fusion / I. Ayari, J. P. Haton // Proc. IEEE Symp. on Emerging Technologies and Factory Automation (Paris, France, October 1995). – P. 710-713.

20. Barthélemy, J.P. Median linear orders: heuristics and a branch and bound algorithm / J.P. Barthélemy, A. Guenoche, O. Hudry // European Journal of Operational Research. – 1989 – Vol. 42. – № 2. – P. 313-325.

21. Berg, D.E. Voting in agreeable societies / D. E. Berg, S. Norine, F. E. Su, R. Thomas, P. Wollan // Amer. Math. Monthly, 2006. – Vol. 117. – P. 27-39.

22. Bhatnagar, G. A new contrast based multimodal medical image fusion framework / G. Bhatnagar, Q.M. J. Wu, Zh. Liu // IEEE Transactions on Multimedia. – 2013. – Vol.15. – Iss. 5. – P. 1014-1024.

23. Betzler, N. Fixed-parameter algorithms for Kemeny rankings / N. Betzler, M.R. Fellows, J. Guo, R. Niedermeier, F.A. Rosamond // Theoretical Computer Science. – 2009. – Vol. 410. – Iss. 45. – P. 4554-4570.

24. Bleiholder, J. Data fusion / J. Bleiholder, F. Naumann // ACM Computing Surveys. – 2008. – Vol.41. – № 1. – P. 1-41.

25. Box, G.E.P. A note on the generation of random normal deviates / G.E.P. Box, M.E. Muller // The Annals of Mathematical Statistics. – 1958. – Vol. 29. – № 2. – P. 610-611.

26. Bruzzone, L. Data fusion experience: from industrial visual inspection to space remote-sensing application / L. Bruzzone, D. Fernandez, G. Vernazza // Proc. Academic and Industrial Cooperation in Space research (Vienna, Austria, November 1998). – P.147-151.

27. Cam, H. Energy-Efficient Secure Pattern Based Data Aggregation for Wireless Sensor Networks / H. C¸am, S. Ozdemir, P. Nair, D. Muthuavinashiappan, and H. Ozgur Sanli // Computer Communications. – 2006. – Vol. 29. – №. 4. – P. 446-455.

28. Carlson, R.J. Voter compatibility in interval societies / R.J. Carlson // HMC Senior Theses, Harvey Mudd College, 2013.

29. Collett, M.A. Aggregating measurement data influenced by common effects / M.A. Collett, M.G. Cox, T.J. Esward, P.M. Harris, J.A. Sousa // Metrologia. – 2007. – Vol. 44. – № 5. – P. 308-318.

30. Conitzer, V. Improved Bounds for Computing Kemeny Rankings / V. Conitzer, A. Davenport, J. Kalagnanam // Proc. 21st National Conf. on Artificial intelligence (Boston, USA, July 2006). – P. 620-626.

31. Cox, M.G. The evaluation of key comparison data: determining the largest consistent subset / M.G. Cox // Metrologia. – 2007. – Vol.44. – P. 187-200.

32. Cremer, F. S. Sensor fusion for anti-personnel land mines detection / F.Cremer, E. den Breejes, S. Klamer // Proc. 3rd Eurofusion Conf. (October 1998). – P. 63-70.

33. Dasarathy, B.V. Sensor fusion potential exploitation-innovative architectures and illustrative applications / B.V. Dasarathy // Proceedings of the IEEE. – 1997. – Vol. 85. – Iss.1. – P. 24-38.

34. Dasgupta, P. On the robustness of majority rule / P. Dasgupta, E. Maskin // Journal of the European Economic Association. – 2008. – Vol. 6. – № 5. – P. 949-973.

35. Datasheet SHT1x Humidity and Temperature Sensor IC. http://www.mouser.com/ds/2/682/Sensirion_Humidity_SHT1x_Datasheet_V5-469722.pdf (дата обращения: 28.03.2017).

