

Новое направление в искусственном интеллекте: измерительный искусственный интеллект

С. В. Прокопчина

Финансовый университет при Правительстве РФ

svprokopchina @ mail.ru

Аннотация. Статья посвящена новому научному направлению: измерительному искусственному интеллекту. В работе даны основные определения и приведены атрибуты измерительного искусственного интеллекта (ИИИ). Определена область задач, в которых необходимо применение методов и средств этого направления. Приводятся примеры ИИИ на основе методологии регуляризирующего байесовского подхода, которые иллюстрируют реализацию принципов прослеживаемого, доверительного и объяснимого искусственного интеллекта и обеспечения устойчивости получаемых решений в условиях значительной неопределенности. Даны методологические основы метрологии решений систем искусственного интеллекта.

Ключевые слова: искусственный интеллект, метрология, регуляризирующий байесовский подход

I. ВВЕДЕНИЕ

Создание нового направления в искусственном интеллекте (ИИ), определяемого как измерительный искусственный интеллект (ИИИ), обусловлено насущной потребностью методологии искусственного интеллекта, связанной с осознанной в настоящее время потребностью получения надежных и достоверных решений систем ИИ в условиях значительной неопределенности (неточности, неполноты, нечеткости) исходной информации.

Времена полного доверия разработчиков средств ИИ к надежности исходной измерительной и статистической экспериментальной информации уже прошли. Основные требования со стороны разработчиков систем ИИ и практических пользователей включают, в том числе, требование обеспечения доверия к получаемым средствами ИИ решениям (доверительный искусственный интеллект). Подтверждением сказанному служит новая информация о принятых и разрабатываемых стандартах по ИИ серии ISO 42000, утвержденных международным институтом стандартов. Так, в стандарте ISO 42001 (Определения и терминология ИИ) подчеркивается необходимость определения качества и детального выбора данных для применения в системах искусственного интеллекта. Однако, особенно широкого выбора наборов надежных исходных данных нет по следующим причинам.

Во-первых, для основного спектра систем ИИ, таких например, как нейросети с обучением, требуются значительные массивы экспериментальных данных, которые на практике или трудно или вообще невозможно сформировать.

Во-вторых, степень надежности данных, особенно статистических, как правило, не определяется в эксперименте. Учитывая технологии их получения, чаще всего они являются субъективно искаженными и содержат разнообразные погрешности.

В третьих, получаемые в статистических экспериментах данные являются либо уникальными из-за уникальности и специфичности условий их получения, либо сгруппированы в малые выборки, что, безусловно, затрудняет их использование в системах ИИ.

В четвертых, в зависимости от типа информационных потоков неопределенность исходной информации, снижающая надежность исходных данных, имеет различную природу. Так, для измерительной информации ненадежность данных имеется за счет погрешности измерений, являющейся интеграцией погрешности методов измерений и погрешности приборов. То есть, речь идет о неточности измерений. Для неструктурированной информации неопределенность обусловлена природной непрерывностью потоков информации (тексты, декларативные знания, видео- и аудиоинформация). Поэтому в целях контроля ее достоверности и устойчивости получаемых решений требуются регуляризация и структуризация информационного пространства. Методологии и технологиям решений этой задачи посвящены работы [1, 5, 6].

В пятых, и может быть, особенно важных аспектах оценка качества (метрологическое обоснование) собственных решений систем ИИ, как правило, не разработано. Для получения метрологического обоснования, и, как следствие, получения высококачественных решений ИИ, важно иметь методологии и технологии как метрологического обоснования исходной информации, так и обоснования качества самих алгоритмов ИИ с целью контроля трансформации погрешности исходной информации через эти алгоритмические фильтры. Речь идет о метрологии получаемых решений в виде знаний, так как, в отличие от измерительных систем, в качестве получаемых решений в системах ИИ выступают знания.

Таким образом, для создания доверительного ИИ необходимы методологические разработки, соответствующие сказанному выше, то есть необходимо привлечение измерительного подхода в методологии ИИ на глубоком уровне методологической детализации и интеграции. Все это и обуславливает актуальность и необходимость нового направления с названием «измерительный искусственный интеллект».

С другой стороны, разработка современных измерительных систем достаточно давно идет по пути интеграции со средствами аналитической обработки результатов измерений. Это обусловлено значительным расширением круга измерительных задач и их усложнением. Использование в системах измерений методов ИИ дает возможность их эффективного решения.

Таким образом, сближение, интеграция и взаимопроникновение методологий ИИ и измерительной науки могут расцениваться как обоюдодоползные, и в настоящее время являются объективной реальностью и мощной перспективой их развития.

1. Определения и основные понятия ИИИИ.

Как было отмечено, в системах ИИ измерительный подход практически не используется, (используется только в работах школы байесовских интеллектуальных технологий [1, 2]). С другой стороны, современные тенденции развития теории и практики измерений направлены на интеграцию измерительных систем с системами искусственного интеллекта (ИИ). Это связано с необходимостью интеллектуальной обработки разнообразных потоков экспериментальной информации с целью получения полного решения задач пользователя в виде аналитических оценок, интерпретации состояния сложных объектов и ситуаций, а также в виде управленческих решений, выводов, рекомендаций. Однако многие примеры применения систем ИИ в измерительных системах, по сути, являются примерами механического соединения двух подсистем (систем ИИ и измерительных систем), функционирующих по законам своих методологий и технологий, которые не интегрированы в единую методологическую среду и единое информационное пространство. Поэтому, при переходе от систем измерений к системам вычислительной обработки данных возникает разрыв цепи информационных преобразований, который ведет к добавлению дополнительной, неконтролируемой энтропии в информационный процесс обработки данных, что вызывает неустойчивость получаемых решений и значительно снижает их качество.

Кроме того, при таком подходе к созданию интеллектуальных средств измерений не обеспечиваются возможности прослеживаемости, прозрачности (transparency), интерпретируемости (explainability) надежности и доверительности (trustworthiness), обеспечения устойчивости (stability) получаемых измерительных решений, что является основным перечнем современных требований к системам интеллектуальной обработки данных. Таким образом, с целью обеспечения указанных требований должны быть использованы и разработаны новые подходы, методология которых дает возможность реализации этих требований на основе глубокой интеграции методологий ИИ и теории измерений с целью достижения синергетического эффекта повышения мощности и эффективности систем.

Такое направление развития измерительных систем, то есть измерительных систем, основанных на глубокой связанности методологий и технологий ИИ и теории измерений при полном метрологическом обосновании

получаемых решений, определено в [7], как измерительный искусственный интеллект (Measuring Artificial Intelligence).

При формировании общей терминологии для систем ИИИИ целесообразно обратиться к терминам и определениям интеллектуальных измерений.

Трактовки основных понятий теории интеллектуальных измерений и метрологии приведены в [1]. Их появление и развитие относится к 90-м годам прошлого столетия [4].

Основным принципом, реализуемым во всех указанных в этих работах типах измерений в ситуации неопределенности, является принцип интеллектуализации измерений на основе привлечения знаний в измерительные процессы.

По терминологии, данной в [1], измерительные процессы, связанные с совместной обработкой данных и знаний и получением на основе этой интеграции метрологически обоснованных решений, называются интеллектуальными измерениями (ИИИ).

Методология ИИИ может быть успешно использована для создания теории и практической реализации систем ИИИИ.

Объектами измерений в ИИИ являются сложные системы, активно взаимодействующие с внешней средой. Таким образом, переходя к терминологии ИИИИ можно определить следующие понятия ИИИИ.

Модель объекта ИИИИ представляет собой сложную иерархическую систему взаимосвязанных факторов, отражающих свойства самого объекта и факторов влияния внешней среды. На основе этой концепции строится рабочая модель объекта ИИИИ, которая в системах ИИИИ отображается в виде иерархической модели дерева свойств сложного объекта и среды его функционирования, построенная на основе системного подхода (цифровой двойник реального объекта). При этом рабочая модель отличается от концептуальной модели ИИИИ количеством оцениваемых свойств объекта и среды. В рабочей модели их может быть существенно меньше из-за отсутствия необходимой исходной информации. Однако наличие в концептуальной модели более широкого списка показателей задает основу для развития модели в процессе ее использования в системах ИИИИ. Методы и технологии реализации таких моделей основаны на теории байесовских интеллектуальных технологий. [1]. Такие модели названы моделями с динамическими ограничениями ИИИИ.

Неструктурированность исходной информации определяет необходимость выбора и применения специальных методов, ориентированных на указанные условия и обеспечивающих качество и устойчивость измерительных решений ИИИИ.

Под измерительным решением ИИИИ здесь понимается разнотипный (выраженный в числовой или неколичественной формах) результат решения прикладной задачи в виде численного значения свойства, степени его проявления в виде лингвистической оценки, вывода, рекомендации с определенными метрологическими показателями качества этого решения

(точности, надежности, достоверности, риска, количества полученной новой информации и других). Особое внимание надо уделить вопросам устойчивости измерительных решений ИИ. Этот вопрос является одним из самых актуальных для развития методов и средств ИИ.

Именно для решения задачи обеспечения устойчивости в условиях неопределенности разработан регуляризирующий байесовский подход (РБП). Его создание относится к началу 90-х годов прошлого столетия. К настоящему времени на его основе и реализующих его технологий и средств решены задачи в сферах индустрии, энергетики, экологии, экономики, регионального и социального развития, природной среды [1, 3, 4, 6].

Подробное теоретическое обоснование РБП, как и методологические основы обеспечения устойчивости решений приведены в работах автора [1, 2, 5]. В данной статье целесообразно отметить, что для обеспечения устойчивости применяется система специально разработанных шкал, названных шкалами с динамическими ограничениями (ШДО), методы построения и применения которых детально рассмотрены в ряде работ, например в [1, 4]. При шкалировании исходной информации на ШДО происходит регуляризация решений.

II. Основные принципы методологии измерительного искусственного интеллекта на основе регуляризирующего байесовского подхода и байесовских интеллектуальных измерений приведены ниже.

1. Реализация вычислительного процесса производится в рамках единого регуляризованного пространства решений (метрических динамических компактов измерительных решений, определения и теоретическое обоснование которых дано в [1]).
2. Обеспечение устойчивости получаемых измерительных решений достигается на основе шкалирования исходной информации.
3. Использование и интеграция разнотипной информации (числовой и лингвистической) на основе модифицированной формулы Байеса [1].
4. Прослеживаемость процесса получения измерительных решений обеспечивается путем последовательной обработки информации на шкалах типа ШДО, согласно иерархической структуре модели объекта.
5. Интерпретация и объяснимость измерительных решений (объяснимый искусственный интеллект (ИИ) на каждом этапе обработки информации).
6. Контроль уровня качества (точности, надежности, достоверности) измерительных решений ИИИ на методологии БИТ для метрологического обоснования и сопровождения измерительных решений ИИИ (доверительный ИИ).
7. Динамичность и развиваемость моделей объектов и моделей ИИИ.

8. Когнитивность и визуализация решений.
9. Возможность использования, обработки и интеграции (свертки) как числовой, так и качественной, лингвистической неструктурированной информации.
10. Условность полученных решений ИИИ (в виде моделей, рекомендаций, оценок, выводов, сценариев).
11. Учет условий получения исходной информации и реализации вычислений в технологиях и средствах ИИИ.

III. Цифровая платформа «Инфоаналитик» как средство создания систем измерительного искусственного интеллекта.

Информационные технологии ИИИ на основе РБП позволяют производить измерение и аналитическую обработку любой количественной и качественной информации в числовом, лингвистическом, графическом, картографическом видах с требуемой точностью, надежностью, достоверностью.

На основе методологии РБП и реализации технологий БИИ и БИТ разработана компьютерная платформа «Инфоаналитик», предназначенная для быстрой разработки интеллектуальных измерительных систем, которые служат примерами реализации ИИИ.

Платформа «Инфоаналитик» представляет собой измерительную нейросеть сверточного типа с обучением [2]. При обучении сети производится ее настройка на модели и условия прикладной задачи. Исходная информация поступает от внешних источников: измерительных приборов, экспертных сообщений, интернет-информации. Номенклатура внешних источников зависит от типа решаемой прикладной задачи.

На этапе предварительной подготовки в режиме обучения сети средствами платформы разрабатываются специальные типы шкал для каждого источника информации, которые обеспечивают интеграцию платформы с источниками информации согласно методикам интеграции РБП. Такие системы могут быть разработаны специалистами-экспертными прикладной задачи без участия программистов – разработчиков.

Круг решаемых прикладных задач, а также конкретных систем ИИИ достаточно широк. Применение платформы «Инфоаналитик» для решения прикладных задач показано на нескольких нижеследующих и приведенных в [1, 3, 4, 5, 6] примерах.

На основе методологии и средств РБП создана интеллектуальная система для мониторинга состояния и управления сетью горячего водоснабжения (ГВС).

Основными задачами такой системы являются:

1. Сбор и предварительная обработка (с целью сжатия информации) экспериментальной информации от приборов учета и контроля.
2. Сбор нечисловой информации и знаний о состоянии сети ГВС.

3. Математическая обработка данных с выдачей прогноза надежности сетей ГВС.
4. Когнитивная визуализация оценок состояния и прогноза ГВС.
5. Генерация управленческих рекомендаций и инструкций.
6. Подготовка обоснования для реконструкции и плановой замены трубопроводов.
7. Сбор информации для составления схем трубопроводов и паспортизации сетей ГВС.

Результатами работы сетей являются комплексные оценки состояния сложных объектов ГВС и рекомендации по обеспечению их устойчивого функционирования. Важной частью таких систем являются встроенные средства полного метрологического обоснования всех получаемых решений. Системы имеют иерархическую архитектуру, согласно уровням управления сложными объектами, которая обладает возможностью саморазвития на основе вновь получаемой информации. Это достигается, благодаря моделям и шкалам с динамическими ограничениями, на основе которых строятся все алгоритмы системы. Реализованы примеры использования интеллектуальных сенсорных сетей для мониторинга и управления энергогенерирующими и водоснабжающими системами. Такие интеллектуальные измерительные сети позволяют производить комплексную оценку состояния распределенных систем в любой момент времени, определить основные риски и потенциалы как отдельных участков, так и системы в целом. Кроме того, можно привести примеры использования ИИИ в задачах управления развитием региональной экономики, производственных организаций, управления персоналом (оценка мотивации персонала и комфортности корпоративного климата), управления социальными системами и иные примеры, опубликованные в [1–7] и других работах.

II. ЗАКЛЮЧЕНИЕ

Новое направление развития методологии и систем ИИ, связанное с привлечением измерительного подхода, которое в работе определено как измерительный искусственный интеллект, позволяет значительно расширить сферу прикладных задач и применения методов и технологий ИИ, обеспечить повышение и контроль качества получаемых измерительных решений в сложных условиях информационной неопределенности.

Применение РБП в измерительных задачах ИИИ позволяет создавать интеллектуальные измерительные системы с обеспечением прозрачности, интерпретируемости, достоверности, устойчивости получаемых решений.

СПИСОК ЛИТЕРАТУРЫ

- [1] Прокопчина С.В. Интеллектуальные измерения на основе регуляризирующего байесовского подхода. Изд-во «Научная библиотека», 2021, 595 с.
- [2] Прокопчина С.В. Новый тип нейронных сетей: байесовские измерительные нейронные сети (BIN), основанные на методологии регуляризирующего байесовского подхода // Мягкие измерения и вычисления. 2020. Том 35. № 10, с. 17-24.
- [3] Прокопчина С.В. Реализация принципов объяснимого искусственного интеллекта в информационных системах на основе регуляризирующего байесовского подхода // Мягкие измерения и вычисления. 2022. №6, с. 7-15.
- [4] Прокопчина С.В. Регуляризация как основа обеспечения устойчивости решений в системах искусственного интеллекта (на примере регуляризирующего байесовского подхода) // Мягкие измерения и вычисления, 2023. № 7, с. 7-18.
- [5] Прокопчина С.В. Измерение «неизмеримых» показателей на основе регуляризирующего байесовского подхода // Мягкие измерения и вычисления. 2022. № 6, с. 7-15.
- [6] Прокопчина С.В. Методология создания «доверительного» искусственного интеллекта в системах на основе регуляризирующего байесовского подхода // Мягкие измерения и вычисления. 2023. № 8, с. 17-28.
- [7] Прокопчина С.В. Измерительный искусственный интеллект: методологические основы, принципы создания и информационные технологии на основе регуляризирующего байесовского подхода // Мягкие измерения и вычисления. 2023. № 10-2, с. 23-36.

Методика метрологического анализа измерительной ситуации, допускающей учет функциональных взаимосвязей между измеряемыми величинами, и примеры ее применения

В. А. Гаранин

Санкт-Петербургский политехнический университет Петра Великого

garanin_va@spbstu.ru

Аннотация. Работа посвящена метрологическому анализу ситуаций совместного измерения значений нескольких искомых величин, для которых существуют и априорно известны такие зависимости между ними, что по крайней мере для одной из измеряемых величин могут быть с их помощью получены одна или более независимые оценки из значений других измеряемых величин. Для расчета пределов возможной погрешности результатов обработки таких совместных измерений, с учетом упомянутых априорных сведений, предложена универсальная методика, позволяющая принять во внимание неопределенности результатов выполненных измерений и параметров уравнений, описывающих известные взаимосвязи. Разработанная методика применима для случаев, когда взаимосвязи между величинами выражены в виде одного или нескольких уравнений. Приведены примеры выполнения метрологического анализа характерных измерительных ситуаций с ее помощью.

Ключевые слова: метрологический анализ; совместные измерения; оценка погрешностей; согласование результатов измерений; взаимосвязи между измеряемыми величинами

I. ВВЕДЕНИЕ

При выполнении совместных измерений значений физических величин возникает возможность осуществить уточнение результатов их математической обработки, если в наличии есть дополнительная априорная информация об измерительной ситуации. Наиболее частым способом формализации такого рода сведений является представление их в виде системы уравнений или неравенств. Указанную процедуру перевода дополнительной информации в уточнение конечных результатов математической обработки результатов измерений в литературе принято называть согласованием (неточных) данных [1]. Цель этой операции заключается в получении эффективных в статистическом смысле оценок значений искомых величин на основе результатов их совместного измерения, не противоречащих известной

математической модели объекта или процесса, для которых выполнены данные измерения.

Как правило, процедура согласования сводится к математической задаче отыскания условного оптимума: введение в задачу обработки результатов совместных измерений дополнительных сведений (математической модели) очерчивает границы множества возможных решений, в пределах которого необходимо найти оптимальное, учитывая погрешности всех данных – как результатов измерений, так и параметров модели.

Для отыскания оценок значений измеряемых величин, получаемых в ходе процедуры согласования, на практике используют как методы аналитической [1, 2], так и численной оптимизации [3–6], общим для которых является построение на принципе максимального правдоподобия и привлечение предположения о том, что случайные погрешности результатов измерений независимы и подчинены нормальному закону.

Обычно метрологический анализ, подразумевающий выполнение расчета погрешностей результатов согласования, на промышленных объектах выполняется двумя способами: до монтирования измерительной системы – с помощью статистического эксперимента, после – в рамках поверочных работ, подразумевающих сравнение результатов с известной мерой или с результатами от образцового средства измерения.

До недавнего времени аналитические методы оценки погрешности результата согласования с использованием сведений о неопределенности исходных данных и параметров учитываемой математической модели были разработаны только для линейных задач [1, 7] в предположении наличия у результатов измерений только несмещенных и некоррелированных случайных погрешностей. В работах [8–12] была предпринята попытка построения универсального подхода к метрологическому анализу результатов согласования совместных измерений для нелинейных задач и при произвольном распределении погрешностей результатов измерений. Данная работа содержит представление обобщенной методики метрологического анализа результатов согласования, основанной на упомянутом подходе и свободной от ограничений, ранее предложенных в литературе методов.

которая в общем случае решается методом наименьших квадратов.

III. МЕТОДИКА РАСЧЕТА ПОГРЕШНОСТЕЙ РЕЗУЛЬТАТОВ СОГЛАСОВАНИЯ ЗНАЧЕНИЙ СОВМЕСТНО ИЗМЕРЕННЫХ ВЗАИМОСВЯЗАННЫХ ВЕЛИЧИН

С использованием описанных подходов была разработана методика для предварительного метрологического анализа измерительной ситуации, допускающей учет взаимосвязи между измеряемыми величинами, состоящая из следующих шагов.

1) Формализуем априорную информацию о зависимостях между значениями подлежащих измерению величин в виде системы алгебраических уравнений

$$\mathbf{F}_M(\mathbf{x}, \mathbf{a}) = \mathbf{0}_m,$$

где $\mathbf{F}_M = (f_1, f_2, \dots, f_m)^T$ – вектор-функция, элементы которого являются достаточно гладкими функциями; $\mathbf{a}^T = (\mathbf{a}_1^T, \mathbf{a}_2^T, \dots, \mathbf{a}_m^T)$ – вектор параметров математической модели, составленный из векторов параметров для каждой функции, входящей в \mathbf{F}_M ; $\mathbf{0}_m$ – вектор-столбец из m элементов, заполненный нулями; вектор $\mathbf{x} = (x_1, x_2, \dots, x_N)^T$ содержит подлежащие измерению величины.

2) Если измерительная ситуация предполагает многократное совместное измерение взаимосвязанных величин \mathbf{x} , то тогда для каждой из измеряемых величин следует рассчитать оценку ее значения по полученным результатам ее многократных измерений

$$y_j = EV(\hat{\mathbf{x}}_j), \quad j = 1, 2, \dots, N,$$

где EV – выбранный алгоритм оценки значения величины x_j по результатам $\hat{\mathbf{x}}_j = (\hat{x}_{j1}, \hat{x}_{j2}, \dots, \hat{x}_{jn})^T$ ее многократного измерения, n – количество многократных измерений. При однократном измерении ($n = 1$) имеем $y_j = \hat{x}_j$.

3) Оцениваем среднеквадратическое отклонение σ_{y_j} случайной погрешности полученных оценок y_j измерения величин x_j :

- при однократном измерении $\sigma_{y_j} = \sigma_{\hat{x}_j}$,
- если алгоритм EV представляет собой вычисление среднеарифметического, то $\sigma_{y_j} = \sigma_{\hat{x}_j} / \sqrt{n}$.

4) Рассчитываем оценку коэффициента k_j , указывающего на потенциально возможное уточнение за счет процедуры согласования по соотношению

$$k_j^2 = 1 + \sigma_{y_j}^2 / \sigma_{\hat{x}_j}^2,$$

где для случая одного уравнения взаимосвязи ($m = 1$) дисперсия $\sigma_{\hat{x}_j}^2$ находится из уравнения (2), а для случая системы уравнений $\mathbf{F}_M(\mathbf{x}, \mathbf{a}) = \mathbf{0}_m$ – из системы уравнений (3).

Для расчетов в данном пункте методики могут быть использованы программы [13, 14].

Если все значения $k_j < 1,2$, $j = 1, 2, \dots, N$, то использование процедуры согласования данных нецелесообразно, поскольку достигаемое уточнение

оказывается на уровне ошибок округления, которому традиционно подвергаются метрологические характеристики результатов измерений.

5) Если это необходимо, вычисляем границы интервала I_{k_j} возможных значений коэффициента k_j как

$$I_{k_j} = \left[\min_{q=1,2} k_{jq}, \max_{q=1,2} k_{jq} \right],$$

$$k_{j1,2}^2 = 1 + \frac{1}{\sigma_{y_j}^2} \cdot \left(\left(1 \pm \frac{\delta_j}{100\%} \right) \cdot \sigma_{\hat{x}_j} \right)^2,$$

где δ_j – относительная доля (%) ожидаемой неучтенной нелинейности вектор-функции $\mathbf{F}_M(\mathbf{x}, \mathbf{a})$ по значению x_j в окрестностях точки \mathbf{t} , образованных пределами погрешности ее компонент, т.е. $\mathbf{t} \pm \Delta \mathbf{t}$. В первом приближении мера нелинейности δ_j , %, может быть задана исходя из общих соображений о характере функциональной связи между измеряемыми величинами.

IV. ПРИМЕРЫ ПРИМЕНЕНИЯ

Пример 1. Рассмотрим следующую измерительную ситуацию. Пусть выполнено $n = 100$ совместных независимых измерений величин $\mathbf{x} = (x_1, x_2)^T$, значения которых связаны единственным уравнением $\mathbf{F}_M(\mathbf{x}, \mathbf{a}) = (f_1) = (x_1 \cdot a - x_2) = (0)$, где вектор параметров $\mathbf{a} = (a)$. Неопределенность результатов измерений \hat{x}_{1i} и \hat{x}_{2i} , $i = 1, 2, \dots, n$, задана через оценки дисперсий $s_{\hat{x}_1}^2$ и $s_{\hat{x}_2}^2$.

Положим, что выполненное количество измерений достаточно велико, чтобы без возникновения значимой для конечных результатов погрешности использовать в формулах в качестве действительных значений среднеквадратических отклонений их результатов измерений искомым величин значения их оценок s (при указанном n величина соответствующей статистической погрешности не превышает с доверительной вероятностью 0,95 значения 15%).

Пусть x_1 и x_2 – это величины напряжения на входе и выходе некоторого четырехполюсника, величина a – коэффициент передачи этого четырехполюсника, а погрешности результатов измерения значений величин x_1 и x_2 носят случайный характер и подчинены нормальному закону. В этом случае согласно методу максимального правдоподобия эффективными оценками значений искомым величин x_1 и x_2 будут являться соответствующие среднеарифметические y_1 и y_2 от результатов измерения каждой из них, т.е.

$$y_j = EV(\hat{\mathbf{x}}_j) = \frac{1}{n} \cdot \sum_{i=1}^n \hat{x}_{ji}, \quad j = 1, 2.$$

Необходимо определить степень потенциального уточнения результатов измерения искомой величины x_1 за счет учета дополнительной априорной информации о взаимосвязи между x_1 и x_2 в сравнении с результатом $y_1 = EV(\hat{\mathbf{x}}_1)$.

В результате измерительного эксперимента с последующей обработкой полученных результатов прямых измерений по алгоритму EV получены следующие численные значения (при $a = 2,00$):

$$y_1 = 2,00 \text{ В}; y_2 = 4,00 \text{ В}; s_{y_1}^2 = 0,04 \text{ В}^2; s_{y_2}^2 = 0,16 \text{ В}^2.$$

Полагая, что функциональная связь между величинами x_1 и x_2 известна точно, запишем для него уравнение (1):

$$(\partial f_1 / \partial x_1)^2 \cdot \sigma_{\hat{x}_1}^2 \approx (\partial f_1 / \partial x_2)^2 \cdot s_{y_2}^2.$$

Отсюда

$$\sigma_{\hat{x}_1}^2 \approx \left(\frac{\partial f_1}{\partial x_2} / \frac{\partial f_1}{\partial x_1} \right)^2 \cdot s_{y_2}^2 = \frac{s_{y_2}^2}{a^2}.$$

Оценим степень уточнения k_1 оценки значения величины x_1 , достигаемого за счет учета априорной информации о взаимосвязи между x_1 и x_2 :

$$k_1 = \sqrt{1 + \frac{s_{y_1}^2}{\sigma_{\hat{x}_1}^2}} = \sqrt{1 + \frac{s_{y_1}^2 \cdot a^2}{s_{y_2}^2}} = \sqrt{1 + \frac{0,02^2 \cdot 2^2}{0,04^2}} = \sqrt{2}.$$

Полученное значение обусловлено выполнением равенства $s_{y_1} / y_1 = s_{y_2} / y_2$ в рассмотренном примере и по сути дела соответствует ситуации уменьшения неопределенности при усреднении двух равнозначных результатов измерения одной величины.

Пример 2. Пусть выполнено однократное совместное измерение массового расхода вещества в нескольких трубопроводах, сходящихся в один общий канал. Измеряемые величины связаны уравнением:

$$F_M(\mathbf{x}, \emptyset) = (f_1) = (x_0 - \sum_{j=1}^4 x_j) = (0),$$

где x_0 – результат измерения массового расхода вещества в общем трубопроводе, x_1, \dots, x_4 – результаты измерения расхода в остальных трубопроводах.

В ходе измерений получены следующие результаты:

$$\hat{x}_0 = 12,60 \text{ (кг/мин)}; \hat{x}_1 = 6,60 \text{ (кг/мин)}; \hat{x}_2 = 1,50 \text{ (кг/мин)}; \\ \hat{x}_3 = 1,50 \text{ (кг/мин)}; \hat{x}_4 = 3,00 \text{ (кг/мин)}.$$

Априорно известно, что случайная составляющая погрешности результатов измерения расхода для каждого датчика не превышает абсолютных значений:

$$\Delta_{\hat{x}_0} = 60 \text{ (г/мин)}; \Delta_{\hat{x}_1} = 35 \text{ (г/мин)}; \Delta_{\hat{x}_2} = 8 \text{ (г/мин)}; \\ \Delta_{\hat{x}_3} = 15 \text{ (г/мин)}; \Delta_{\hat{x}_4} = 30 \text{ (г/мин)},$$

где Δ обозначает полуширину доверительного интервала одинаковой для всех датчиков вероятностной меры $Q = 0,95$.

Определим значение k_1 – величины возможного уточнения результата измерения \hat{x}_1 массового расхода вещества в первом трубопроводе за счет присутствующей дополнительной информации. Будем полагать, что погрешность модели взаимосвязи между измеряемыми величинами пренебрежимо мала.

Определим предельное значение $\Delta_{\hat{x}_1}$ через уравнение (1), используя то, что $\Delta_{\hat{x}_j} / \sigma_{\hat{x}_j} = \text{const}(j), j = 0, 1, \dots, 4$:

$$(\partial f_1 / \partial x_1)^2 \cdot \Delta_{\hat{x}_1}^2 \approx \sum_{j \in \{0,2,3,4\}} (\partial f_1 / \partial x_j)^2 \cdot \Delta_{\hat{x}_j}^2.$$

Отсюда получаем, что $\Delta_{\hat{x}_1}^2 \approx 75 \text{ (г/мин)}$.

Поскольку измерения однократные, то $y_1 = \hat{x}_1$ и согласно (1) получаем, что $k_1 = 1,21$ (раза).

Таким образом, рассмотренная измерительная ситуация допускает уточнение результатов прямого измерения массового расхода в первой трубе до 1,2 раз. Небольшой потенциал уточнения обусловлен относительно высокой неопределенностью результатов выполняемых измерений.

Представленные в данном разделе результаты находятся в хорошем соответствии с результатами расчетов, получаемых при использовании более сложных моделей [15-17] согласования результатов совместных измерений (полу- и полностью непараметрических) для нелинейных задач, характеризующих измерительные ситуации.

V. ЗАКЛЮЧЕНИЕ

В данной работе предложена методика метрологического анализа измерительной ситуации, допускающей учет функциональных взаимосвязей между измеряемыми величинами для стационарных задач с достаточно гладкими функциями упомянутых взаимосвязей. Представленная методика сочетает в себе вычислительную простоту и достаточную с метрологической точки зрения достоверность получаемых оценок. В работе приведены примеры ее применения для нескольких измерительных ситуаций.

БЛАГОДАРНОСТИ

Автор статьи благодарит Семенова Константина Константиновича (СПбПУ) за плодотворные обсуждения и полезные рекомендации.

СПИСОК ЛИТЕРАТУРЫ

- [1] L.K. Reznik and G.N. Solopchenko, "Use of priori information on functional relations between measured quantities for improving accuracy of measurement," // Measurement, vol. 3, pp. 98-106, 1985.
- [2] M.J. Bagajewicz and E. Cabrera, "Data reconciliation in gas pipeline systems," // Industrial & engineering chemistry research, vol. 42(22), pp. 5596-5606, 2003.
- [3] S. Bai, J. Thibault and D.D. McLean, "Dynamic data reconciliation: Alternative to Kalman filter," // Journal of Process Control, vol. 16(5), pp. 485-498, 2006.
- [4] D.M. Prata, M. Schwaab, E.L. Lima and J.C. Pinto, "Nonlinear dynamic data reconciliation and parameter estimation through particle swarm optimization: Application for an industrial polypropylene reactor," // Chemical Engineering Science, vol. 64(18), pp. 3953-3967, 2009.
- [5] D. Wang and J.A. Romagnoli, "A framework for robust data reconciliation based on a generalized objective function," // Industrial & engineering chemistry research, vol. 42(13), pp. 3075-3084, 2003.
- [6] D. Wang and J.A. Romagnoli, "Robust data reconciliation based on a generalized objective function," // IFAC Proceedings Volumes, vol. 35(1), pp. 191-196, 2002.
- [7] S. Narasimhan and C. Jordache, Data reconciliation and gross error detection: An intelligent use of process data. Houston, Texas: Gulf Publ. Comp., 2000.
- [8] V.A. Garinin and K.K. Semenov, "The systematic approach for estimating the potential increase of measurement results accuracy achieved by the use of dependencies between measurands," // Proc. of the XXIV International Conference on Soft Computing and Measurements (SCM-2021), pp. 15-18, 2021.
- [9] V.A. Garinin and K.K. Semenov, "Measured Data Reconciliation Processed within Nonlinear Constraints Models in Cyber-Physical Systems," // Lecture Notes in Networks and Systems, vol. 460, pp. 78-95, 2023.
- [10] V.A. Garinin and K.K. Semenov, "Estimating the Accuracy Increase During the Measuring Two Quantities with Linear Dependence," // Lecture Notes in Networks and Systems, vol. 95, pp. 235-246, 2020.

- [11] V.A. Garanin and K.K. Semenov, "Semi-nonparametric approach for measured data reconciliation based on the Gram-Charlier series expansion," // Measurement: Sensors, vol. 18, paper 100351, 2021.
- [12] V.A. Garanin and K.K. Semenov, "Evaluation of the Potential Refinement of the Results of Multiple Measurements with Asymmetric Distribution, Achieved by Considering the Relationships between the Measurands," // Proc. of the XXVI International Conference on Soft Computing and Measurements (SCM-2023), pp. 32-36, 2023.
- [13] Свид. о регистрации программы для ЭВМ RU 2023617421 / А.А. Целищева, К.К. Семенов, В.А. Гаранин. Программа для вычисления приближенной оценки погрешности корней нелинейных уравнений с неточными параметрами на основе локальной полиномиальной аппроксимации. Заявка № 2023615897 от 30.03.2023. Оpubл. 10.04.2023.
- [14] Свид. о регистрации программы для ЭВМ RU 2023617702 / А.А. Целищева, К.К. Семенов, В.А. Гаранин. Программа для вычисления приближенной оценки погрешности решений систем нелинейных уравнений с неточными параметрами на основе локальной полиномиальной аппроксимации. Заявка № 2023615916 от 30.03.2023. Оpubл. 12.04.2023.
- [15] Свид. о регистрации программы для ЭВМ RU 2023610445 / В.А. Гаранин, К.К. Семенов. Программа для непараметрического согласования результатов совместных измерений взаимосвязанных величин с использованием метода проекционного восстановления плотности распределения вероятностей. Заявка № 2022686089 от 27.12.2022. Оpubл. 11.01.2023.
- [16] Свид. о регистрации программы для ЭВМ RU 2023610961 / В.А. Гаранин, К.К. Семенов. Программа для оценки повышения точности измерений за счет учета известных функциональных взаимосвязей между измеряемыми величинами. Заявка № 2022686101 от 27.12.2022. Оpubл. 16.01.2023.
- [17] Свид. о регистрации программы для ЭВМ RU 2023611165 / В.А. Гаранин. Программа для непараметрического согласования результатов совместных измерений взаимосвязанных величин с использованием метода ядерного восстановления плотности распределения вероятностей. Заявка № 2022686069 от 27.12.2022. Оpubл. 17.01.2023.

Assessment and Modeling of the Sewerage Network Using SewerGEMS Software in the City Center of Karbala, Iraq

Ihsan Kadhim Abed

*Lebanese University Faculty Of
Engineering, Civil Engineering*
Beirut, Lebanon
Ihsan.abed@ul.edu.Lb

Fadi Hage Chehade

*Lebanese University Faculty Of
Engineering, Civil Engineering*
Beirut, Lebanon
fadihagechehade@hotmail.com

Zaid Farhood Maki

*Lebanese University Faculty Of
Engineering, Civil Engineering*
Beirut, Lebanon
makkizaid@gmail.com

Abstract—The sewerage system is a vital element of urban infrastructure. Environmental pollution can be attributed to the presence of floodwater caused by pipe blockages, pipe breaks, or inadequate drainage capacity. In recent decades, there has been a significant surge in both the population size and the pace of population growth. As a result, urbanisation and industrialization have experienced a rise. Recent patterns indicate a growth in urban areas. As cities grow, it becomes necessary for the Municipality to provide essential public service infrastructure, like drinking water pipelines, sewer systems, street lighting, and roadways. An example of a public service institution is sewer networks, which play a crucial role in maintaining public health and cleanliness. The future entails the implementation of an efficient drainage system to mitigate the risk of flooding caused by abrupt population surges during festivals or religious visits. In this research, we conduct a modeling of the sewerage network previously implemented in the historic city center of Karbala - Iraq, using sewer gems. This region is characterized by attracting many visitors to its religious shrines and since this network was implemented during the year 1980 for a limited number of residents, this area suffers from frequent sewage overflows during the pilgrim because it cannot accommodate the increasing numbers of residents, which also causes environmental pollution and obstruction of traffic[reference: Karbala Sewerage Directorate].

Keywords—*Sewer Network, Sewege overflow, SewerGEMS, manhole, population, Karbala*

I. INTRODUCTION

Sewerage systems are an essential component of the infrastructure of any society. The primary objective of implementing a sewer network is to efficiently transport sanitary waste from an urban region, ensuring that it does not pose any public health concerns. The installation of a sewerage network transports the wastewater generated by individuals, business enterprises, and industrial establishments to facilities where it can be treated before being released back into the natural environment [1]. The expense associated with sewerage projects is considerable, but this problem can be solved through the use of computer software such as SewerGEMS. A sewage network model was developed in the city center of Karbala, Iraq, to address the growing population during the Pilgrim season. The model predicts the amount of people and the timing of wastewater collection based on the area's zoning rules. The SewerGEMS software automatically determines pipe diameters based on the flow velocity and gradient needs of the pipe. The software is used to determine parameters such as discharge, velocity, ground level, inversion level, and depth of cut. The

computer software package SewerGEMS is superior to manual approaches for building an economical sewer network. It provides an optimal cost and practical plan that can accommodate a vast network.

II. LITERATURE REVIEW

The sewerage network serves as the fundamental component in the design. Cost savings made during the design phase of this unit will have an impact on the total cost of the sewerage system. Bentley SewerGEMS v8i is a unique and comprehensive software that enables dynamic modelling of sanitary and combined sewer systems across several platforms, including GIS and CAD [2]. SewerGEMS is a software that simplifies the modelling process by allowing more time to solve wastewater engineering challenges. It also increases capacity and reduces sewer overflows, helping utilities meet sewer design criteria established by regulatory authorities. The software offers sophisticated engineering capabilities for the design, planning, maintenance, and operation of sanitary sewage systems [3].

Comparing the SewerGEMS computer software to manual design methods reveals its economic advantages, as it offers cost optimisation and realistic planning capabilities for sewerage networks [4], [5]. The SewerGEMS software is versatile, allowing the creation of a numerical model of the sewage system using geographic information systems [6]. In addition, I utilise SewerGEMS in conjunction with AutoCAD to assess the functionality of a sewage network system and propose appropriate corrective actions [6]. Nishant Sourabh [7] similarly examined the hydraulic state of the SVNIT campus, identifying areas of excessive flow and offering recommended solutions and measures.

III. METHODOLOGY

This paper will present a modeling work for the sewerage network implemented in the past in the center of the historic religious city of Karbala, using the SewerGEMS for the purpose of treating the recurring occurrence of sewage overflows during the Pilgrimage because it does not accommodate the excess drainage resulting from the population increase for which the network was designed. Fig. 1, shows the work steps.