36. Davenport, A. Ranking Pilots in Aerobatic Flight Competitions / A. Davenport, D. Lovell // IBM Research Report RC23631 (W0506-079), T.J. Watson Research Center, NY, 2005.

37. Díaz-Ramírez, A. Wireless Sensor Networks and Fusion Information Methods for Forest Fire Detection / A. Díaz-Ramírez. L.A. Tafoya. J.A. Atempa. P. Mejía-Alvarez // Procedia Technology. – 2012. – Vol. 3. – P. 69-79.

38. Dietrich, I. On the Lifetime of Wireless Sensor Networks / I. Dietrich, F. Dressler // ACM Transactions on Sensor Networks. – 2009. – Vol. 5. – Iss. 1. – P. 1-38.

39. Dopazo, E. Rank aggregation methods dealing with ordinal uncertain preferences / E. Dopazo, M.L. Martinez-Cespedes // Expert Systems with Applications. – 2017. – Vol. 78. – Iss. C. – P. 103-109.

40. Durrant-Whyte, H. Multisensor data fusion / H. Durrant-Whyte, T.C. Henderson // Springer Handbook of Robotics. – 2008. – Part C. – P. 585-610.

41. Duta, M. The Fusion of Redundant SEVA Measurements / M. Duta, M. Henry // IEEE Transactions on Control Systems Technology. – 2005. –Vol. 13. – Iss.2. – P. 173-184.

42. Edwards, I. Fusion of NDT data / I. Edwards, X.E. Gross, D.W. Lowden, P. Strachan // British Journal of NDT. – 1993. – Vol. 35. – № 12. – P. 710-713.

43. Fagin, R. Comparing and Aggregating Rankings with Ties / R. Fagin, R. Kumar, M. Mahdian, D. Sivakumar, E. Vee // Proc. 23rd ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems (New York, USA, 2004). – P. 47-58.

44. Fan, Zh.P. An Approach to Solve Group-Decision-Making Problems With Ordinal Interval Numbers / Zh.P. Fan,Y. Liu // IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). – 2010. – Vol. 40. – Iss. 5. – P. 1413-1423.

45. Hall, D.L. An Introduction to Multisensor Data Fusion / D.L. Hall, J. Llinas // Proceedings of the IEEE. – 1997. – Vol. 85. – Iss.1. – P. 6-23.

46. Halpern, J. Reasoning about Uncertainty / J. Halpern // MIT Press, Cambridge: 2003. – 473 p.

47. Hoover, A. A real-time occupancy map from multiple video streams / A. Hoover, B.D. Olsen // Proc. IEEE Intern. Conf. on Robotics and Automation (Detroit, USA, May 1999). – P. 2261-2266.

48. IEEE – Standards Association. IEEE 802.11: Wireless LANs http://standards.ieee.org/about/get/802/802.11.html (дата обращения: 03.02.2017).

49. Intel Lab Data. http://db.lcs.mit.edu/labdata/labdata.html (дата обращения: 28.03.2017).

50. Kalman, R.E. A new approach to linear filtering and prediction problems / R.E. Kalman // Journal of Basic Engineering. – 1960. – Vol. 3. – P. 35-45.

51. Kearfott, R.B. Applications of Interval Computations / R.B. Kearfott, V. Kreinovich // Springer Science & Business Media: 2013. – 428 p.

52. Keener, J.P. The Perron-Frobenius theorem and the ranking of football teams / J.P. Keener // SIAM review. – 1993. – Vol. 35. – Iss. 1. – P. 80-93.

53. Kelly, G. Data fusion: From metrology to process measurement / G. Kelly // Proc. 16th IEEE Instrumentation and Measurement Technology Conf. (Venice, Italy, May 1999). – Vol. 3. – P. 1325-1329.

54. Kemeny, J. Mathematics without numbers / J. Kemeny // Daedalus. – 1959. – Vol. 88. – P. 571-591.

55. Khaleghi, B. Multisensor data fusion: A review of the state-of-the-art / B. Khaleghi , A. Khamis, F.O. Karray // Information Fusion. – 2013. –Vol. 14. – Iss.1. – P. 28-44.