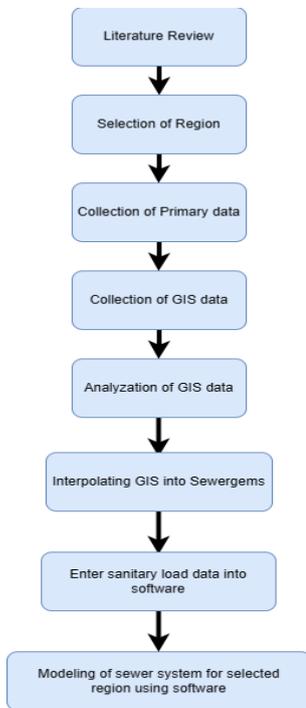


Fig. 1. The stages of research methodology work

IV. DESCRIPTION OF THE STUDY AREA

Karbala, spanning 5,034 square kilometres, is a city in Iraq situated approximately 100 kilometres (62 miles) southwest of Baghdad. Its precise geographical coordinates are Latitude: 32°36'51" N, Longitude: 044°01'29" E. Karbala consists of two districts, namely "Old Karbala" and "New Karbala". The residential district encompasses Islamic schools and government facilities. The city centre of Karbala has an approximate population of 495,411, whereas the entire Karbala Governorate has a population of 975,000 (City Population, 2018) [8]. Fig. 2 depicts a map of Iraq, specifically highlighting the city of Karbala along with neighbouring cities.

Karbala is renowned as a significant religious city that attracts Muslims worldwide at specified annual periods. The city experiences the highest influx of visitors during the Arbaeen visit, which occurs on the twentieth day of the month of Safar in the Hijri calendar. Fig. 3, displays the visitor count to the governorate over a period of twenty days (from 1st Safar to 20th Safar) for the past seven years [9].

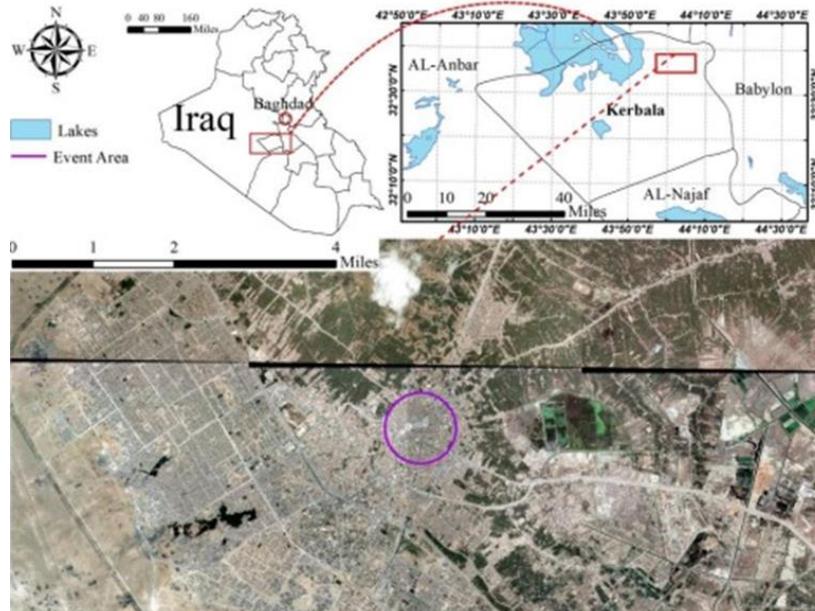


Fig. 2. the location of Karbala Governorate on a map of Iraq

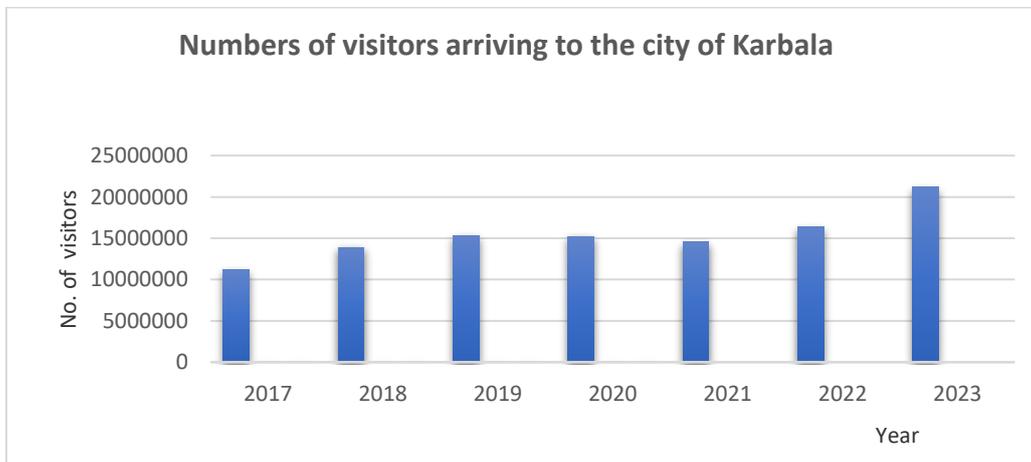


Fig. 3. The number of visitors arriving to the city of Karbala

In 1980, sewage networks were established in the city centre of Karbala, utilising pipes of varying diameters. The network comprises primary and secondary pipelines, as depicted in Fig. 4. These pipelines are linked to concrete manholes, which come in round and rectangular forms, with varying intervals between each manhole.



Fig. 4. Aerial photo showing Karbala city center with sewage networks

Due to the aging of the city's sewage network and its design to accommodate a limited number of residents, sewage overflow occurs in many locations, as in the Fig. 1, affecting the environment and traffic during the period of religious visitation.



Fig. 5. Sewage overflow in a number of locations in the city center of Karbala

V. PROJECT NECESSITY

The necessity of the project is to establish a new sewage network that is compatible with the huge amounts of drainage during the period of religious visits to the city to prevent recurring sewage overflows and preserve the environment from pollution.

VI. SEWERGEMS

SewerGEMS simplifies the modelling process, allowing us to allocate more time to address wastewater engineering challenges such as enhancing capacity and mitigating sewer overflows. This eventually enables utilities to adhere to sewer standards established by regulatory authorities. SewerGEMS offers experts an advanced engineering tool for the planning, design, maintenance, and operation of sanitary sewage systems. This tool encompasses the following aspects [10].

1-Utilise the what-if scenario management tools of SewerGEMS to gain insights into the functioning of your wastewater system, enhancing decision making and reducing reaction time more effectively.

2-Enhanced model accuracy: A properly calibrated model reduces the likelihood of making suboptimal decisions and guarantees the utilisation of the most reliable data provided to it. The inclusion of SCADA connectivity in SewerGEMS calibration tools allows experts to have confidence in the accuracy of their model findings.

3-SewerGEMS facilitates the efficient exchange of information across utilities' design, engineering, GIS, and operation departments, as well as their consultants, through its comprehensive GIS, CAD, and SCADA interoperability. This enhanced information mobility is cost-effective and enables seamless sharing and flow of data. Users may utilise the data that the utility has invested in once and reuse it across several departments.

SewerGEMS offers engineers a user-friendly platform to analyse, design, and manage sanitary or combined conveyance sewer systems. It includes advanced hydraulic and hydrology features, as well as various wet weather calibration methods, making it suitable for tasks ranging from urban sewer planning to overflow remediation analysis to

optimised best management practices design. SewerGEMS utilises Bentley CONNECT services by linking a hydraulic model with a CONNECT project [10]. This facilitates seamless sharing access to the model for all team members.

7-Hydraulic modeling and analysis

The primary goal of a hydraulic analysis of the network is to scrutinise the architectural model that was previously constructed in the Arc map environment. Hydraulic models are the most successful methodology in various types of sewer network study [11]. The objective of the control system is to utilise the available capacity of the pipes in order to minimise overflow in the wastewater system. The primary objectives of hydraulic design are to ascertain the necessary diameter of sewer pipes, compute and regulate the slope, and manage the flow velocity within the sewer pipes. Velocity plays a crucial role in fluid dynamics and is subject to specific limitations based on design criteria. Additionally, the sewer pipes are intentionally built with incomplete filling, necessitating the specification of the fullness-to-emptiness ratio in hydraulic calculations. The model was subjected to a load using the "area load" method in this investigation. The dispersion of discharge in the pipelines has also been achieved utilising "Equal Flow Distribution." The distribution of flow in all pipelines is proportional to the area occupied by each pipeline. Given that certain pipelines may have varying numbers of loads, this situation will not actually occur in actuality. This approach is referred to as the optimal loading framework, as it was specifically designed to assess the feasibility of modelling. The following design scenario was considered for the hydraulic design:

a) The type of pipe material selected is, Polyvinyl chloride (PVC).

- b) Pipeline's diameter range is between 30 to 1800 cm.
- c) The high and low-velocity limits are assumed 500-30cm/sec.
- d) Depth of burial of pipelines according within 100 to 500cm
- e) The slope for piping was set among 0.5 and 10%percent.
- f) All pipelines are considered (75%) full.

The generated model has undergone validation to assess the input data and design settings scenario. Upon completion of the validation procedure, the model was executed to conduct hydraulic analysis. The software quantifies and evaluates numerical and analytical factors.

8-Presenting and Evaluating the Results

The SewerGEMS programme provides multiple techniques for displaying the results of the research. The "design report" offers a succinct summary of the scrutinised network. The relevant information comprises the designer's specifications, project title, organisation name, analysis date, and design. Furthermore, it offers succinct information regarding the particular case that was utilised to construct the model. It precisely represents the exact amount of each element used in the network. This analysis employs a grand total of 863 pipes and manholes, which encompasses three outfall structures. Furthermore, it furnishes the precise measurements of the lengths and diameters for every range of pipeline diameters employed in the model (Fig. 6). A total of 49,070.9 metres of pipelines were used in this modelling project.

Hydraulic Model Inventory: sewer net.16-12.stsw

Title	Sanitation network for the city center of Karbala - Iraq		
Engineer	Insan Kadhim Abed		
Company			
Date	12/7/2023		
Notes			
Scenario Summary			
ID	1		
Label	Base		
Notes			
Active Topology	Base Active Topology		
User Data Extensions	Base User Data Extensions		
Physical	Base Physical		
Boundary Condition	Base Boundary Condition		
Initial Settings	Base Initial Settings		
Hydrology	Base Hydrology		
Output	Base Output		
Infiltration and Inflow	Base Infiltration and Inflow		
Rainfall Runoff	Base Rainfall Runoff		
Water Quality	Base Water Quality		
Sanitary Loading	Base Sanitary Loading		
Headloss	Base Headloss		
Operational	Base Operational		
Design	Base Design		
System Flows	Base System Flows		
SCADA	Base SCADA		
Energy Cost	Base Energy Cost		
Surface Definition	Base Surface Definition		
Solver Calculation Options	Base Calculation Options		
Network Inventory			
Conduits	863	Outfalls	6
-Circle	863	Catchments	0
-Box	0	Low Impact Development Controls	0
-Ellipse	0	Ponds	0
-Virtual	0	Pond Outlet Structures	0
-Irregular Channel	0	Headwalls	0
-Trapezoidal Channel	0	Pumps	0
-Triangular Channel	0	Wet Wells	0
-Rectangular Channel	0	Pressure Junctions	0
-Pipe-Arch	0	SCADA Elements	0
Laterals	0	Pump Stations	0
Channels	0	Variable Speed Pump Batteries	0
Gutters	0	Air Valves	0
Pressure Pipes	0	Grids	0
Catch Basins	0	Surface Polygons	0
Manholes	863	Surface Polylines	0
Property Connections	0	Boundary Lines	0
Taps	0	Boundary Points	0

Circle Inventory			
Circle - 0.3 m	21,094.0 m	Circle - 1.0 m	763.5 m
Circle - 0.4 m	3,772.0 m	Circle - 1.2 m	2,142.5 m
Circle - 0.5 m	3,682.6 m	Circle - 1.3 m	276.3 m
Circle - 0.6 m	4,427.4 m	Circle - 1.5 m	3,101.8 m
Circle - 0.7 m	2,897.4 m	Circle - 1.8 m	5,126.1 m
Circle - 0.8 m	1,254.3 m	Total Length	49,070.9 m
Circle - 0.9 m	533.0 m		

sewer net.16-12.stsw
12/23/2023

Bentley Systems, Inc. Haestad Methods Solution
Center
76 Watertown Road, Suite 2D Thomaston, CT
06787 USA +1-203-755-1666

SewerGEMS
[10.04.00.158]
Page 2 of 2

Fig. 6. Summary of hydraulic design

The software allows for the presentation of distinct tables, enabling the examination of design specifics to obtain further information regarding manholes and pipelines. The programme report can illustrate the materials used in the pipeline design, the length of each branch, and all the necessary information. An alternative approach to visualise the outcomes is through diagrammatic representation. The graph depicting the temporal aspects of different model components can be utilised to elucidate the fluctuations within each component within the framework of a map. The variation in the "hydraulic gradient" within a manhole during a 24-hour period is seen in Fig. 7. By implementing the extended time strategy in the design example, a definitive result is achieved.

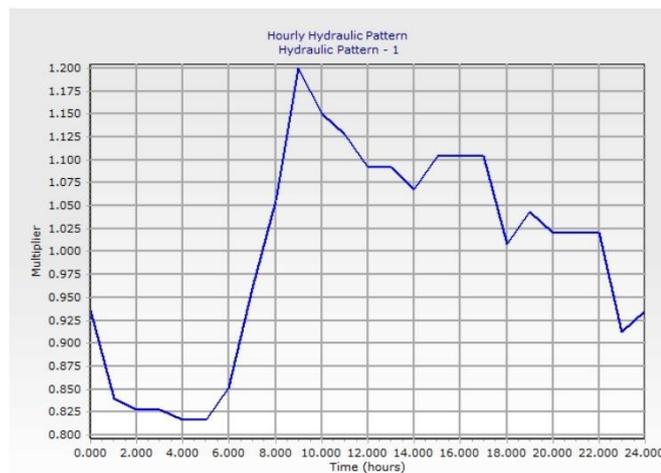


Fig. 7. Hydraulic gradient for manhole 259

Displaying the sewer network profile is an essential component. The comprehensive analysis of all pipelines is crucial for the successful execution of sanitation initiatives. The crucial path, also referred to as the farthest path, is observable. The profile may encompass surface elevation, manhole depth, invert elevation, pipeline gradient, width, and any other necessary data.

VII. CONTROL OF PARAMETERS

Certain modules can serve as control parameters to evaluate the accuracy and exactness of the design and hydraulic analysis. One component that can be analysed is the velocity, which refers to the speed of flow. The velocity limitations for sewage networks range from 30 to 500 cm/s, with an optimal range of 60 to 180 cm/s, as determined by the design requirements. Therefore, it can be concluded that in wastewater collection modelling, the velocity must fall within the appropriate range, otherwise the ideal level. The flow velocity in all tubes ranged from 30 to 487 cm/s. Similar

to the conventional practice of incorporating pipeline diameters from the software library, other variables can be monitored, verified, and quantified based on the flow volume. Upon executing the programme, the selection of all diameters was determined according to the pipe diameter that is commonly accessible, as indicated in the diameter column. SewerGEMS can also be utilised to validate factors associated with manholes. The elevation of the ground and the vertical distance of the manholes, known as invert altitude, are two crucial factors to take into account while dealing with manholes.

VIII. CONCLUSION

The objective of this project was to develop a model for upgrading an existing sewage drainage network, originally designed for restricted capacity, to handle a significantly larger volume of drainage during pilgrimages. Shapefiles were generated separately for pipelines and manholes within the ArcMap framework to construct the architectural model. The existing architectural model is imported into the SewerGEMS software and modified to create a hydrological model. The hydraulic assessment of the model was conducted using SewerGEMS software. The geometric modelling and hydraulic analysis demonstrate that wastewater collection systems are feasible and suitable solutions for the city of Karbala. However, after creating numerous simulations of different situations, the hydraulic model design for the city was ultimately finished using the SewerGEMS programme, taking into account the highest possible discharge. The successful development and operation of the hydraulic model can be deduced from the study of hydraulic findings and the evaluation of control parameters such as flow rate, pipeline diameters, slopes, and profiles. These models can be used to analyse the sizes of cities at the district level. The absence of a coherent urban design poses challenges in developing such models; this analysis seeks to incorporate supplementary tools into the gravity system for both planned and unplanned scenarios geographical areas.

REFERENCES

- [1] K. Muruges, B. Krishna and B. Manoj Kumar, "Desing of Sanitary Sewer Network using Sewer GEMS V8i Software," International journal of Science Technology and Engineering, vol. 2, no. 01, 2015.
- [2] Abhishek P., Yogesh P., Sachin M., Satyajee P., Nikhil T., Yugandhara I. "Design of Sewer System for Village using SewerGEMS," International Research Of Engineering and Technology, vol. 08, no. 07, 2021.
- [3] Vidhi B., Yash B., Raj M., Kushal P. "Design of Sewer Network for Dolarana Vasana Village using SewerGEMS Software," International Journal of Science Technolgy & Engineering, vol. 4, no. 2, 2017.
- [4] Shradda T., Mangesh B. "Design of Underground Drainage System in Rural Area Using SewerGEMS Software," International Research Journal of Engineering ana Technology, vol. 07, no. 06, pp. 126-133, 2020.

- [5] Savadhary k., Sweety W. Shivani T., Rajashri C., Priyanka T. "Design of Sewer System for Holkarwadi Village by using SewerGEMS Software," International Journal for Research in Applied Science & Engineering Technology, vol. 9, no. 6, 2021.
- [6] Naveen Kumar Rai "Sewerage System Assessment Using SewerGEMS V8i and AutoCad Civil 3D," International Journal of Engineering Science Invention, vol. 9, no. 5, pp. 24-29, 2020.
- [7] Nishant S., Timbadiya P. "Hydraulic and Condition Assessment of Existing Sewerage Network :A Case Study of an Educational Institute," Journal of Institution of Engineers (India): Series A, vol. 99, pp. 555-563, 2018.
- [8] "www.karbala.gov.iq," [Online].
- [9] [Online]. Available: <http://c-karbala.com>.
- [10] Shraddhn,T.,Mangesh,B. "Design of Underground Drainage System in Rural Area Using SewerGEMS Software," International Research Journal of Engineering and Technology, vol. 07, no. 06, 2020.
- [11] Zhang D., Martinez N., Lindholm G., Ratnaweera H. "Manage Sewer In-Line Storage Control Using Hydraulic Model and Recurrent Neural Network," Water Resources Management, vol. 32, no. 6, pp. 2079-2098, 2018.

Повышение эффективности модели доступа к удаленной памяти MPI для систем с распределенной памятью путем реализации односторонней рассылки

М. Абуэльсауд¹, А. А. Пазников²

Санкт-Петербургский электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)

¹welleamy@gmail.com, ²apaznikov@gmail.com

Аннотация. В настоящее время обработка больших объемов расширяющихся данных эффективно и последовательно представляет собой значительное вызов. Традиционные высокопроизводительные компьютеры с распределенной памятью (HPC), основанные на модели передачи сообщений, сталкиваются с встроенными трудностями синхронизации, что ограничивает их способность следовать за темпом развития. Удаленный доступ к памяти (RMA), также известный как односторонние обмены MPI позволяет процессу напрямую читать из памяти другого процесса или записывать в нее, обходя необходимость в обмене сообщениями. К сожалению, в текущем стандарте MPI RMA отсутствует интерфейс коллективных операций. Тем не менее, RMA имеет потенциал снижать затраты на синхронизацию, позволяя одновременный доступ к общим структурам данных, распределенным среди памяти процессов MPI. Существующие стандарты односторонних обменов MPI предлагают только линейный интерфейс, который затрудняет параллелизацию и далек от эффективности. Для заполнения этой пробела мы предлагаем алгоритмическое проектирование для эффективных коллективных (параллелизуемых) операций в парадигме RMA. Наше исследование в первую очередь рассматривает выгоды коллективных операций, используя в качестве примера алгоритм рассылки. Наши реализации превосходят традиционные методы, демонстрируя многообещающий потенциал этой техники, как показывают дальнейшие тесты производительности.

Ключевые слова: MPI, высокопроизводительные вычисления, параллельное программирование, удаленный доступ к памяти, MPI RMA, модель общей памяти MPI

I. ВВЕДЕНИЕ

В высокопроизводительных вычислениях (HPC) множество параллельных приложений зависят от коллективных операций, осуществляемых с помощью интерфейса передачи сообщений (MPI) для распределенных коммуникационных задач, таких как рассылка и сокращение обмена сообщениями. Исследования, проведенные в операционных условиях HPC, демонстрируют, что коллективные операции MPI могут потреблять более двух третей времени на коммуникацию в приложениях MPI, часто становясь основными источниками узких мест в производительности [1, 2].

В рамках MPI удаленный доступ к памяти (RMA) предлагает эффективную среду для обмена данными между процессами, обходя прямое вовлечение процессора. Хотя этот метод доказывает свою эффективность во множестве сценариев, он сталкивается с трудностями при обработке операций рассылки, когда один процесс распространяет данные на несколько получателей. Традиционные реализации MPI обычно требуют двусторонней связи для каждого процесса, участвующего в рассылке, что может привести к неэффективности и увеличенной задержке, особенно в обширных развертываниях [3]. Более того, интеграция операций RMA приводит к значительному снижению затрат, поскольку она устраняет расходы, связанные с сопоставлением тегов, управлением предварительным прибытием сообщений и сложностями, присущими буферизации в системах передачи сообщений точка-точка [4, 5].

Выполнение программ часто требует множественных идентичных операций удаленного доступа к памяти (RMA) на различных сегментах памяти [5–8]. Однако текущий стандарт MPI [9] ограничивает пользователей простыми линейными (последовательными) алгоритмами, которые далеки от оптимальных, когда семантика программы позволяет параллельный доступ к данным. Таким образом, основной целью этого исследовательского проекта является разработка алгоритмов для коллективных операций (collectives) в рамках модели RMA. Ожидается, что эти алгоритмы значительно превзойдут линейные алгоритмы, что приведет к сокращению времени выполнения и энергопотребления в программах MPI. Для иллюстрации концепций и принципов мы сосредотачиваемся на алгоритме рассылки (one-to-all), который широко используется в схемах коммуникации. Выбор этого алгоритма основан на его относительной простоте в понимании и реализации. В этом подходе один процесс (корневой) передает данные всем остальным процессам, участвующим в коллективной операции. Определившись с простым алгоритмом, мы стремимся облегчить анализ и отладку, что позволит легче выявлять потенциальные проблемы или узкие места. Изучение сложностей проектирования коллективных операций в рамках модели MPI RMA представляет уникальные вызовы по сравнению с традиционными интерфейсами передачи

сообщений. Среди этих вызовов – необходимость синхронизации данных, где процессы должны координировать и синхронизировать свой доступ к общим данным в удаленной памяти во время коллективных операций на основе RMA. Обеспечение правильной синхронизации и предотвращение конфликтов данных становится сложным, когда несколько процессов одновременно получают доступ и изменяют общие данные. Решение этих проблем требует тщательного анализа и разработки эффективных стратегий для достижения эффективной и конфликтно-свободной синхронизации данных в контексте коллективных операций в рамках модели MPI RMA.

Данное исследование направлено на изучение модели MPI для коллективных операций, ее особенностей и интеграции в распределенные системы памяти. Мы исследуем основные концепции, лежащие в основе этого подхода, рассмотрим его потенциальные преимущества и недостатки, и обсудим стратегии преодоления его вызовов. Кроме того, мы исследуем будущие направления модели MPI RMA для коллективных операций в контексте улучшения парадигм HPC и предоставим примеры реального использования [10].

II. РАССЫЛКА В ОБЩЕЙ ПАМЯТИ RMA

Этот раздел исследует эффективный способ реализации алгоритма рассылки с использованием модели удаленного доступа к памяти (RMA). Рассылка выбрана, поскольку она является как фундаментальной, так и часто используемой в параллельном программировании. В операции рассылки данные изначально находятся в памяти одного процесса. Затем операция реплицирует эти данные для всех процессов, участвующих в том же MPI-коммуникаторе. В рамках модели RMA все участвующие процессы должны разделять одиночный объект окна RMA, названный 'win', с выделенной достаточной памятью для хранения передаваемых данных [4]. Спецификация MPI позволяет использовать возможности сетевого оборудования для односторонней связи. При использовании удаленного доступа к памяти (RMA) в MPI группы процессов, инкапсулированные внутри MPI-коммуникатора, связываются с MPI-окнами. Эти окна представляют собой области памяти, из которых могут быть считаны или в которые могут быть записаны удаленные узлы. Другими словами, операции RMA выполняются в рамках MPI-окна, облегчая передачу данных между локальным процессом MPI (исходным) и удаленным процессом MPI (целью). Интерфейс RMA включает поддержку операций чтения (MPI_Get()), записи (MPI_Put()) и атомарных операций с памятью (MPI_Accumulate()) для данных, предоставленных внутри MPI-окна, показанных на рисунке 1 [11].

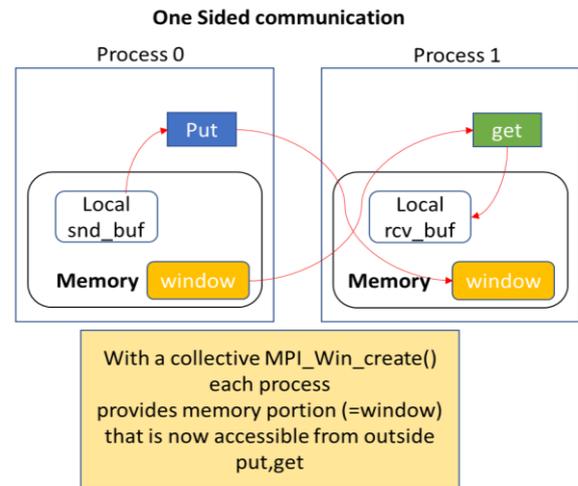


Рис. 1. Идет операция RMA Put. Процесс 0 записывает данные в память процесса 1

В рамках фреймворка Message Passing Interface (MPI), коммуникация через общую память облегчает обмен данными между процессами, находящимися на одном вычислительном узле, путем прямого доступа к области общей памяти. Этот подход заметно повышает эффективность коммуникации, особенно когда процессы находятся на одной физической машине, как показано на рис. 2, благодаря врожденной низкой задержке и высокой пропускной способности, связанных с архитектурами общей памяти [12].

- Выделение памяти: MPI предоставляет процедуры для выделения областей общей памяти, доступных для нескольких процессов. Например, можно использовать MPI_Win_allocate shared для выделения окна общей памяти.
- Доступ к памяти: После выделения области общей памяти, процессы могут напрямую читать из этого пространства памяти и записывать в него. Этот прямой доступ исключает необходимость в явной передаче сообщений через сеть, что снижает накладные расходы на коммуникацию.
- Освобождение памяти: Когда область общей памяти больше не нужна, MPI предоставляет процедуры для освобождения памяти. Например, можно использовать MPI_Win_free для освобождения окна общей памяти.

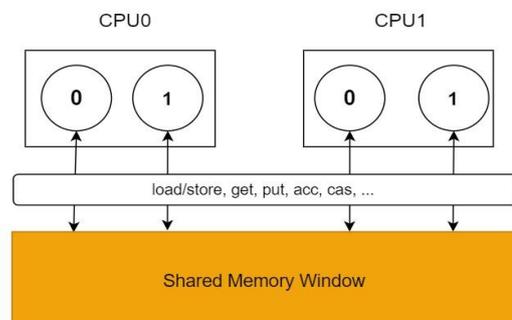


Рис. 2. Расширение общей памяти между процессами с использованием MPI RMA и окно общей памяти выделяется в каждом узле

А. Бинарное дерево

Алгоритм бинарного дерева обеспечивает эффективный метод трансляции данных в средах Message Passing Interface (MPI). В этом разделе описывается его реализация и анализируется его временная сложность. В начальном раунде, изображенном на рис. 4, процесс 0 сохраняет буфер *b* в памяти процессов 1 и 2. Затем, во втором раунде, процесс 1 передает буфер в память процессов 3 и 4. Во время третьего раунда процесс 2 помещает сообщение в буфер процессов 5 и 6. Наконец, в четвертом раунде процесс 3 перемещает сообщение в память процесса 7 [13].

С каждым раундом количество активно транслирующих узлов удваивается. Этот экспоненциальный рост приводит к значительному сокращению числа этапов передачи 2. В результате, последний процесс, получающий данные (например, ранг 7 на рис. 3), требует только $\log_2(p)$ раундов коммуникации для завершения трансляции. В алгоритме RMA Vcast Binary 1:

Функция сначала переупорядочивает ранги процессов, представляя структуру бинарного дерева, где корневой процесс имеет ранг 0.

- Он вычисляет новый ранг (*rank*), вычитая ранг корневого процесса *root* из текущего ранга процесса (*my_rank*), а затем берет модуль от общего числа процессов (*prgos*). Это гарантирует, что новый ранг находится в диапазоне рангов процессов на строке (1).
- Он вычисляет ранги дочерних процессов (*child1* и *child2*) в структуре бинарного дерева на основе текущего ранга процесса на строках (4,5).
- Для каждого дочернего процесса (*child1* и *child2*) функция проверяет, меньше ли ранг дочернего процесса общего числа процессов (*p*) на строках (6,12).
- Если существует дочерний процесс, он отображает виртуальный ранг дочернего процесса на реальный ранг, добавляя ранг корневого процесса и беря модуль от общего числа процессов на строках (7,13).
- Блокировка окна MPI, связанного с дочерним процессом, на строках (8,14).
- Затем вызывается операция `MPI_Put` для отправки данных из *src_buf* (9,15).
- Разблокировка окна MPI, связанного с дочерним процессом, на строке (10,16).

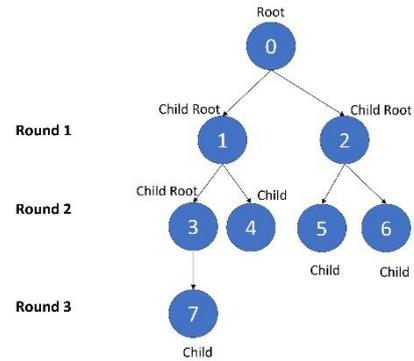


Рис. 3. Иллюстрация передачи данных бинарного дерева между 8 процессами

Input data:	<i>src_buf</i> – буфер происхождения <i>win</i> – окно RMA <i>comm</i> – коммуникатор
1	<i>srank</i> = <code>comp_srank(myrank, root, p)</code>
2	for <i>r</i> = 0 to <i>p</i> do
3	<i>rank</i> = <code>comp_rank(rank, root, p)</code>
4	<i>child1</i> = <code>2 * rank + 1</code>
5	<i>child2</i> = <code>2 * rank + 2</code>
6	if <i>child1</i> < <i>p</i> then
7	<i>child1</i> = <code>comp_rank(rank, root, p)</code>
8	MPI_Win_lock(MPI_LOCK_SHARED, win);
9	MPI_Put(<i>src_buf</i> , <i>length</i> , <i>datatype</i> , <i>child1</i> , win);
10	MPI_Win_unlock(<i>child1</i> , win);
11	end if
12	if <i>child2</i> < <i>p</i> then
13	<i>child2</i> = <code>comp_rank(rank, root, p)</code>
14	MPI_Win_lock(MPI_LOCK_SHARED, win);
15	MPI_Put(<i>src_buf</i> , <i>length</i> , <i>datatype</i> , <i>child2</i> , win);
16	MPI_Win_unlock(<i>child2</i> , win);
17	end if
18	end for

Алгоритм 1 RMA Vcast binary

Ограничение такое же, как у алгоритма биномиального дерева, упомянутого в подразделе II-B

В. Биномиальное дерево

В этом разделе подробно описывается реализация алгоритма биномиального дерева для трансляции данных в среде Message Passing Interface (MPI), используя возможности Remote Memory Access (RMA). Алгоритм использует итерационные раунды, число которых определяется логарифмом по основанию 2 от общего числа процессов (*p*). Каждый раунд включает в себя специфические операции обмена данными на основе рангов процессов [14], [15], [16], [17]. Рассмотрим сценарий с восемью процессами (*p* = 8). Алгоритм разворачивается на протяжении трех раундов:

- **Раунд 1:** Процесс 0 (корень) использует операции MPI RMA для размещения буфера данных в область общей памяти, доступную для процесса 1.
- **Раунд 2:** Процесс 0 снова использует MPI RMA для размещения буфера данных в области общей памяти как процесса 1, так и процесса 2.
- **Раунд 3:**
 - Процесс 0 использует MPI RMA для размещения буфера данных в области общей памяти процессов 1, 2 и 4.
 - Процесс 1, получивший данные в предыдущих раундах, использует MPI RMA для

размещения буфера данных в области общей памяти процессов 3 и 5.

– Аналогично, процесс 2 использует MPI RMA для размещения буфера данных в области общей памяти процесса 6.

– Наконец, процесс 6, получивший данные в предыдущем раунде, использует MPI RMA для размещения буфера данных в области общей памяти процесса 7

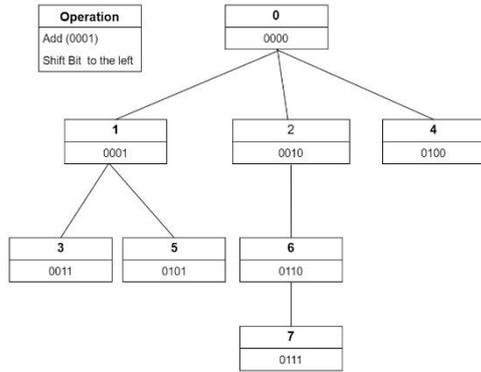


Рис. 4. Визуализация биномиального дерева с 8 процессами, использующими битовый сдвиг для эффективной коммуникации.

Временная сложность алгоритма биномиального дерева с использованием MPI RMA является алгоритмом putloop 2 Анализ сложности:

- На строке 3 цикл выполняется, пока $mask < N$, и $mask$ удваивается на каждой итерации, тогда количество итераций $\log_2(N)$.
- На строке 9 внутри цикла выполняется операция MPI_Put для переноса данных из src_buf , которая выполняет операцию копирования размером с длину сообщения
- Пусть M представляет длину сообщения, тогда общая сложность составляет: $O(M \log(N))$.

Input data:	src_buf – буфер происхождения win – окно RMA $comm$ – коммуникатор
1	$srank = comp_srank(myrank, root, p)$
2	$mask = 1$
3	while $mask < p$ do
4	if $(srank \& mask) == 0$ then
5	$rank = srank mask$
6	if $rank < p$ then
7	$rank = comp_rank(rank, root, p)$
8	MPI_Win_lock(MPI_LOCK_SHARED, win);
9	MPI_Put(src_buf , length, datatype, rank, win);
10	MPI_Win_unlock(child2, win);
11	else
12	break
13	end if
14	$mask = mask \ll 1$
15	end if
16	end while

Алгоритм 2 put loop

Анализ сложности алгоритма RMA_Bcast_binomial:

- В алгоритме 3 на строке 1 цикл выполняется p раз, при каждой итерации вызывается функция putloop.
- Каждый вызов sendloop имеет сложность $O(M \log(N))$.
- Полная сложность составляет $O(N \times M \times \log(N))$.

$$T_{loop} = O(N \times M \times \log(N)) \quad (1)$$

$$T_{total} = N \times T_{loop} \quad (2)$$

Из (1) и (2)

$$T_{total} = N \times O(N \times M \times \log(N)) \quad (3)$$

Input data:	src_buf – буфер происхождения win – окно RMA $comm$ – коммуникатор
1	for $r = 0$ to p do
2	$root = r$
3	put_loop(src_buf , $dest_buf$, rank, root, win, comm)
4	end for

Алгоритм 3 RMA Bcast binomial

Ограничения алгоритма биномиального дерева:

- Производительность алгоритма может ухудшиться при большом количестве процессов из-за увеличения накладных расходов на коммуникацию и синхронизацию.
- Этот подход может быть эффективным для небольшого и среднего количества процессов, но может столкнуться с увеличением задержек и накладными расходами при большом количестве процессов.
- Алгоритм зависит от наличия достаточного объема памяти и пропускной способности сети для обмена данными.

III. ЭКСПЕРИМЕНТАЛЬНАЯ ОЦЕНКА

В рамках данного исследования эксперименты проводились на вычислительном кластере, размещенном на платформе Yandex Cloud. Кластер использовал процессоры Intel Ice Lake, конкретно Intel® Xeon® Gold 6338 с тактовой частотой 2.00 ГГц. Каждый экземпляр виртуальной машины в составе кластера был сконфигурирован с 4 виртуальными центральными процессорами (vCPU), 8 ГБ оперативной памяти (RAM) и 20 ГБ дискового пространства.

Для создания бенчмарка были использованы вычислительные нагрузки, соответствующие нашим исследовательским контекстам. Мы провели тесты на сервере, включающие обработку 1000 сообщений, распределенных по восьми отдельным пакетам данных. Размеры этих пакетов были увеличены последовательно в двоичном формате, начиная с минимального значения 16 байт и удваиваясь на каждом шаге, пока не был достигнут пиковый размер в 33 мегабайта.

Согласно графическому представлению, показанному на рис. 5, алгоритм бинарного дерева демонстрирует превосходную производительность, за которым следует алгоритм биномиального дерева, а затем линейный алгоритм. Например, использование алгоритма бинарного дерева позволяет передавать данные двум процессам примерно за 0,01 миллисекунды, в то время как использование линейного алгоритма требует примерно 0,1 миллисекунды для выполнения той же задачи.

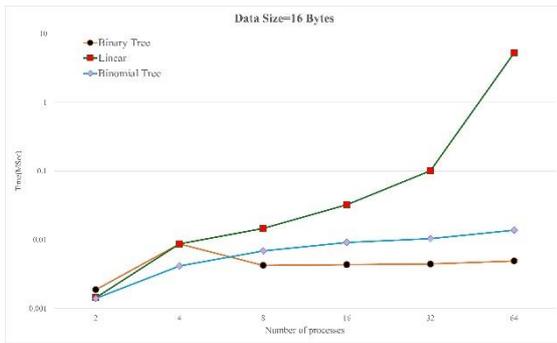


Рис. 5. Время рассылки в зависимости от выбора алгоритма (16 байт).

Как показано на рис. 6, алгоритм бинарного дерева демонстрирует превосходную производительность, за которым следует алгоритм биномиального дерева, а затем линейный алгоритм. Например, передача данных 32 процессам при использовании алгоритма бинарного дерева требует примерно 1 миллисекунды, в то время как использование линейного алгоритма для той же задачи требует примерно 10 миллисекунд.

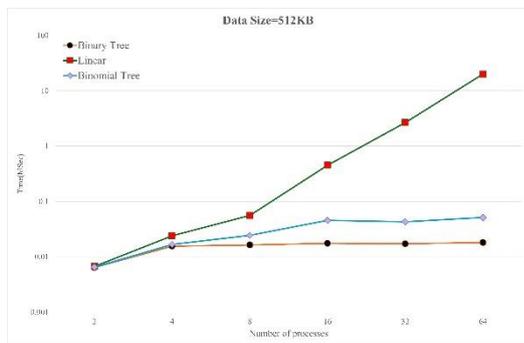


Рис. 6. Время рассылки в зависимости от выбора алгоритма (512 КБ данных)

Исходя из рис. 7, алгоритм бинарного дерева показывает наивысшую производительность, за которым следует алгоритм биномиального дерева, а затем линейный алгоритм. Например, передача данных 16 процессам с использованием алгоритма бинарного дерева занимает примерно 10 миллисекунд, в то время как для выполнения той же задачи с использованием линейного алгоритма требуется примерно 100 миллисекунд.

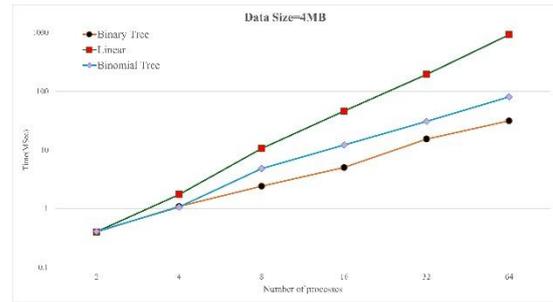


Рис. 7. Время рассылки в зависимости от выбора алгоритма (4 МБ данных)

Фигура 8 показывает время (измеренное в миллисекундах), необходимое для рассылки 32 МБ данных всем процессам с использованием трех различных алгоритмов: бинарного дерева, линейного и биномиального дерева. На оси x представлено количество процессоров, а на оси y – время (в миллисекундах) для рассылки данных.