56. Khudonogova, L.I. Energy-accuracy aware active node selection in wireless sensor networks / L.I. Khudonogova, S.V. Muravyov // Сборник материалов XII Международной IEEE Сибирской конференции по управлению и связи SIBCON-2016 (Москва, 12-14 мая 2016 г.). – С. 7491835.

57. Kumar, M. Maximum likelihood wavelet fusion for aerospace NDE applications / M. Kumar, P. Ramuhalli // Proc. IEEE Intern. Conf. on Electro Information Technology (Lincoln, USA, May 2005). – P. 9037026.

58. Lakhtaria, K.I. Technological advancements and applications in mobile ad-hoc networks : research trends / K.I. Lakhtaria // IGI Global, Hershay: 2012. – 507 p.

59. Land, A.H. An automatic method of solving discrete programming problems / A.H. Land, A.G. Doig // Econometrica. – 1960. – Vol. 28. – № 3. – P. 497-520.

60. Li, B. Fault-tolerant interval estimation fusion by Dempster-Shafer Theory / B. Li, Y. Zhu, Li X. R. // Proc. 5th Intern. Conf. on Information Fusion (Annapolis, USA, July 2002). – P. 7412261.

61. Liang, H. The fusion process of interval opinions based on the dynamic bounded confidence /H. Liang, C.C. Li, Y. Dong,Y. Jiang // Information Fusion. – 2016. – Vol. 29. – Iss. C. – P. 112-119.

62. Linn, R.J. A survey of data fusion systems / R. J. Linn, D.L. Hall // Proc. SPIE Conf. on Data Structure and Target Classification (Orlando, USA, August 1991). – P. 13-36.

63. Liu, H.J. Data fusion based on interval Dempster-Shafer Theory for emitter platform identification / H.J. Liu, B. Wang, Zh. Liu, Y.Y. Zhou // Proc. Intern. Conf. on Information Engineering and Computer Science (Wuhan, China, December 2009). – P. 11032825.

64. Mana, G. Model selection in the average of inconsistent data: an analysis of the measured Planck-constant values / G. Mana, E. Massa, M. Predescu // Metrologia. – 2012. – Vol.49. – P. 492-500.

65. Marzullo, K. Tolerating Failures of Continuous-Valued Sensors / K. Marzullo // ACM Transaction on Computer System. – 1990. – Vol. 8. –№ 4. – P. 284-304.

66. Matin, M.A. Overview of Wireless Sensor Network / M.A. Matin, M.M. Islam // InTechOpen: 2012.

67. Messner, M. Robust Political Equilibria under Plurality and Runoff Rule / M. Messner, M.K. Polborn // SSRN Electronic Journal. – 2002. – P. 1-35.

68. Mica2Dot. Wireless microsensor mote. https://www.eol.ucar.edu/isf/facilities/isa/internal/CrossBow/DataSheets/mica2dot.pdf (дата обращения: 28.03.2017).

69. Moore, R.E. Introduction to Interval Analysis / R.E. Moore, R.B. Kearfott, M.J. Cloud // SIAM: 2009. – 234 p.

70. Mukhopadhyay, S.C. Wireless Sensor Networks and Ecological Monitoring / S.C. Mukhopadhyay, J.-A. Jiang // Smart Sensors, Measurement and Instrumentation. – 2013. – Vol. 3. – P. 1-297.

71. Murav’ev, S.V. Aggregation of preferences as a method of solving problems in metrology and measurement technique / S.V. Murav’ev // Measurement Techniques. – 2014. – Vol. 57. – № 2. – P. 132-138.

72. Muravyov, S.V. Axiomatic definition of quantity as a basis for teaching metrology in Mechanical Engineering / S.V. Muravyov, B. Ramamoorthy // Journal of Physics: Conference Series. – 2016. – Vol. 772. – № 1. – P. 012050.