Исходя из данных на рис. 8, алгоритм бинарного дерева демонстрирует наиболее эффективную производительность, за которым следует алгоритм биномиального дерева, и, наконец, линейный алгоритм. Например, рассылка данных 64 процессам с использованием алгоритма бинарного дерева занимает примерно 100 миллисекунд, в то время как для выполнения той же задачи с использованием линейного алгоритма требуется примерно 1000 миллисекунд.

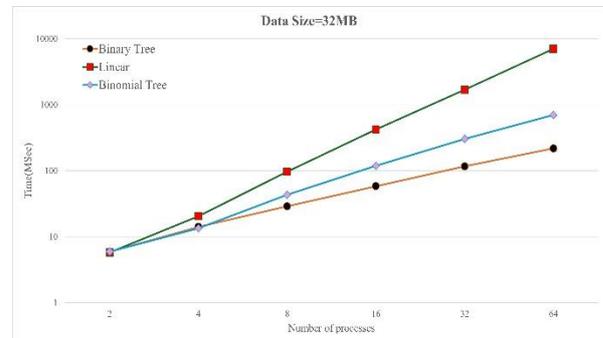


Рис. 8. Производительность рассылки с большим объемом данных (32 МБ): Бинарное дерево против Линейного против Биномиального дерева.

IV. ЗАКЛЮЧЕНИЕ

Исследование прояснило эффективность нескольких алгоритмов рассылки при улучшении модели удаленного доступа к памяти MPI для распределенных систем памяти, с особым упором на реализацию односторонней рассылки. После тщательного анализа были сделаны несколько важных выводов:

Бинарное дерево: Оно выделяется как наиболее эффективный алгоритм, демонстрируя логарифмическое увеличение времени рассылки при увеличении числа процессов, что указывает на масштабируемость и эффективность. Кроме того, биномиальное дерево: Расположенное между бинарным деревом и линейными алгоритмами, оно предлагает сбалансированный

профиль производительности, обеспечивая компромисс между эффективностью и сложностью. Линейный: Несмотря на свою простоту, он демонстрирует наименее эффективное время рассылки среди тройки, показывая линейное увеличение времени рассылки с увеличением числа процессов, что подчеркивает ограничения масштабируемости и эффективности.

Более того, необходимо признать потенциальные вызовы и ограничения, связанные с предложенным подходом:

1. Деградация производительности: Эффективность алгоритма может уменьшиться с увеличением количества процессов из-за роста накладных расходов на коммуникацию и синхронизацию. По мере масштабирования системы ее врожденные ограничения могут препятствовать эффективности.

2. Зависимость от количества процессов: Хотя алгоритм эффективен для небольшого и среднего числа процессов, увеличение числа процессов может привести к увеличению накладных расходов и задержек. Это подчеркивает необходимость тщательного рассмотрения масштабируемости в реальных распределенных системах.

3. Зависимость от ресурсов: Эффективность алгоритма зависит от доступной памяти и пропускной способности сети для безпроблемной коммуникации. Недостаточные ресурсы могут затруднить оптимальную производительность, что требует стратегий выделения и оптимизации ресурсов.

В заключение, хотя предложенный алгоритм проявляет обнадеживающие черты производительности, его эффективность зависит от того, насколько тщательно учитываются масштабируемость, доступность ресурсов и динамика системы. Будет крайне важно учитывать эти факторы, чтобы полностью использовать метод и применять его к распределенным системам памяти.

СПИСОК ЛИТЕРАТУРЫ

- [1] Chunduri, S., Parker, S., Balaji, P., Harms, K., & Kumaran, K. (2018). Characterization of MPI usage on a production supercomputer. In Proceedings - International Conference for High Performance Computing, Networking, Storage, and Analysis, SC 2018. doi: 10.1109/SC.2018.00033.
- [2] Hoefler, T., Siebert, C., & Rehm, W. (2007). A practically constant-time MPI broadcast algorithm for large-scale InfiniBand clusters with multicast. In Proceedings - 21st International Parallel and Distributed Processing Symposium, IPDPS 2007; Abstracts and CD-ROM. doi: 10.1109/IPDPS.2007.370475.
- [3] Hoefler, T., et al. (2015). Remote memory access programming in MPI-3. ACM Transactions on Parallel Computing, 2(2). doi: 10.1145/2780584.
- [4] Abuelsoud, M. M., Kogutenko, A. A., & Naveen. (2024). Enhancing MPI remote memory access model for distributed-memory systems through one-sided broadcast implementation. J Phys Conf Ser, 2697(1), 012035. doi: 10.1088/1742-6596/2697/1/012035.
- [5] Tipparaju, V., Nieplocha, J., & Panda, D. (2003). Fast collective operations using shared and remote memory access protocols on clusters. Proceedings - International Parallel and Distributed Processing Symposium, IPDPS 2003. doi: 10.1109/IPDPS.2003.1213188.
- [6] Petrović, D., Shahmirzadi, O., Ropars, T., & Schiper, A. (2012). High-performance RMA-based broadcast on the Intel SCC. In Annual ACM Symposium on Parallelism in Algorithms and Architectures. doi: 10.1145/2312005.2312029.
- [7] Tipparaju, V., Krishnan, M., Nieplocha, J., Santhanaraman, G., & Panda, D. (2003). Exploiting non-blocking remote memory access communication in scientific benchmarks. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2913. doi: 10.1007/978-3-540-24596-4_27.
- [8] Sur, S., Bondhugula, U. K. R., Mamidala, A., Jin, H. W., & Panda, D. K. (2005). High performance RDMA based all-to-all broadcast for InfiniBand clusters. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). doi: 10.1007/11602569_19.
- [9] Lusk, E., Huss, S., Saphir, B., & Snir, M. (2009). MPI: A message-passing interface standard Version 3.0. International Journal of Supercomputer Applications, 8(3/4).
- [10] Gropp, W., & Lusk, E. (1999). Reproducible measurements of MPI performance characteristics. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). doi: 10.1007/3-540-48158-3_2.
- [11] Hjelm, N., Dosanjh, M. G. F., Grant, R. E., Groves, T., Bridges, P., & Arnold, D. (2018). Improving MPI multi-threaded RMA communication performance. In ACM International Conference Proceeding Series. doi: 10.1145/3225058.3225114.
- [12] Gropp, W., Lusk, E., & Skjellum, A. (2019). Using MPI - 2nd Edition Portable Parallel Programming with the Message Passing Interface (Scientific and Engineering Computation).
- [13] Nuriyev, E., Rico-Gallego, J. A., & Lastovetsky, A. (2022). Model-based selection of optimal MPI broadcast algorithms for multi-core clusters. J Parallel Distrib Comput, 165. doi: 10.1016/j.jpdc.2022.03.012.
- [14] Wadsworth, D. M., & Chen, Z. (2008). Performance of MPI broadcast algorithms. In IPDPS Miami 2008 - Proceedings of the 22nd IEEE International Parallel and Distributed Processing Symposium, Program and CD-ROM. doi: 10.1109/IPDPS.2008.4536478.
- [15] Tu, B., Fan, J., Zhan, J., & Zhao, X. (2012). Performance analysis and optimization of MPI collective operations on multi-core clusters. Journal of Supercomputing, 60(1). doi: 10.1007/s11227-009-0296-3.
- [16] Bruck, J., Ho, C. T., Kipnis, S., Upfal, E., & Weathersby, D. (1997). Efficient algorithms for all-to-all communications in multiport message-passing systems. IEEE Transactions on Parallel and Distributed Systems, 8(11). doi: 10.1109/71.642949.
- [17] Hoefler T., & Lumsdaine A. (2008). Optimizing non-blocking collective operations for InfiniBand. In IPDPS Miami 2008 - Proceedings of the 22nd IEEE International Parallel and Distributed Processing Symposium, Program and CD-ROM. doi: 10.1109/IPDPS.2008.4536138.

Исчисление дискретных рисков в условиях вероятностной неопределенности

Т. А. Уразаева

Поволжский государственный технологический университет,
Республика Марий Эл, г. Йошкар-Ола

urazaevata@volgatech.net

Аннотация. Предложена модель исчисления абстрактных рисков эволюции систем произвольной природы с дискретным множеством состояний в условиях вероятностной неопределенности. Выявлены алгебраические структуры, возникающие на универсальном множестве абстрактных рисков и их частей. Сформулирован аналог полиномиальной теоремы на этом множестве, имеющей место при условии коммутативности эволюционных преобразований состояния системы. Предложен ряд естественных конгруэнций на универсальном множестве абстрактных рисков и их частей, позволяющих также повысить эффективность вычислений на практике.

Ключевые слова: алгебраическая теория риска; вероятностная неопределенность; мультимножество; полиномиальная теорема; риск; система; эволюция

I. ВВЕДЕНИЕ

Любая система (техническая, социально-экономическая и т. п.) подвержена эволюции. Это означает, что система, находясь в начальный момент времени t_0 в состоянии $s(t_0)$, под действием решения субъекта принятия решения (СПР), если он существует, δ , и с учетом состояния внешней среды ω_0 , в некоторый следующий момент времени t_1 окажется в состоянии $s(t_1) = H(t_1, s(t_0), [\delta, \omega_0])$, где H – «историческое» отображение. При этом спектр возможных состояний при фиксированном решении СПР δ определяется множеством возможных состояний природы $\omega_0 \in \Omega_0$. В качестве модели природы мы будем использовать вероятностное пространство $(\Omega_0, \mathcal{A}_0, P_0)$, где Ω_0 – множество состояний природы, \mathcal{A}_0 – σ -алгебра подмножеств множества Ω_0 , P_0 – вероятностная мера на измеримом пространстве $(\Omega_0, \mathcal{A}_0)$. Также будем предполагать, что каррированная функция $H(t_1, s(t_0), [\delta, \cdot])$ измерима на $(\Omega_0, \mathcal{A}_0)$. В этом случае можно говорить о риске развития системы в условиях вероятностной неопределенности, если на множестве состояний $\{H(t_1, s(t_0), [\delta, \cdot])(\omega_0) : \omega_0 \in \Omega_0\}$ задан порядок предпочтения, минимально образующий полурешетку [9].

Обычно переход от состояния $s(t_0)$ к состоянию $s(t_1)$ происходит как череда преобразований состояния системы: $H(t_1, s(t_0), [\delta, \cdot]) \omega_0 = c_k \circ c_{k-1} \circ \dots \circ c_1(s(t_0))$, где каждое преобразование состояния c_i , $i = 1, 2, \dots, k$, может быть связано как с многошаговым воздействием решения СПР, с многошаговым влиянием внешней среды, так и с преобразованиями, связанными с процессами в различных подсистемах рассматриваемой системы. При этом каждое преобразование состояния системы может моделироваться своим вероятностным пространством. В случае, если эти вероятностные пространства конечны, задача исчисления риска развития системы оказывается в ряде случаев доступной для реализации на современных компьютерах.

II. ОПИСАНИЕ МОДЕЛИ ИСЧИСЛЕНИЯ

Пусть S – множество всех возможных состояний системы. Будем рассматривать подмножество $C \subset S^S$ всех отображений множества состояний на себя, таких, которые коммутативны относительно композиции. В этом случае можно показать, что система $\langle C, \circ \rangle$ образует коммутативный моноид [4].

Введем понятие проекции прямого произведения двух множеств $A \times B$ на первый и второй сомножители:

$$\text{pr}_A : A \times B \rightarrow A, \text{pr}_B : A \times B \rightarrow B$$

через сопоставления следующего вида:

$$\text{pr}_A : (x, y) \mapsto x, \text{pr}_B : (x, y) \mapsto y.$$

Пусть $\mathbb{I} = [0, 1] \subset \mathbb{R}$, где \mathbb{R} – множество действительных чисел. Пользуясь понятием проекций, определим на множестве $C \times \mathbb{I}$ операцию умножения следующим образом:

$$z = x \cdot y \stackrel{\text{Def}}{\Leftrightarrow} \text{pr}_C z = \text{pr}_C y \circ \text{pr}_C x, \text{pr}_\mathbb{I} z = \text{pr}_\mathbb{I} x \cdot \text{pr}_\mathbb{I} y.$$

Заметим, что система $\langle C \times \mathbb{I}, \cdot \rangle$ тоже образует коммутативный моноид.

Введем, далее, множество мультимножеств вида:

$$\mathfrak{A} = \{A : \text{Supp } A \subset C \times \mathbb{I}\}$$

и операции сложения и умножения на нем:

$$A + B \stackrel{Def}{=} \left\{ k_{A+B}(u) * u : k_{A+B}(u) = k_A(u) + k_B(u), \right. \\ \left. u \in \text{Supp } A \cup \text{Supp } B \right\},$$

$$A \cdot B \stackrel{Def}{=} \left\{ k_{A \cdot B}(u) * u : k_{A \cdot B}(u) = \sum_{\substack{x \in \text{Supp } A, \\ y \in \text{Supp } B, \\ x \cdot y = u}} k_A(x) k_B(y), \right. \\ \left. u \in \bigcup_{\substack{x \in \text{Supp } A, \\ y \in \text{Supp } B}} \{x \cdot y\} \right\}.$$

С основными идеями и обозначениями теории мультимножеств можно познакомиться в фундаментальной работе А. Б. Петровского [5].

Можно показать, что алгебра $\langle \mathfrak{A}, +, \cdot \rangle$ является коммутативным полукольцом с единицей [11].

Из множества \mathfrak{A} естественным образом можно выделить два подмножества $\mathfrak{X} \subset \mathfrak{Y} \subset \mathfrak{A}$:

1) множество рисков:

$$\mathfrak{X} \stackrel{Def}{=} \left\{ \mathbf{R} : \mathbf{R} \in \mathfrak{A}, \sum_{x \in \text{Supp } \mathbf{R}} k_{\mathbf{R}}(x) \text{pr}_{\mathbb{I}} x = 1 \right\},$$

2) множество рисков и их частей:

$$\mathfrak{Y} \stackrel{Def}{=} \left\{ \mathbf{R} : \mathbf{R} \in \mathfrak{A}, \sum_{x \in \text{Supp } \mathbf{R}} k_{\mathbf{R}}(x) \text{pr}_{\mathbb{I}} x \leq 1 \right\}.$$

Множество рисков содержит, таким образом, полные группы событий (мультимножества, элементы которых являются парами «преобразование системы – вероятность этого преобразования», при этом сумма

вероятностей всех элементов дает 1). Операция умножения на этом множестве фактически означает вычисление совместного риска из двух рисков, представленных сомножителями, при условии их независимости.

Множество рисков и их частей, соответственно, содержит не только полные группы событий, но и всевозможные их разбиения.

Нетрудно проверить, что алгебры $\langle \mathfrak{X}, \cdot \rangle$ и $\langle \mathfrak{Y}, \cdot \rangle$ представляют из себя также коммутативные моноиды. Кроме того на множестве \mathfrak{Y} действует частичная операция сложения. Эту операцию следует понимать, как операцию агрегации риска на любых подмножествах разбиений любого абстрактного риска.

Подводя итог данного раздела, можно определить алгебраическую систему, представляющую универсальное множество абстрактных рисков и их частей в условиях вероятностной неопределенности, как моноид (относительно операции вычисления совместного риска), вложенный в коммутативное полукольцо с единицей, на котором (моноиде) определена частичная операция (полукольцевая) сложения (агрегации риска). Данный вывод демонстрирует фундаментальную алгебраическую природу риска, по крайней мере в условиях вероятностной неопределенности.

III. АНАЛОГ ПОЛИНОМИАЛЬНОЙ ТЕОРЕМЫ

Для элементов коммутативного моноида как частного случая полугруппы определена натуральная степень [4]. Интерпретацией, например, n -й степени некоторого абстрактного риска развития системы является результирующий риск n независимых проявлений исходного риска в системе (n подсистем, n последовательных периодов эволюции системы в условиях, когда известен риск одного периода, и т. п.).

Можно доказать следующую весьма полезную для практики теорему.

Теорема. Пусть $m = |\text{Supp } A|$. Занумеруем элементы носителя мультимножества $A : \biguplus_{i=1}^m x_i = \text{Supp } A$. Пусть, далее, определены предикаты:

$$\pi_1[n](l_1, l_2, \dots, l_m) = (l_i \in \{0, 1, \dots, n\}, i = 1, 2, \dots, m, \sum_{i=1}^m l_i = n),$$

$$\pi_2[A, n, w](l_1, l_2, \dots, l_m) = \pi_1[n](l_1, l_2, \dots, l_m), \biguplus_{i=1}^m x_i^{l_i} = \text{Supp } A, \prod_{i=1}^m x_i^{l_i} = w$$

и

Тогда

$$A^n = \left\{ k_{A^n}(w) * w : k_{A^n}(w) = \sum_{\pi_2[A, n, w](l_1, l_2, \dots, l_m)} \frac{n!}{\prod_{i=1}^m l_i!} \prod_{i=1}^m [k_A(x_i)]^{l_i}, \right. \\ \left. w \in \bigcup_{\pi_1[n](l_1, l_2, \dots, l_m)} \prod_{i=1}^m x_i^{l_i}, \biguplus_{i=1}^m x_i = \text{Supp } A, m = |\text{Supp } A| \right\}$$

для любого $A \in \mathfrak{A}$.

Приведенная теорема является прямым аналогом классической полиномиальной теоремы, переформулированной для мультимножеств и введенной выше операции их умножения [3].

Практическая ценность этой теоремы заключается в том, что, если вычислять степень посредством последовательного умножения, то такой алгоритм вычисления степени может быть отнесен к классу сложности $\mathbf{EXPTIME} \cap \mathbf{EXSPACE}$, в то время как алгоритм, основанный на применении приведенной теоремы, относится к классу сложности $\mathbf{PSPACE} \setminus \mathbf{NL}$.

Напомним [10], что класс сложности $\mathbf{EXPTIME}$ – это множество задач, решаемых детерминированной машиной Тьюринга за время $O(2^{p' n})$, а $\mathbf{EXSPACE}$ – это множество задач, для решения которых детерминированной машине Тьюринга понадобится память порядка $O(2^{p'' n})$. Здесь p' и p'' – полиномиальные функции. Класс сложности \mathbf{PSPACE} – это множество задач, которые могут быть решены машиной Тьюринга с полиномиальным ограничением пространства, а \mathbf{NL} – это множество простых задач, решаемых недетерминированной машиной Тьюринга с использованием $O(\log n)$ памяти.

IV. АРИФМЕТИКА ВЫЧЕТОВ

На универсальном множестве рисков и их частей можно определить ряд конгруэнций, имеющих конкретный прикладной смысл. Соответственно можно определить ряд вычетов, использование которых способно повысить эффективность вычислений.

Введем в рассмотрение две функции:

$$\tilde{k} \mathbf{R} = \left\{ 1 * \left(\tilde{c}, \sum_{\substack{x \in \text{Supp} \mathbf{R}, \\ \text{pr}_C x = \tilde{c}}} k_{\mathbf{R}} x \text{ pr}_1 x \right) : \tilde{c} \in \bigcup_{x \in \text{Supp} \mathbf{R}} \text{pr}_C x \right\},$$

$$\tilde{\sigma} \mathbf{R} = \underset{\text{Def}}{k_{\mathbf{R}} x * x : \text{pr}_1 x > 0, x \in \text{Supp} \mathbf{R}},$$

где $\mathbf{R} \in \mathfrak{F}$.

Можно показать, что на универсальном множестве рисков и их частей возможно введение трех вычетов:

- 1) наименьший (простейший) вычет: $\left| \mathbf{R} \right|_{\kappa} \underset{\text{Def}}{=} \tilde{k} \mathbf{R}$,
- 2) значимый вычет: $\left| \mathbf{R} \right|_{\sigma} \underset{\text{Def}}{=} \tilde{\sigma} \mathbf{R}$,
- 3) наименьший значимый вычет: $\left| \mathbf{R} \right|_{\kappa \sigma} \underset{\text{Def}}{=} \tilde{\sigma} \circ \tilde{k} \mathbf{R}$,

где $\mathbf{R} \in \mathfrak{F}$.

Наименьший (простейший) вычет представляет собой отображение, которое сопоставляет произвольному абстрактному риску или его части (мультимножеству) мультимножество, все компоненты которого характеризуются единичной кратностью, и которые

получены сбором компонент исходного мультимножества с одинаковыми преобразованиями системы и вероятностью, равной сумме соответствующих кратностей и вероятностей, иначе, суммарной вероятности всех исходов с данным преобразованием состояния системы.

Значимый вычет представляет собой отображение, которое сопоставляет произвольному абстрактному риску или его части (мультимножеству) мультимножество, все компоненты которого характеризуются ненулевой вероятностью.

Наименьший значимый вычет реализует отображение, являющееся композицией двух ранее рассмотренных отображений. (Можно доказать коммутативность композиции в данном случае.)

Для перечисленных вычетов можно показать выполнение следующих свойств:

$$\left| \mathbf{A} \right|_{\kappa} + \left| \mathbf{B} \right|_{\kappa} = \left| \mathbf{A} + \left| \mathbf{B} \right|_{\kappa} \right|_{\kappa} = \left| \left| \mathbf{A} \right|_{\kappa} + \left| \mathbf{B} \right|_{\kappa} \right|_{\kappa} = \left| \mathbf{A} + \mathbf{B} \right|_{\kappa},$$

$$\left| \mathbf{A} \right|_{\kappa} \cdot \left| \mathbf{B} \right|_{\kappa} = \left| \mathbf{A} \cdot \left| \mathbf{B} \right|_{\kappa} \right|_{\kappa} = \left| \left| \mathbf{A} \right|_{\kappa} \cdot \left| \mathbf{B} \right|_{\kappa} \right|_{\kappa} = \left| \mathbf{A} \cdot \mathbf{B} \right|_{\kappa},$$

$$\left| \mathbf{A} \right|_{\sigma} + \left| \mathbf{B} \right|_{\sigma} = \left| \mathbf{A} + \left| \mathbf{B} \right|_{\sigma} \right|_{\sigma} = \left| \left| \mathbf{A} \right|_{\sigma} + \left| \mathbf{B} \right|_{\sigma} \right|_{\sigma} = \left| \mathbf{A} + \mathbf{B} \right|_{\sigma},$$

$$\left| \mathbf{A} \right|_{\sigma} \cdot \left| \mathbf{B} \right|_{\sigma} = \left| \mathbf{A} \cdot \left| \mathbf{B} \right|_{\sigma} \right|_{\sigma} = \left| \left| \mathbf{A} \right|_{\sigma} \cdot \left| \mathbf{B} \right|_{\sigma} \right|_{\sigma} = \left| \mathbf{A} \cdot \mathbf{B} \right|_{\sigma},$$

$$\left| \mathbf{A} \right|_{\kappa \sigma} + \left| \mathbf{B} \right|_{\kappa \sigma} = \left| \mathbf{A} + \left| \mathbf{B} \right|_{\kappa \sigma} \right|_{\kappa \sigma} =$$

$$= \left| \left| \mathbf{A} \right|_{\kappa \sigma} + \left| \mathbf{B} \right|_{\kappa \sigma} \right|_{\kappa \sigma} = \left| \mathbf{A} + \mathbf{B} \right|_{\kappa \sigma},$$

$$\left| \mathbf{A} \right|_{\kappa \sigma} \cdot \left| \mathbf{B} \right|_{\kappa \sigma} = \left| \mathbf{A} \cdot \left| \mathbf{B} \right|_{\kappa \sigma} \right|_{\kappa \sigma} =$$

$$= \left| \left| \mathbf{A} \right|_{\kappa \sigma} \cdot \left| \mathbf{B} \right|_{\kappa \sigma} \right|_{\kappa \sigma} = \left| \mathbf{A} \cdot \mathbf{B} \right|_{\kappa \sigma},$$

где $\mathbf{A}, \mathbf{B} \in \mathfrak{F}$, или $\mathbf{A}, \mathbf{B} \in \mathfrak{A}$.

Использование системы описанных вычетов с учетом перечисленных их свойств позволяет на практике в некоторых случаях существенно снизить объем вычислений при анализе и оценке риска эволюции систем.

V. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Представленные в докладе результаты позволяют говорить о создании основ исчисления (абстрактных) дискретных рисков в условиях вероятностной неопределенности. Отметим, что доказательство всех

алгебраических фактов, приведенных в докладе, представлено в монографии автора [6]. Также необходимо отметить, что разработан прототип пакета прикладных программ, реализующий примитивы работы с рисками в представленном контексте [7, 8].

Отметим, что использование результатов, представленных в данном докладе, например, позволило повысить точность оценки меры риска «Value-at-Risk» для розничных кредитных портфелей, по отношению к классическому параметрическому методу [6]. При этом в рамках пакета прикладных программ «МультиМИР» версии 1.0 для 4-сценарной модели кредитного договора удается осуществить прямой расчет риска для однородного портфеля, включающего до порядка 500 договоров. Такого рода расчет доступен на обычном персональном компьютере.

Также отметим, что прототип пакета прикладных программ «МультиМИР» версии 1.0 был с успехом использован при разработке информационной системы поддержки принятия решений по управлению персоналом розничной подсистемы коммерческого банка [1], а также при технико-экономическом обосновании варианта резервирования сетевой компоненты отказоустойчивой масштабируемой вычислительной системы специального назначения [2].

VI. ЗАКЛЮЧЕНИЕ

Предложенный Вашему вниманию доклад представляет собой последовательное изложение результатов по обоснованию создания нового направления исследований, а именно Исчисления дискретных рисков в условиях вероятностной неопределенности. В качестве будущих направлений исследований в контексте исчисления дискретных рисков можно рассматривать проблему факторизации рисков, проблему оценки сложности риска, и другие связанные задачи.

БЛАГОДАРНОСТЬ

Хочу выразить благодарность коллективу кафедры Информатики и системного программирования Поволжского государственного технологического университета за конструктивное и плодотворное обсуждение как теоретических аспектов развития общей алгебраической теории риска, так и вопросов практической реализации и оптимизации ряда алгоритмов, представленных в пакете прикладных программ «МультиМИР» версии 1.0.

СПИСОК ЛИТЕРАТУРЫ

- [1] Бородин А.В. Архитектура информационной системы поддержки принятия решений по управлению персоналом розничной подсистемы коммерческого банка // Программные системы и вычислительные методы. 2014. № 2. С. 174-190.
- [2] Бородин А.В. Технико-экономическое обоснование варианта резервирования сетевой компоненты отказоустойчивой масштабируемой вычислительной системы специального назначения // Кибернетика и программирование. 2015. № 6. С. 55-70.
- [3] Виленкин Н.Я., Виленкин А.Н., Виленкин П.А. Комбинаторика. М.: МЦНМО, 2006. 400 с.
- [4] Клиффорд А., Престон Г. Алгебраическая теория полугрупп. Т. 1. М.: Мир, 1972. 285 с.
- [5] Петровский А.Б. Пространства множеств и мультимножеств. М.: Едиториал УРСС, 2003. 248 с.
- [6] Уразаева Т.А. Алгебраические методы анализа риска в развивающихся экономиках. Йошкар-Ола: Поволжский государственный технологический университет, 2017. 276 с.
- [7] Уразаева Т.А. О функциональности пакета прикладных программ «МультиМИР» // Современные проблемы и перспективы социально-экономического развития предприятий, отраслей, регионов. Йошкар-Ола: Поволжский государственный технологический университет, 2014. С. 261-265.
- [8] Уразаева Т.А. Программный комплекс «МультиМИР», версия 1.0. Свидетельство о регистрации программы для ЭВМ RUS 2014614663. Заявка № 2014610697 от 31.01.2014.
- [9] Уразаева Т.А. Реляционная теория риска и ее приложения к теоретико-игровым задачам нечисловой экономики // Статистика и Экономика. 2021. Т. 18. № 2. С. 12-21.
- [10] Arora S., Barak B. Computational Complexity: A Modern Approach. New York: Cambridge University Press, 2009. 594 p.
- [11] Golan J.S. Semirings and their Applications. Dordrecht; Boston: Kluwer Academic Publishers, 1999. 381 p.

О методах повышения точности измерения биосигнала для системы IoB

Е. А. Алтай, Р. К. Ускенбаева

Казахский национальный исследовательский
технический университет

aeldos@inbox.ru

А. В. Федоров

Университет ИТМО

avfedorov@itmo.ru

Аннотация. В данной работе рассмотрены методы фильтрации для повышения точности измерения биосигнала – кардиосигнала при влиянии электромагнитной помехи, оказывающей существенное влияние на работу системы IoB. Синтезирован одночастотный полиномиальный режекторный фильтр, реализующий фильтрацию данной помехи для повышения точности измерения кардиосигнала. Результаты реализации синтезированного режекторного фильтра продемонстрировали, что фильтр позволяет разделить информативную составляющую от искажающей помехи и повышает точность измерения сигнала при его реализации в сравнении с вейвлет фильтрами.

Ключевые слова: точность измерения, кардиосигнал, высокочастотные помехи, фильтрация, отношение сигнал/помеха, IoB

I. ВВЕДЕНИЕ

Развитие современной электронной базы и вычислительной техники привело к широкому распространению цифровых измерительных и управляющих устройств, а также цифровых каналов передачи данных, применяемых в системах Интернета вещей, в частности Интернета тел (IoB). В системе IoB основной акцент с вещей смещен на людей и представляет собой сеть связанных интеллектуальных датчиков, размещенных на поверхности кожи человека для контроля физиологических показателей человека и их передачи по беспроводной системе связи [1–3]. Известно [1–3], что в IoB системах контроль показателей человека может быть осуществлен как кратковременно, так и длительно. Среди перечисленных видов контроля физиологических показателей длительный контроль может выявить ряд значимых изменений и отклонений, которые не могут быть своевременно обнаружены при кратковременном контроле [4]. Поэтому, при реализации системы IoB, основное внимание уделяется длительному контролю физиологических показателей человека [1–3].

Система IoB, при длительном контроле физиологических показателей с помощью специальных интеллектуальных датчиков, осуществляет регистрацию биосигналов сердца, мышц, головного мозга, а также позволяет оценить насыщение крови кислородом, а также вариацию сердечного ритма с целью контроля работы синусового узла [1–3] и ряда других значимых показателей. Однако, несмотря на это, во время длительной регистрации перечисленных биосигналов, а именно биосигналов сердца – кардиосигнала неизбежно

влияют помехи [1–3], что существенно усложняет работу системы регистрации и обработки [4].

Согласно опубликованным данным [6], при длительной регистрации кардиосигнала неизбежно влияние низкочастотных помех дрейфа изолиний и артефактов движения, а также высокочастотных мышечных и электрических помех. Перечисленные помехи, вызванные индивидуальными особенностями человека, относятся к физиологическим помехам. Для их компенсации необходимо многократно проводить эмпирическую оценку параметров натуральных помех такого типа. Электромагнитные помехи, известные как электрические помехи, ввиду их высокой мощности [6] и узкой полосы [5], оказывают существенное влияние на работу системы IoB. Влияние такой помехи приводит к снижению точности измерения амплитудных и временных параметров кардиосигнала и эффективности методов, предназначенных для компенсации данной помехи.

Необходимо подчеркнуть, что точность измерения информативных параметров сигнала при влиянии помехи может быть достигнута за счет реализации методов фильтрации [4–8]. В связи с этим для компенсации электромагнитной помехи, важной и актуальной задачей является оценка эффективности методов фильтрации данного сигнала, обеспечивающих устойчивость и точность обработки при измерении кардиосигнала.

Среди методов фильтрации с задачей компенсации электромагнитной помехи справляются режекторные фильтры [6], аппроксимированные полиномиальными моделями и вейвлет фильтр Добеши [9]. Класс данных фильтров во многом позволяет настраивать свои параметры к параметрам обрабатываемых сигналов для выделения информативных составляющих сигнала.

Известно [9], что режекторный фильтр в основном синтезируется с помощью полинома Баттерворта и реже полиномами Чебышева, Бесселя и Кауэра. В основном это связано с невысокими значениями точности обработки кардиосигнала и выработкой наибольших значений собственной ошибки данных фильтров, ассоциированных с их частотными характеристиками, полосы заграждения помехи и добротности Q режекторного фильтра. Кроме того, добротность также оказывает существенное влияние на точность обработки при измерении кардиосигнала и значение этого показателя может быть низким только в том случае [9], когда полоса заграждения помехи расширена.

В работах [10–11] результат синтеза режекторных фильтров выявил, что в основном для синтеза данных фильтров используется два подхода.

Первый подход основан на применении полосы заграждения B помехи, от значения которой сильно зависит добротность Q . Следовательно, точность фильтрации сигнала на выходе системы при влиянии узкополосной одночастотной помехи будет недостаточно высокой [6]. Кроме того, данный подход при синтезе фильтра сопровождается удвоением порядка режекторных фильтров [10,11]. Например, при использовании полинома первой степени $n=1$ для синтеза, порядок системы фильтрации равняется к двум, то есть $n=2$, что приводит к изменению характеристики фильтра.

Второй подход основан на параллельном соединении передаточных функций полиномиальных фильтров верхних и нижних частот (ФВЧ и ФНЧ), где нижняя граница заграждения фильтра характеризуется частотой среза ω_1 для ФВЧ и верхней частотой среза ω_2 для ФНЧ. Общая передаточная функция $W(s)$ такого режекторного фильтра рассчитывается сложением двух передаточных функций ФВЧ и ФНЧ. Данный подход от первого подхода отличается тем, что вместо полосы заграждения помехи B используются граничные частоты среза ω_1 и ω_2 , задающие ширину режекции помехи, изменяющимся в определенном диапазоне частот.

Вместе с тем, несмотря на широкое использование перечисленных методов режекторной фильтрации электромагнитной помехи, в настоящее время вопрос синтеза именно одночастотного (ω) высокочастотного полиномиального режекторного фильтра для обработки кардиосигнала все еще остается открытым. Таким образом, синтез одночастотного режекторного фильтра, позволяющего повысить точность обработки при измерении кардиосигнала по отношению к данной помехе является важной задачей, требующей комплексного подхода к решению.

В отличие от ранее опубликованных работ [6, 11, 12], посвященных режекторной фильтрации кардиосигнала, в рассматриваемой работе для оценки точности измерения кардиосигнала представлены результаты оценки эффективности методов, повышающих точность обработки данного сигнала при фильтрации помехи. Рассматривается одночастотный и высокочастотный полиномиальный режекторный фильтр и сравнивается его результативность с хорошо зарекомендовавшим себя вейвлет фильтром Добеши.

Целью рассматриваемой работы является повышение точности измерения кардиосигнала по отношению к высокочастотной помехе, оказывающей существенное влияние на работу системы IoB.

II. ПОСТАНОВКА ЗАДАЧИ

Рассматривается система режекторной фильтрации кардиосигналов с общей передаточной функцией $W(s) = G(s)/H(s)$, где $G(s)$ и $H(s)$ – полиномы числителя и знаменателя $W(s)$ по степеням s . Пусть полиномы числителя $G(s)$ и знаменателя $H(s)$ имеют фиксированные коэффициенты при одночастотной (ω) частоте среза фильтра и порядка (n), определяющие вид

размещения корней полинома передаточной функции $W(s)$. Передаточная функция $W(s)$ системы фильтрации имеет порядок $n=2$, полученный при синтезе режекторного фильтра, состоящего из звеньев $n=2$ при $W_1(s)$ и $n=2$ при $W_2(s)$, как $W(s)=W_1(s)+W_2(s)$ и имеет вид:

$$\left\{ \begin{aligned} W_1(s) &= W_{\text{нормированный}}(s) \Big|_{s \rightarrow \frac{s}{\omega_c}} = \frac{1}{\left(\frac{s}{\omega_c}\right)^2 + 1,414 \frac{s}{\omega_c} + 1} = \\ &= \frac{1}{s^2 + 1,414 \omega_c s + \omega_c^2} \\ W_2(s) &= W_{\text{нормированный}}(s) \Big|_{s \rightarrow \frac{\omega_c}{s}} = \frac{1}{\left(\frac{\omega_c}{s}\right)^2 + 1,414 \frac{\omega_c}{s} + 1} = \\ &= \frac{1}{s^2 + 1,414 \omega_c s + \omega_c^2} \end{aligned} \right.$$

Передаточная функция $W(s)$ одночастотного высокочастотного режекторного фильтра одночастотной частотой среза (ω) имеет вид:

$$W(s) = W_1(s) + W_2(s) = \frac{s^2}{s^2 + 1,414 \omega_c s + \omega_c^2} + \frac{\omega_c^2}{s^2 + 1,414 \omega_c s + \omega_c^2} = \frac{s^2 + \omega_c^2}{s^2 + 1,414 \omega_c s + \omega_c^2} \quad (1)$$

Решаемая задача обработки кардиосигнала имеет вид:

$$x(q) = s(q) + \xi(q), \quad (2)$$

где q – отсчеты измерений кардиосигнала, $x(q)$ – зашумленный кардиосигнал, $s(q)$ – не зашумленный сигнал, $\xi(q)$ – высокочастотная электромагнитная помеха. Для формирования аддитивной смеси кардиосигнала и помехи (2) рассматривается модель высокочастотной электромагнитной помехи $\xi(q)$ вида

$$\xi(q) = A \sin\left(2\pi \frac{f_\xi}{f_d} k + \varphi\right),$$

где A – амплитуда гармоники помехи, f_ξ – частота, f_d – частота дискретизации, k – порядковый номер отсчетов помехи, φ – начальная фаза сигнала гармоники помехи.

Ставится задача выделения информативных составляющих кардиосигнала $s(q)$ из аддитивной смеси (2), регистрируемого в виде дискретными отсчетами q , путем узкополосной полиномиальной режекторной фильтрации (1).

Рассматриваемая задача решена при допущении, что амплитуда помехи кардиосигнала известна, что является типовым допущением при решении задачи обработки биосигнала для представления результатов фильтрации.

III. ФИЛЬТРАЦИЯ КАРДИОСИГНАЛА

Согласно допущению, рассмотренного в разделе «Постановка задачи» проведена фильтрация сигнала, результаты которых представлены ниже на рис. 1–4.

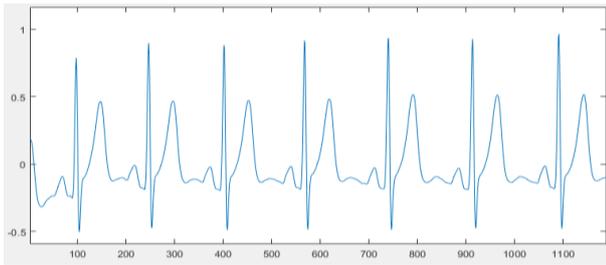


Рис. 1. Рис. 1. Тестовый (чистый) кардиосигнал. По осям ординат – амплитуда (мВ), абсцисс – время (с)

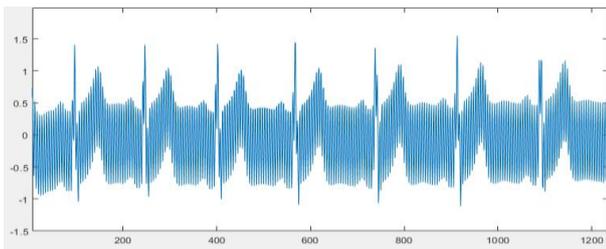


Рис. 2. Зашумленный тестовый кардиосигнал. По осям ординат – амплитуда (мВ), абсцисс – время (с)

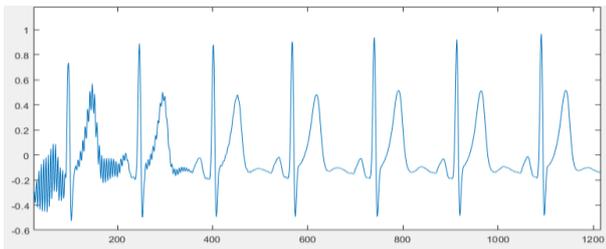


Рис. 3. Отфильтрованный сигнал на выходе синтезированного режекторного фильтра. По осям ординат – амплитуда (мВ), абсцисс – время (с)

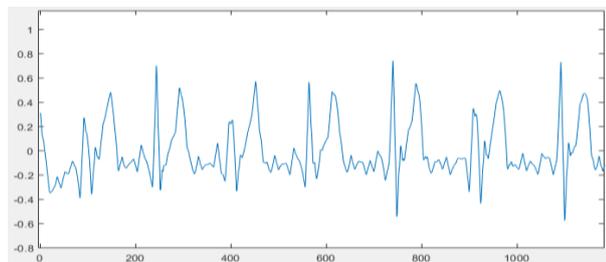


Рис. 4. Отфильтрованный сигнал на выходе вейвлет фильтра [9]. По осям ординат – амплитуда (мВ), абсцисс – время (с)

Из выше представленных амплитудно-временных диаграмм заметим, что обработка электрокардосигнала синтезированным режекторным фильтром позволяет отфильтровать составляющие сигнала при минимальных искажениях в сравнении с вейвлет фильтром. Вейвлет фильтр помимо сглаживания информативных участков сигнала, размывает форму сигнала, что также ранее было обнаружено в работах [13, 14]. При реализации синтезированного режекторного фильтра из-за наличия переходного процесса при фильтрации помехи, влияющей на работу системы IoB, искажается всего лишь участок 1 кардиоцикла сигнала. Наличие в числителе передаточной функции режекторного фильтра параметра $s^2 + \omega^2$ при фильтрации в сигналы вносят собственные искажения в виде затухающих колебаний и

наделяют переходную характеристику [11], что не противоречит результату, полученной в данной статье.