73. Muravyov, S.V. Consensus rankings in prioritized converge-cast scheme for wireless sensor network / S.V. Muravyov, Sh. Tao, M. Ch. Chan, E.V. Tarakanov // Ad Hoc Networks. – 2015. – Vol. 24. – Part A. – P. 160-171.

74. Muravyov, S.V. Dealing with chaotic results of Kemeny ranking determination / S.V.Muravyov // Measurement. – 2014. – Vol. 51. – P. 328-334.

75. Muravyov, S.V. Feasibility estimation of creating fault-tolerant prioritized transmission scheme in WSN / S.V. Muravyov, L.I. Khudonogova // Сборник материалов XI Международной IEEE Сибирской конференции по управлению и связи SIBCON-2015 (Омск, 21-23 мая 2015 г.). – С. 7147265.

76. Muravyov, S.V. Multiple solutions of an exact algorithm for determination of all Kemeny rankings: preliminary experimental results / S.V.Muravyov, E.V.Tarakanov // Proc. Intern. Conf. on Instrumentation, Measurement, Circuits and Systems (Hong Kong, 13-14 December 2011) ASME Press New York. – Vol. 1. – P. 17-20.

77. Muravyov, S.V. Multisensor accuracy enhancement on the base of interval voting in form of preference aggregation in WSN for ecological monitoring / S.V. Muravyov, L.I. Khudonogova // Proc. Intern. Congress on Ultra Modern Telecommunications and Control Systems and Workshops (Brno, Czech Republic, November 2015). – P. 293-297.

78. Muravyov, S.V. Ordinal measurement, preference aggregation and interlaboratory comparisons / S.V. Muravyov // Measurement. – 2013. – Vol. 46. – P. 2927–2935.

79. Muravyov, S.V. Processing of interlaboratory comparison data by preference aggregation method / S.V. Muravyov, I.A. Marinushkina // Measurement Technique. – 2016. – Vol. 58. – № 12. – P.1285-1291.

80. Muravyov, S.V. Rankings as ordinal scale measurement results / S.V. Muravyov // Metrology and Measurement Systems. – 2007. – Vol. 13. – № 1. – P. 9-24.

81. Muravyov, S.V. Representation theory treatment of measurement semantics for ratio, ordinal and nominal scales / S.V. Muravyov, V. Savolainen // Measurement. – 1997. – Vol. 22. – P. 37-46.

82. Muravyov, S.V. Representation of interval data by weak orders yields robustness of the data fusion outcomes / S.V. Muravyov, L.I. Khudonogova, I.A. Marinushkina // Journal of Physics: Conference Series. – 2016. – V. 772. – № 1. – P. 012064.

83. Neves, P. Application of Wireless Sensor Networks to Healthcare Promotion / P. Neves, M. Stachyra, J. Rodrigues // Journal of Communications Software and Systems. – 2008. – Vol. 4. – № 3. – P. 181-190.

84. Neyman, J. Outline of a theory of statistical estimation based on the classical theory of probability / J. Neyman // Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences. – 1937. – Vol. 236. – № 767. – P. 333-380.

85. Oliveira, L.M.L. Wireless Sensor Networks: a Survey on Environmental Monitoring / L.M.L. Oliveira, J.J.P.C. Rodrigues // Journal of communications. – 2011. – Vol. 6. – № 2. – P. 143-151.

86. Parhami, B. Distributed Interval Voting with Node Failures of Various Types / B. Parhami // Proc. 12th IEEE Workshop on Dependable Parallel, Distributed and Network-Centric Systems (California, USA, 2007). – P. 1-7.

87. Parhami, B. Voting algorithms / B. Parhami // IEEE Transactions on Reliability. – 1994. – Vol. 43. – Iss. 4. – P. 617-629.

88. Parhami, B. Voting: a paradigm for adjudication and data fusion in dependable systems / B. Parhami // Dependable Computing Systems: Paradigms, Performance Issues, & Applications. –2005. – Vol. 52. – №. 2 – P. 87-114.

89. Pendrill, L.R. Using measurement uncertainty in decision-making and conformity assessment / L.R. Pendrill // Metrologia. – 2014. – Vol.51. – P. 206-218.

90. Ramya, K. A Survey on Target Tracking Techniques in Wireless Sensor Networks / K. Ramya, K. P. Kumar, V. S. Rao // International Journal of Computer Science & Engineering Survey. – 2012. – Vol.3. – № 4. – P. 93-108.

91. Raol, J. R. Data fusion mathematics: theory and practice / J. R. Raol. – CRC Press: 2015. – 600 p.

92. Rappaport, T. Wireless Communications: Principles and Practice / T. Rappaport // Prentice-Hall PTR: 2001. – 736 p.

93. Schneeweiss, H. Symmetric and asymmetric rounding: a review and some new results / H. Schneeweiss, J. Komlos, A.S. Ahmad // AStA Advances in Statistical Analysis. – 2010. – Vol. 94. – Iss. 3. – P 247-271.

94. Schulze, M. A New Monotonic and Clone-Independent Single-Winner Election Method / M. Schulze // Voting Matters. – 2003. – Iss. 17. – P. 9-19.

95. Sentinella, D.J. Real time data fusion / D.J. Sentinella, A.G. Raines // Proc. IEE Colloquium on Strategic Control of Inter-Urban Road Networks (London, UK, March 1997). – P. 5.

96. Shafer, G. A Mathematical theory of evidence / G. Shafer // Princeton University Press: 1976. – p. 314.

97. Shiryaev, D. On Elections with Robust Winners / D. Shiryaev, L. Yu, E. Elkind // Proc. Intern. conf. on Autonomous agents and multi-agent systems (St. Paul, USA, May 2013). – P. 415-422.

98. Sloane, N.J.A. The Encyclopedia of Integer Sequences / N.J.A. Sloane, S. Plouffe // Academic Press, San Diego: 1995.

99. Sommer, K.D. A Bayesian approach to information fusion for evaluating the measurement uncertainty / K.D. Sommer, O. K., F. P.e Leòn, B. R. L. Siebert // Proc. IEEE Intern.l Conf. on Multisensor Fusion and Integration for Intelligent Systems (Heidelberg, Germany, September 2006) . – P. 507–511.

100. Soodamani, R. A Fuzzy Interval Valued Fusion Technique for Multi-Modal 3D Face Recognition / R. Soodamani, U. M. Mariappan // Proc. IEEE Intern. Carnahan Conf. on Security Technology (Orlando, USA, October 2016). – P. 1-8.

101. Takruri, M. Data Fusion Techniques for Auto Calibration in Wireless Sensor Networks / M. Takruri, S. Challa, R. Yunis // Proc. 12th Intern. Conf. on Information Fusion Seattle (Seattle, USA, July 2009). – P. 132-139.

102. Tan, R. Adaptive Calibration for Fusion-based Wireless Sensor Networks / R. Tan, G. Xing, X. Liu, J. Yao, Zh. Yuan // Proc. IEEE INFOCOM (San Diego, USA, March 2010). – P. 11291695.

103. The On-Line Encyclopedia of Integer Sequences. The OEIS Foundation. https://oeis.org/A000217 (дата обращения: 25.03.17).

104. The On-Line Encyclopedia of Integer Sequences. The OEIS Foundation. https://oeis.org/A002662 (дата обращения: 25.03.17).

105. Thomas, R.R. Lectures in Geometric Combinatorics / R.R. Thomas // American Mathematical Society, 2006.

106. Van Deemen, A.M.A. The probability of the paradox of voting for weak preference orderings / A.M.A. Van Deemen // Social Choice and Welfare. – 1999. – Vol. 16. – P. 171-182.