Известно, что на практике обработки кардосигналов чтобы не учитывать влияния внутренних переходных процессов таких фильтров, при проведении вторичной обработки, полученные результаты фильтрации первых двух участков кардиоцикла отбраковывается. Однако, отбраковка первых двух участков кардиоцикла по сравнению с результатами исследования [15] является наименьшей. В работе [15], из-за сильно искаженных участков обработанных кардосигналов, отбракована 40 кардиоциклов.

Вейвлет фильтрация кардосигнала проводилась на основе мягкой пороговой обработке (soft threshold) при выборе материнской вейвлет функции Добеши 3 порядка гладкости и восьмого уровня разложения [9]. Данные параметры вейвлет фильтрации для оценки и сравнения результативности обработки кардосигнала выбраны из [9]. Следует отметить, что вейвлет функция Добеши выбранного порядка и уровня разложения сигнала в [9] применено исходя из способности данного метода для выделения информативных составляющих сигнала при низких значениях показателя сигнал/помеха.

Показатели отношение сигнал/помеха (ОСП) и среднее квадратическое отклонение (СКО), рассчитанные до и после реализации методов цифровой фильтрации, дают возможность количественно оценить точность измерения кардосигнала между зашумленным тестовым и отфильтрованным тестовым сигналами [12]. При выборе перечисленных показателей значение показателя СКО характеризует точность измерения, показатель ОСП помехоустойчивость (или устойчивость) системы фильтрации по отношению к электромагнитной помехе, оказывающей существенное влияние на работу системы IoB. Результат оценки представлен в табл. 1.

ТАБЛИЦА I. ОЦЕНКА ПОКАЗАТЕЛЕЙ

Показатели	Синтезированный режекторный фильтр	Вейвлет фильтр из [0]
ОСП до, дБ	19,095	19,095
ОСП после, дБ	38,624	21,772
ОСП улучшенный, дБ	19,529	2,677
СКО, мВ	0,0011	0,002

В табл. 1 продемонстрировано, что при влиянии помехи синтезированный полиномиальный режекторный фильтр превосходит вейвлет-фильтр по эффективности обработки сигнала на 19 дБ, тогда как для вейвлет фильтра значение показателя ОСП составляет 2,677 дБ, соответственно. Чем меньше значение показателя СКО на выходе фильтров, тем точнее измеряется составляющие кардосигнала при влиянии помехи.

IV. ЗАКЛЮЧЕНИЕ

В представленной работе получены результаты оценки точности измерения кардосигнала при влиянии электромагнитной помехи, оказывающей существенное влияние на работу системы IoB. Получены результаты расчета передаточной функции одночастотного и высокочастотного режекторного фильтра на основе полинома Баттерворта. Показано, что предложенный одночастотный режекторный фильтр подтверждает

сходство отфильтрованной формы сигнала с исходной формой при влиянии электромагнитной помехи. Результаты оценки показывают, что значение показателя отношение сигнал/помеха на выходе предложенного режекторного фильтра в среднем возрастает в 2 раза и значение ошибки фильтрации в сравнении с известным вейвлет фильтром уменьшается в 2 раза, что характеризует повышение результативности обработки кардиосигнала по точности и устойчивости к влияющей помехе.

СПИСОК ЛИТЕРАТУРЫ

- [1] Celik A., Eltawil A. M. The internet of bodies: The human body as an efficient and secure wireless channel // *IEEE Internet of Things Magazine*. 2022. №. 3. P. 114-120.
- [2] Mäkitalo N. et al. The internet of bodies needs a human data model // *IEEE Internet Computing*. 2020. №. 5. P. 28-37.
- [3] Celik A., Salama K.N., Eltawil A.M. The internet of bodies: A systematic survey on propagation characterization and channel modeling // *IEEE Internet of Things Journal*. 2021. №. 1. P. 321-345.
- [4] Altay Y.A., Kremlev A.S. A brief analysis of the use of non-linear time-frequency filtering for processing ECG signals // *Cardiometry*. 2018. №. 13. P. 96-98.
- [5] Altay Y.A., Kremlev A.S. Analysis and systematization of noise arising by long-term recording of ECG signal // 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EConRus). IEEE, 2018. С. 1053-1057.
- [6] Алтай Е.А., Кремлев А.С. Оценка точности обработки электрокардиосигнала методом каскадной широкополосной режекторной фильтрации // *Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление*. 2020. №. 2. С. 18-28.
- [7] Григорьев Е.Б. Метод и программный комплекс обработки многоканальных записей электрокардиосигнала для диагностики ишемической болезни сердца: автореф. дис. ... канд. техн. наук / СПбГЭТУ «ЛЭТИ». СПб., 2019. 18 с.
- [8] Красичков А.С., Григорьев Е.Б. Оценка статистических характеристик могографической помехи при многоканальной регистрации электрокардиосигнала // *Известия вузов России. Радиоэлектроника*. 2018. №6. С. 118-125.
- [9] Malghan P.G., Hota M.K. A review on ECG signal filtering techniques for rhythm analysis // *Research on Biomedical Engineering*. 2020. V. 36. P. 171-186.
- [10] Алтай Е. Исследование и разработка методов обработки сигналов акустической эмиссии: автореф. дисс. ... канд. техн. наук / Университет ИТМО. СПб., 2023. 270 с.
- [11] Алтай Е.А., Лямин А.В., Скаков Д.М., Келемсеит Н.Е. Каскадный режекторный фильтр с единичной обратной связью и улучшенной переходной характеристикой // V Международная научная конференция по проблемам управления в технических системах. С. 262-265.
- [12] Алтай Е.А. Методы фильтрации для оценки точности и помехоустойчивости системы обработки электрокардиосигналов // *Измерение. Мониторинг. Управление. Контроль*. 2024. №1. С. 53-64.
- [13] Rakshit M., Das S. An efficient ECG denoising methodology using empirical mode decomposition and adaptive switching mean filter // *Biomedical Signal Processing and Control*. 2018. V. 40. P. 140-148.
- [14] Федоров А.В., Алтай Е., Степанова К.А., Кузиванов Д.О. Статистическая оценка влияния величины сигнал/помеха на погрешность измерения параметров акустической эмиссии // *Научно-технический вестник информационных технологий, механики и оптики*. 2022. №. 6. С. 1205-1215.
- [15] Тихоненко ВМ., Шихалиев ДР. Ошибки при оценке данных Холтеровского мониторингирования // *Вестник Аритмологии*. 2015. № 82. С. 12-20.

Выявление требований к метрологическим характеристикам измерителя импульсного магнитного поля

С. К. Терехин

Санкт-Петербургский
государственный
электротехнический
университет «ЛЭТИ»
им. В.И. Ульянова (Ленина)
terehin.serghey@yandex.ru

С. В. Романцов

Санкт-Петербургский
государственный
электротехнический
университет «ЛЭТИ»
им. В.И. Ульянова (Ленина)
romantsov89@gmail.com

Н. В. Романцова

Санкт-Петербургский
государственный
электротехнический
университет «ЛЭТИ»
им. В.И. Ульянова (Ленина)
nvromantsova@mail.com

Аннотация. В докладе приведены требования к характеристикам измерителя импульсного магнитного поля на основе анализа нормативной документации. Приведен анализ нормативно-технических документов, содержащих требования к параметрам молнии. Выявлены основные характеристики измерителя для проведения испытаний импульсным магнитным полем молнии.

Ключевые слова: магнитное поле молнии; требования по параметрам импульсного магнитного поля; измерение импульсного магнитного поля

I. ВВЕДЕНИЕ

Испытания технических средств на устойчивость к импульсным магнитным полям молнии являются важной частью процесса разработки современной аппаратуры. Магнитные поля могут привести к сбоям и причинить непоправимый ущерб техническим средствам. Именно поэтому проведение испытаний позволяет выявить возможные уязвимости и разработать эффективные меры защиты и предотвращения потенциальных повреждений.

Одним из основных преимуществ проведения таких испытаний является возможность эффективного предотвращения серьезных потерь и крупных аварий. Благодаря проведению соответствующих тестов, разработчики и производители способны обнаружить потенциальные проблемы, которые могут возникнуть в результате воздействия магнитного поля молнии, и внести коррективы в конструкцию или функциональность устройств еще на стадии разработки. Проведение испытаний магнитным полем молнии особенно важно в разработке систем связи, авиационных и космических технологий. Уязвимость этих систем может привести к катастрофическим последствиям.

Импульсные магнитные поля возникают при воздействии молнии на строения, металлические конструкции, мачты антенн и коммутационных процессов в линиях электропередач. Для измерения параметров импульсного поля при испытаниях и экспериментальной оценки распределения магнитного

поля внутри экранированного сооружения при ударе молнии, можно использовать измерители импульсного магнитного поля.

II. МОЛНИЯ

При исследовании параметров тока молнии измеряют:

- пиковый ток молнии;
- максимальную производную тока (di/dt);
- среднюю скорость возрастания тока;
- время фронта импульса тока;
- длительность импульса тока;
- величину переданного заряда;
- интеграл действия.

Распределение этих параметров, приведенных в большинстве существующих стандартов по молниезащите, основано на измерениях Бергера. [1]

Измерение характеристик электрических и магнитных полей, вызываемых разрядом молнии, необходимы для изучения потенциально опасного влияния электромагнитных полей на различные электрические цепи и системы. Чувствительные электронные цепи уязвимы для электромагнитного воздействия. Измеренные электрические и магнитные поля могут быть использованы для проверки различных моделей молнии и оценки параметров молниевых разрядов.

III. МОДЕЛИРОВАНИЕ ИМПУЛЬСНОГО МАГНИТНОГО ПОЛЯ

Для испытаний на стойкость к импульсному магнитному полю, испытываемое техническое средство помещают в центре индукционной катушки, создающей равномерное магнитное поле. К катушке подключается импульсный генератор, который должен обеспечивать протекание тока требуемой формы и амплитуды. Для получения магнитного поля с неоднородностью не более +3 дБ используют катушки, состоящие из двух параллельных коаксиальных плоских колец (катушка Гельмгольца). Расстояние между средними плоскостями колец должно равняться половине среднего диаметра

обмотки колец. Средний диаметр кольца должен быть не менее чем в 2,5 раза больше габаритных размеров испытуемого объекта.

Изделие подвергают воздействию магнитного поля в соответствии с реальными условиями эксплуатации. Силу тока выбирают с таким расчетом, чтобы получить в центре катушки магнитное поле требуемой напряженности.

Средства измерений, используемые в процессе испытаний, должны быть удалены от магнитной катушки на расстояние, обеспечивающее независимость их показаний от влияния магнитного поля, создаваемого катушкой. Изделие считается выдержавшим испытание, если во время и после испытания его характеристики соответствуют требованиям, установленным в техническом условии на изделие.

Метрологическое обеспечение при экспериментальных исследованиях должно давать возможность получать в исследуемой точке при каждом испытании данные о значениях амплитудно-временных характеристик полей с суммарной погрешностью не выше $\pm 20\%$ для напряженности магнитного поля. Для измерения значений амплитудно-временных характеристик напряженности магнитного поля применяются датчики магнитного поля.

IV. ТРЕБОВАНИЯ К ПАРАМЕТРАМ МОЛНИИ В СТАНДАРТАХ

В различных национальных и международных стандартах приведены различные амплитудно-временные параметры и используются разные методики испытаний, что усложняет разработку испытательного оборудования.

В разделе 23 КТ-160G/14G [2] молния имитируется компонентами А, А_h, А/5, В, С*, С и D (рис. 1). Импульсы могут подаваться отдельно или одновременно по два или более.

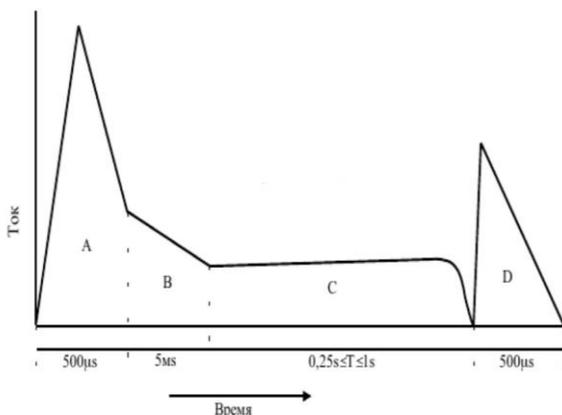


Рис. 1. Формы импульсов тока в разделе 23 КТ-160G/14G

Компонент А имеет интеграл действия (интеграл квадрата, изменяющегося по времени тока за время его протекания) $2 \times 10^6 \text{ A}^2\text{s}$, пиковую амплитуду 200 кА, полное время действия не более 500 мкс. Время нарастания тока от уровня 0,1 до 0,9 величины пикового тока не более 50 мкс. Может быть однополярным или колебательным (рис. 2).

Компонент А_h имеет интеграл действия $0,8 \times 10^6 \text{ A}^2\text{s}$, пиковую амплитуду 150 кА, полное время действия не более 500 мкс. Время нарастания тока от уровня 0,1 до 0,9 величины пикового тока не более 50 мкс. Может быть однополярным или колебательным (рис. 2).

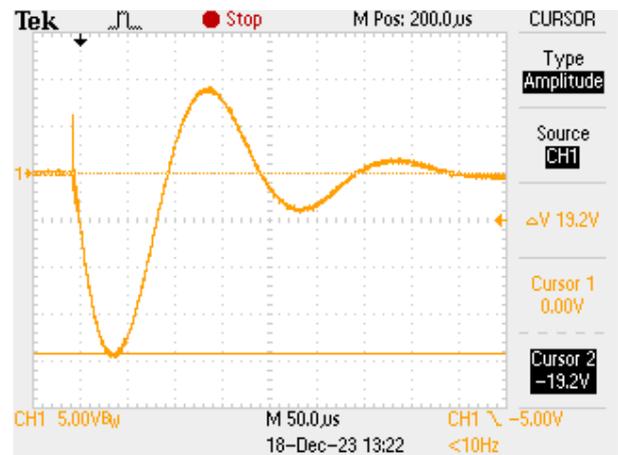


Рис. 2. Осциллограмма компонента А при испытаниях

Компонент А_h имеет интеграл действия $0,8 \times 10^6 \text{ A}^2\text{s}$, пиковую амплитуду 150 кА, полное время действия не более 500 мкс. Время нарастания тока от уровня 0,1 до 0,9 величины пикового тока не более 50 мкс. Может быть однополярным или колебательным.

Компонент А/5 имеет интеграл действия $0,08 \times 10^6 \text{ A}^2\text{s}$, пиковую амплитуду 40 кА, полное время действия не более 500 мкс. Время нарастания тока от уровня 0,1 до 0,9 величины пикового тока не более 50 мкс. Может быть однополярным или колебательным.

Компонент В имеет усредненную амплитуду 2 кА, обеспечивая перенос заряда 10 Кл полное время действия не более 5 мс. Должен быть однополярным.

Компонент С имеет усредненную амплитуду от 200 до 800 А, обеспечивает перенос заряда 200 Кл за время от 0,25 до 1,0 с. Должен быть однополярным.

Компонент С имеет усредненную амплитуду от 200 до 800 А, обеспечивает перенос заряда 200 Кл за время от 0,25 до 1,0 с. Должен быть однополярным.

Компонент С* имеет усредненную амплитуду не менее 400 А, обеспечивает перенос заряда 18 Кл за время от 45 мс. Должен быть однополярным.

Компонент D имеет интеграл действия $0,25 \times 10^6 \text{ A}^2\text{s}$, пиковую амплитуду 100 кА, полное время действия не более 500 мкс. Время нарастания тока от уровня 0,1 до 0,9 величины пикового тока не более 50 мкс. Может быть однополярным или колебательным.

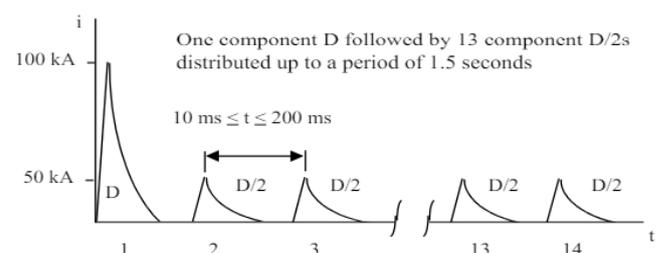


Рис. 3. Молния с несколькими ударами

Множественные молниевые разряды изображены на рис. 3 [3].

Множественные молниевые импульсы приведены на рис. 4.

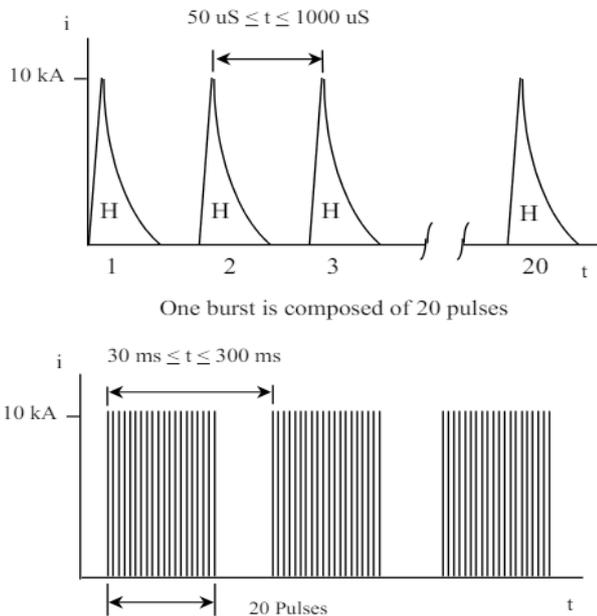


Рис. 4. Множественные молниевые импульсы

Квалификационные требования КТ-160G/14G не требуют испытаний импульсными магнитными полями молнии, поэтому приведенные параметры можно рассматривать только для справки.

В ГОСТ Р 50649-94 «Совместимость технических средств электромагнитная. Устойчивость к импульсному магнитному полю. Технические требования и методы испытаний» [4] форма испытательного импульса магнитного поля соответствует форме импульса тока имитатора микросекундной помехи большой энергии по ГОСТ Р 50007 [5] (длительность фронта 6,4 мкс ±30 %, длительность импульса 16 мкс ±20 %), напряженность от 100 до 1000 А/м.

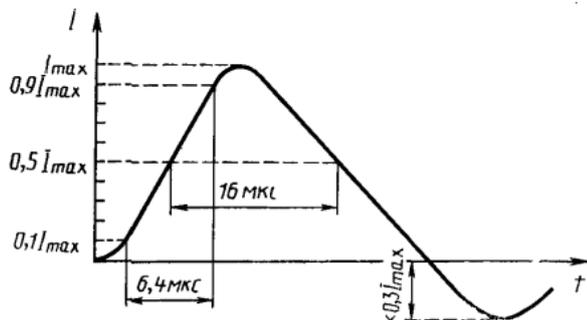


Рис. 5. Форма импульса тока на короткозамкнутом выходе испытательного генератора для создания микросекундного импульса помехи большой энергии

Для контроля напряженности магнитного поля, создаваемого испытательной установкой при испытаниях, должен использоваться датчик магнитного поля. В стандарте рекомендуют применять измерительную систему, состоящую из датчиков

магнитного поля с полосой пропускания не менее 5–10 МГц и осциллографа.

Международный стандарт ГОСТ 30585-98 «Совместимость технических средств электромагнитная. Стойкость к воздействию грозовых разрядов. Технические требования и методы испытаний» требует испытывать технические средства импульсным магнитным полем с длительностью фронта 2 мкс и длительностью полуспада 50 мкс. На рис. 3 длительность фронта $t_f = 1,25 T_1$, длительность полуспада $t_i = T_2$.