107. Voorbraak, F. On the justification of Dempster's rule of combination / F. Voorbraak // Artificial Intelligence. – 1991. – Vol. 48. – P. 171-197.

108. Walley, P. Statistical Reasoning with Imprecise Probabilities / P. Walley // Chapman and Hall CRC: 1991. – 719 p.

109. Wang, P. A Defect in Dempster–Shafer Theory / P. Wang // Proc. 10th Conf. on Uncertainty in Artificial Intelligence (San Mateo, USA, 1994). – P 560-566.

110. Wang, Y.M. A preference aggregation method through the estimation of utility intervals / Y.M. Wang, J.B. Yang, D.L. Xu // Computers and Operations Research. – 2005. – Vol. 32. – № 8. – P. 2027-2049.

111. Weisstein, E.W. Triangular Number. MathWorld – A Wolfram Web Resource. http://mathworld.wolfram.com/TriangularNumber.html (дата обращения: 25.03.17).

112. White, F.E. Data Fusion Subpanel of the Joint Directors of Laboratories Technical Panel for C3 / F.E. White // Data Fusion Lexicon, NOSC, San Diego, USA , 1991.

113. Wilrich, P.-Th. Rounding of measurement values or derived values / P.-Th. Wilrich // Measurement. – 2005. – Vol. 37. – P. 21-30.

114. Wimmer, G. Proper rounding of the measurement results under normality assumptions / G. Wimmer, V. Witkovsk´y, T. Duby // Measurement Science and Technology. – 2000. – Vol. 11. – P. 1659-1665.

115. Wimmer, G. Proper rounding of the measurement results under the assumption of uniform distribution / G. Wimmer, V. Witkovsk´y // Measurement science review. – 2002. – Vol. 2. – Sec. 1. – P. 1-7.

116. Xu, N. A Wireless Sensor Network for Structural Monitoring / N. Xu, S. Rangwala, K.K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan, D. Estrin // Proc. 2nd Intern. Conf. on Embedded networked sensor systems (Baltimore, USA, November 2004). – P. 13-24.

117. Ye, W. An energy-efficient MAC protocol for wireless sensor networks / W. Ye, J. Heidemann, D. Estrin // Proc. IEEE INFOCOM (June 2002 New York, USA). – P. 7492167.

118. Young, H.P. Optimal voting rules / H.P. Young // Journal of Economic Perspectives. – 1995. – Vol. 9. – № 1. – P. 51-64.

119. Zhao, G. Wireless Sensor Networks for Industrial Process Monitoring and Control: A Survey / G. Zhao // Network Protocols and Algorithms. – 2011. – Vol. 3. – № 1. – P. 46-63.

120. Zhu, Y. Optimal interval estimation fusion based on sensor interval estimates with confidence degrees / Y. Zhu, B. Li // Automatica (Journal of IFAC). – 2006. – Vol. 42. – Iss. 1. – P. 101-108.

121. ZigBee Alliance. http://www.zigbee.org/ (дата обращения: 03.02.2017).

ПРИЛОЖЕНИЕ А

Акты внедрения результатов диссертационной работы

МИНОБРНАУКИ РОССИИ

**Федеральное государственное
автономное образовательное
учреждение высшего образования
«Национальный исследовательский
Томский государственный университет»
(ТГУ, НИ ТГУ)**

Ленина пр., 36, г. Томск, 634050
Тел. (3822) 52-98-52, факс (3822) 52-95-85
E-mail: rector@tsu.ru
http://www.tsu.ru
ОКПО 02069318, ОГРН 1027000853978
ИНН 7018012970, КПП 701701001

_____ № _____
на № _____ от _____

"УТВЕРЖДАЮ"
Проректор по научной работе
_____И.В. Ивонин
"___" _____2017 г.

А К Т
о внедрении результатов
кандидатской диссертации Худоноговой Л.И. "Комплексирование интервальных измерительных данных методом агрегирования предпочтений"
в лаборатории мониторинга окружающей среды Томского государственного университета

Комиссия в составе: Отмахов В.И., д.т.н., зав. лабораторией, Петрова Е.В., к.х.н., научный сотрудник лаборатории, Шелковников В.В., к.х.н., доцент кафедры аналитической химии, – составила настоящий акт в том, что результаты диссертационной работы Худоноговой Л.И.:

- метод комплексирования интервальных измерительных данных, состоящий в нахождении наилучшего дискретного значения в ранжировании консенсуса, найденном для набора наведенных интервалами ранжирований дискретных значений;
- алгоритм повышения точности результатов измерений на базе метода комплексирования.

используются в лаборатории мониторинга окружающей среды для обработки данных экологического мониторинга.

Реализованные в составе разработанного Худоноговой Л.И. программного обеспечения IntFusion метод комплексирования интервальных данных IF&PA и алгоритм SensAcc обеспечивают повышение точности результатов экологических измерений на 1-2 порядка при увеличении их робастности и достоверности по сравнению с традиционными методами.

Зав. лабораторией Отмахов В.И.

Научный сотрудник Петрова Е.В.

Доцент кафедры аналитической химии Шелковников В.В.

TOMSK POLYTECHNIC UNIVERSITY
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Ministry of Education and Science of the Russian Federation
Federal State Autonomous Educational Institution of Higher Education
"National Research Tomsk Polytechnic University" (TPU)
30, Lenin ave., Tomsk, 634050, Russia
Tel. (3822) 60 63 33, (3822) 70 17 79,
Fax (3822) 56 38 65, e-mail: tpu@tpu.ru, tpu.ru
OKPO (National Classification of Enterprises and Organizations): 02069303,
Company Number: 1027000890168,
VAT / KPP (Code of Reason for Registration) 7018007264/701701001, BIC 046902001

Министерство образования и науки Российской Федерации
федеральное государственное автономное образовательное учреждение высшего образования
«Национальный исследовательский Томский политехнический университет» (ТПУ)
Ленина, пр., д. 30, г. Томск, 634050, Россия
тел.: (3822) 60 63 33, (3822) 70 17 79,
факс: (3822) 56 38 65, e-mail: tpu@tpu.ru, tpu.ru
ОКПО 02069303, ОГРН 1027000890168,
ИНН/КПП 7018007264/701701001, БИК 046902001

_____ № _____

"УТВЕРЖДАЮ"
Директор Института кибернетики
_____С.А. Байдали
"26" 05 2017 г.

Акт

о внедрении в учебный процесс результатов диссертации Худоноговой Л.И. на тему: "Комплексирование интервальных измерительных данных методом агрегирования предпочтений", представленной на соискание ученой степени кандидата технических наук

Комиссия в составе: председателя – зав. кафедрой систем управления и мехатроники (СУМ), к.т.н., доцента Губина В.Е., и членов – д.т.н., профессора Рыбина Ю.К. и к.т.н., доцента Цимбалиста Э.И. – составила настоящий акт в том, что результаты диссертационной работы Худоноговой Л.И.

- программное обеспечение IntFusion, реализующее метод IF&PA, предназначенный для комплексирования интервальных измерительных данных на основе агрегирования предпочтений;
- алгоритм SensAcc повышения точности результатов измерений на базе метода комплексирования;
- алгоритм ActiveNode выбора подмножества активных узлов в кластере беспроводной сенсорной сети на основе метода IF&PA, обеспечивающий снижение энергопотребления узлов в кластере

используются при проведении лекционных, практических и лабораторных занятий по дисциплинам "Автоматизация измерений, контроля и испытаний", "Информационно-измерительные системы" и "Сенсорные сети". Результаты обеспечивают возможность применения инновационных технологий преподавания и повышают качество учебного процесса.

Зав. кафедрой СУМ		Губин В.Е.
Профессор кафедры СУМ		Рыбин Ю.К.
Доцент кафедры СУМ		Цимбалист Э.И.