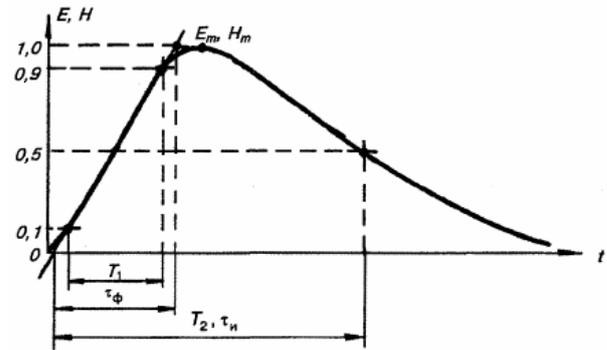


Рис. 6. Форма испытательного импульса напряженности грозовых электромагнитных полей

Параметры напряженности заданы от 50 до 300 А/м [6].

Параметры тока молнии, приведены в ГОСТ Р МЭК 62305-1-2010 [6].

V. ЗАКЛЮЧЕНИЕ

Нормативные документы определяют параметры электромагнитной среды, содержат требования к проектированию систем, определяют порядок испытаний оборудования.

Метрологическое обеспечение при экспериментальных исследованиях должно давать возможность получать в исследуемой точке при каждом испытании данные о значениях амплитудно-временных характеристик полей с суммарной погрешностью не выше ±20 % для напряженности магнитного поля.

СПИСОК ЛИТЕРАТУРЫ

- [1] Berger K., Anderson R.B., Kroninger H. Parameters of lightning flashes // Electra. 1975. Vol. 41. P.23–37.
- [2] Квалификационные требования КТ-160G/14G Условия эксплуатации и окружающей среды для бортового оборудования (Внешние воздействующие факторы – ВВФ). Требования нормы и методы испытания (в трех книгах). 2015.
- [3] MIL-STD-464C US. Department of defense interface standard electromagnetic environmental effects requirements for systems, 2010.
- [4] ГОСТ Р 50649-94 Совместимость технических средств электромагнитная. Устойчивость к импульсному магнитному полю. Технические требования и методы испытаний. Москва. 1994. 19 с.
- [5] ГОСТ Р 50007-92 Совместимость технических средств электромагнитная. Устойчивость к микросекундным импульсным помехам большой энергии. Технические требования и методы испытаний. Москва 1992. 15 с.
- [6] ГОСТ 30585-98 Совместимость технических средств электромагнитная. Стойкость к воздействию грозовых разрядов. Технические требования и методы испытаний. Минск, БелГИСС, 2005. 28 с.
- [7] ГОСТ Р МЭК 62305-1-2010 Менеджмент риска. Защита от молнии. Часть 1. Общие принципы. М.: Стандартинформ, 2011. 45 с.

Пифагоровы нечеткие числа и нечеткий анализ загрузки многоядерного процессора

Т. Г. Черноусова

Московский государственный технический университет им. Н.Э. Баумана
(национальный исследовательский университет)

E-mail chtg@bmstu.ru

Аннотация. Теория нечетких множеств, предложенная Л. Заде, получила свое дальнейшее развитие в теориях интуиционистских и пифагоровых нечетких множеств, позволивших вести перспективные прикладные исследования в ряде областей, в частности, исследования многоядерных процессоров. Анализируя ключевые понятия интуиционистских и пифагоровых нечетких множеств, показаны их принципиальные отличия между собой. На базе пифагоровых нечетких чисел предложен алгоритм нечеткого анализа загрузки ядер процессора, архитектурное решение которого выполнено по технологии Intel big. LITTLE.

Ключевые слова: пифагоровы нечеткие числа; функция оценки; функция точности; алгоритм нечеткого анализа загрузки многоядерного процессора.

I. ВВЕДЕНИЕ

Теория нечетких множеств (НМ) американского математика Л. Заде [1] послужила основой для развития теории интуиционистских нечетких множеств (ИНМ) (Intuitionistic Fuzzy Set (IFS)). Первая работа по ИНМ опубликована К. Атанасовым в 1983 году [2]. До появления его работы элементы нечеткого множества (НМ) (fuzzy set (FS)) рассматривались как элементы, характеризующиеся только степенью принадлежности μ , а степень непринадлежности ν находилась как разность между единицей и степенью принадлежности μ . В интуиционистской модели, предложенной К. Атанасовым, любой элемент НМ характеризуется не только степенью принадлежности μ его к данному множеству, но и степенью непринадлежности ν , которые отчасти не зависят друг от друга. Однако по свойству ИНМ сумма степеней принадлежности μ и непринадлежности ν любого его элемента не может быть больше единицы: $0 \leq \mu + \nu \leq 1$. Для преодоления ограничений, имеющих в теории ИНМ, разработана теория пифагоровых нечетких множеств (ПНМ) (Pythagorean Fuzzy Set (PFS)), которая впервые изложена Р. Ягером [3] и основана на теории пифагоровых чисел. Согласно Р. Ягеру для любого элемента ПНМ его степени принадлежности μ и непринадлежности ν этому множеству связаны между собой неравенством: $0 \leq \mu^2 + \nu^2 \leq 1$.

В статье представлен алгоритм нечеткого анализа загрузки многоядерного процессора, архитектурное решение которого построено согласно технологии Intel big. LITTLE. Оценка загрузки процессора

осуществляется с помощью пифагоровых нечетких чисел (ПНЧ) (Pythagorean fuzzy number (PFN)) [4]. Сравнение нечеткой загрузки высокопроизводительных Р-ядер и энергоэффективных Е-ядер процессора поможет существенно снизить временные затраты и энергопотребление на системном уровне при планировании исполнения ряда сложных приложений [5–9].

II. ХАРАКТЕРИСТИКА ИНМ И ПНМ

Определение НМ. НМ A_1 есть набор упорядоченных пар на универсуме X :

$$A_1 = \{(u, \mu_{A_1}(u)) | u \in X\}, \quad (1)$$

где функция принадлежности $\mu_{A_1}(u)$ элемента u множеству A_1 такая, что

$$\mu_{A_1} : X \rightarrow [0, 1]. \quad (2)$$

Функция принадлежности $\mu_{A_1}(u)$ устанавливает степень или меру принадлежности элемента u НМ A_1 [1, 2]. Из (2) следует, что $\mu_{A_1}(u)$ принимает значения на отрезке $[0, 1]$. Заметим, если $\mu_{A_1}(u) = 0$ или $\mu_{A_1}(u) = 1$, то $\mu_{A_1} : X \rightarrow \{0, 1\}$. Это значит, что $\mu_{A_1}(u)$ принимает значения на множестве $\{0, 1\}$. Тогда НМ A_1 (1) преобразуется в традиционное четкое множество.

Определение ИНМ. ИНМ B есть набор упорядоченных троек на универсуме X :

$$B = \{(u, \mu_B(u), \nu_B(u)) | u \in X\}, \quad (3)$$

где функция принадлежности $\mu_B(u)$ элемента u множеству B такая, что $\mu_B : X \rightarrow [0, 1]$, а функция непринадлежности $\nu_B(u)$ элемента u множеству B такая, что $\nu_B : X \rightarrow [0, 1]$. Из данного определения следует, что $\mu_B(u)$ и $\nu_B(u)$ принимают значения на отрезке $[0, 1]$, причем для любого $u \in X$ имеет место быть неравенство [13–15]:

$$0 \leq \mu_B(u) + \nu_B(u) \leq 1. \quad (4)$$

Укажем, что $\mu_B(u)$ устанавливает степень или меру принадлежности множеству B , а $\nu_B(u)$ устанавливает степень или меру непринадлежности множеству B . Тогда степень нерешительности (hesitation degree) ИНМ B или интуиционистский нечеткий индекс η_B ИНМ B определяется следующим образом:

$$\eta_B = 1 - (\mu_B(u) + \nu_B(u)), \quad (5)$$

При этом $\eta_B : X \rightarrow [0, 1]$. Если $\mu_B(u) + \nu_B(u) = 1$, то согласно (5) $\eta_B = 0$. Это значит, что ИНМ B становится НМ. Исходя из этого, можно заключить, что НМ есть частный случай ИНМ.

Определение ПНМ. ПНМ есть набор упорядоченных троек на универсуме X :

$$A = \{(u, \mu_A(u), \nu_A(u)) \mid u \in X\}, \quad (6)$$

где функция принадлежности $\mu_A : X \rightarrow [0, 1]$ и функция непринадлежности $\nu_A : X \rightarrow [0, 1]$ для всех $u \in X$ связаны между собой неравенством:

$$0 \leq \mu_A^2(u) + \nu_A^2(u) \leq 1. \quad (7)$$

Обозначим степень нерешительности ПНМ A или пифагоров нечеткий индекс (Pythagorean fuzzy index (PFI)) ПНМ A через π_A . Тогда согласно (5), (7) пифагоров нечеткий индекс (ПНИ) π_A для $\forall u \in X$ находится следующим образом:

$$\pi_A = \sqrt{1 - \mu_A^2(u) - \nu_A^2(u)}, \quad (8)$$

при этом $\pi_A : X \rightarrow [0, 1]$. Из (8) следует, что $\pi_A^2 + \mu_A^2(u) + \nu_A^2(u) = 1$. Если в этом равенстве $\pi_A = 0$, тогда $\mu_A^2(u) + \nu_A^2(u) = 1$. В табл. 1 показана сравнительная характеристика между ИНМ B (2) и ПНМ A (5).

Определение ПНЧ. Если задано ПНМ A (6), в котором функции принадлежности μ_A и непринадлежности ν_A удовлетворяют (7), а ПНИ π_A удовлетворяет (8), тогда говорят, что пара (μ_A, ν_A) есть ПНЧ. Обозначим эту пару через α :

$$\alpha = (\mu_A, \nu_A). \quad (9)$$

ПНЧ (9) характеризуется функцией оценки (score function) $S(\alpha)$, а также функцией точности (accuracy function) $H(\alpha)$ [3, 4]:

$$S(\alpha) = 0,5(1 + \mu_A^2 - \nu_A^2), \quad (10)$$

$$H(\alpha) = \mu_A^2 + \nu_A^2. \quad (11)$$

ТАБЛИЦА 1. СРАВНЕНИЕ ИНМ B И ПНМ A

ИНМ B	ПНМ A
1	2
$0 \leq \mu_B(u) + \nu_B(u) \leq 1$	$0 \leq \mu_A^2(u) + \nu_A^2(u) \leq 1$
$\eta_B = 1 - (\mu_B(u) + \nu_B(u))$	$\pi_A = \sqrt{1 - \mu_A^2(u) - \nu_A^2(u)}$
$\eta_B + \mu_B(u) + \nu_B(u) = 1$	$\pi_A^2 + \mu_A^2(u) + \nu_A^2(u) = 1$

Пусть заданы ПНЧ $\alpha_1 = (\mu_{A1}, \nu_{A1})$, $\alpha_2 = (\mu_{A2}, \nu_{A2})$. Согласно (10), (11) ПНЧ α_1 описывается $S(\alpha_1)$ и $H(\alpha_1)$:

$$S(\alpha_1) = 0,5(1 + \mu_{A1}^2 - \nu_{A1}^2), \quad (12)$$

$$H(\alpha_1) = \mu_{A1}^2 + \nu_{A1}^2. \quad (13)$$

Наряду с этим ПНЧ α_2 определяется $S(\alpha_2)$ и $H(\alpha_2)$:

$$S(\alpha_2) = 0,5(1 + \mu_{A2}^2 - \nu_{A2}^2), \quad (14)$$

$$H(\alpha_2) = \mu_{A2}^2 + \nu_{A2}^2. \quad (15)$$

Выполним сравнение ПНЧ α_1, α_2 по функциям μ_{A1} и μ_{A2} , а также ν_{A1} и ν_{A2} [4]:

$$\text{если } \mu_{A1} \geq \mu_{A2} \ \& \ \nu_{A1} \leq \nu_{A2}, \text{ то } \alpha_1 \geq \alpha_2. \quad (16)$$

Проведем сравнение ПНЧ α_1 и α_2 по функциям $S(\alpha_1)$ (12), $S(\alpha_2)$ (14) [4]:

$$\text{Если } S(\alpha_1) > S(\alpha_2), \text{ то } \alpha_1 > \alpha_2; \quad (17)$$

$$\text{если } S(\alpha_1) < S(\alpha_2), \text{ то } \alpha_1 < \alpha_2. \quad (18)$$

В случае, когда $S(\alpha_1) = S(\alpha_2)$, тогда производится сравнение функций $H(\alpha_1)$ (13) и $H(\alpha_2)$ (15) [4]:

$$\text{если } H(\alpha_1) > H(\alpha_2), \text{ то } \alpha_1 > \alpha_2; \quad (19)$$

$$\text{если } H(\alpha_1) < H(\alpha_2), \text{ то } \alpha_1 < \alpha_2; \quad (20)$$

$$\text{если } H(\alpha_1) = H(\alpha_2), \text{ то } \alpha_1 = \alpha_2. \quad (21)$$

Учитывая (13) и (15), пифагоровы нечеткие индексы π_{A1}, π_{A2} ПНЧ α_1 и α_2 находятся из равенств:

$$H(\alpha_1) + \pi_{A1}^2 = 1; \quad (22)$$

$$H(\alpha_2) + \pi_{A_2}^2 = 1. \quad (23)$$

Сравнение (22), (23) приводит к следующему результату:

$$\text{если } H(\alpha_1) > H(\alpha_2), \text{ то } \pi_{A_1} < \pi_{A_2} \text{ и } \alpha_1 > \alpha_2; \quad (24)$$

$$\text{если } H(\alpha_1) < H(\alpha_2), \text{ то } \pi_{A_1} > \pi_{A_2} \text{ и } \alpha_1 < \alpha_2; \quad (25)$$

$$\text{если } H(\alpha_1) = H(\alpha_2), \text{ то } \pi_{A_1} = \pi_{A_2} \text{ и } \alpha_1 = \alpha_2. \quad (26)$$

Итак, (16) – (26) позволяют найти соотношения между ПНЧ α_1 и α_2 , а также между ПНИ π_{A_1} , π_{A_2} .

III. ОЦЕНКА НЕЧЕТКОЙ ЗАГРУЗКИ ПРОЦЕССОРОВ

Определение загрузки процессора. Под загрузкой многопоточного многоядерного процессора понимается процент времени, в течение которого выполняется приложение или иное задание. Если этот процент времени представлен как ПНЧ (9), то будем говорить о нечеткой загрузке процессора.

Например, процессор для выполнения приложения предоставляет 100 % своих вычислительных ресурсов, но в процессе его исполнения используется только 88,8 % ресурсов. По определению ПНЧ (9), данный результат воспроизводится таким образом: $\alpha = (0.188, 0.118)$. По формулам (10), (11) вычислим функции оценки и точности:

$$S(\alpha) = 0.5(1 + 0.7885 - 0.0139) = 0.8873;$$

$$H(\alpha) = 0.7885 + 0.0139 = 0.8024.$$

По формуле (8) определим ПНИ π_A рассматриваемого ПНЧ:

$$\pi_A = \sqrt{1 - 0.8024} = \sqrt{0.1976} = 0.4445.$$

Согласно табл. I имеем:

$$\pi_A^2 + H(\alpha) = 0.1976 + 0.8024 = 1.$$

В табл. II показаны результаты загрузки одного из первых четырехъядерных процессоров Intel Core 2 Quad Q6600 (2,4 ГГц, 4 потока) при выполнении многопоточного приложения 1, решающего уравнение Лапласа [10]. Многопоточность этого приложения реализована посредством OpenMP. В зависимости от количества потоков измерения производились с помощью команд `gprof`, `time` операционной системы (ОС) LINUX. Считаем, что затраченное время на исполнение приложения 1 только одним потоком, является базовым. По отношению к базовому 2-поточное исполнение приложения 1 позволило ускорить его работу в 3.1 раза, 4-поточное исполнение – уже в 7.5 раз, а 8-поточное – в 4.3 раза. Таким образом, результаты измерений показали максимальную загрузку процессора при 4-поточном исполнении приложения 1.

ТАБЛИЦА II. ВЫПОЛНЕНИЕ ПРИЛОЖЕНИЯ 1

Потоки	Linux time-Real	Linux time – % Change	Linux time-Speedup
1	2	3	4
1	98.275	Base	Base
2	31.440	-68.0%	3.1
4	13.174	-86.6%	7.5
8	22.805	-76.8%	4.3

Согласно формулам (7) – (11) расчеты функций оценки $S(\alpha_i)$ и точности $H(\alpha_i)$, а также π_i для нечеткой загрузки процессора при исполнении приложения 1 в зависимости от числа потоков представлены в табл. III. Для базового однопоточного исполнения приложения 1 имеем: $H(\alpha_1) = S(\alpha_1) = 1$, $\pi_1 = 0$.

Отметим, что анализ нечеткой загрузки процессора, выполненный по $S(\alpha_i)$, $H(\alpha_i)$ и π_i является нечетким. Обозначим число потоков через I . Тогда для $\forall i(i = \overline{2, I})$ определим максимальное $S(\alpha_i)$ ($S_{\max}(\alpha_i)$), максимальное $H(\alpha_i)$ ($H_{\max}(\alpha_i)$) и минимальное π_i ($\pi_{i\min}$) по следующему алгоритму 1:

Шаг 1: Вычисление π_i по формуле (8).

Шаг 2: Вычисление $S(\alpha_i)$ по формуле (10).

Шаг 3: Вычисление $H(\alpha_i)$ по формуле (11).

Шаг 4: Нахождение $S_{\max}(\alpha_i)$ по соотношениям (17) – (21).

Шаг 5: Нахождение $\pi_{i\min}$ по соотношениям (24) – (26) одновременно влечет за собой определение $H_{\max}(\alpha_i)$ по тем же соотношениям.

Найденные по алгоритму 1 $S_{\max}(\alpha_4)$, $H_{\max}(\alpha_4)$, $\pi_{4\min}$ совместно указывают на максимальную нечеткую загрузку многоядерного процессора. Согласно табл. III максимальная нечеткая загрузка процессора производится при 4-поточном выполнении приложения 1, что соответствует данным табл. II.

Intel является первой компанией, которая реализовала концепцию одновременной многопоточности (Simultaneous Multithreading (SMT)), создав HT – технологию (Hyper – Threading Technology или HT – Technology). По этой технологии ОС рассматривает любое физическое ядро как два логических ядра, обеспечивая тем самым высокий уровень многозадачности. Другой инновацией Intel является использование `big.LITTLE` технологии, по которой высокопроизводительные P-ядра и энергоэффективные E-ядра с пониженным энергопотреблением располагаются на одном кристалле процессора. Так Intel Core i9 – 12900K, построенный на архитектуре Alder Lake, имеет на одном кристалле восемь

высокопроизводительных Р-ядер с НТ-технологией и восемь энергоэффективных Е-ядер. Р-ядра рассматриваемого процессора являются 2-поточными, а Е – ядра - однопоточными. Таким образом, Intel Core i9-12900К имеет 24 потока (16 +8). Обратим внимание, что Alder Lake располагает аппаратной технологией Intel Thread Director, согласно которой производится телеметрический сбор данных о состоянии процессора и передача их ОС Windows11. Учитывая полученные данные, планировщик ОС принимает решение о распределении потоков по конкретным ядрам процессора.

ТАБЛИЦА III. Нечеткий анализ приложения 1

Потоки	α_i	$H(\alpha_i)$	$S(\alpha_i)$	π_i
1	2	3	4	5
1	α_1 1.0, 0.0	1	1	0
2	α_2 0.68, 0.32	0.56	0.68	0.66
4	α_3 0.866, 0.134	0.77	0.86	0.48
8	α_4 0.768, 0.232	0.64	0.77	0.60

Произведем анализ выполнения одних и тех же приложений на Р-ядрах, Е-ядрах и покажем загрузку Intel Core i9-12900К при совместном их использовании. Отметим, что во всех рассматриваемых случаях базовая частота процессора составляет 3.9 ГГц.

В табл. IV представлены результаты выполнения приложения 2. В соответствии с алгоритмами машинного обучения данное приложение производит классификацию изображений. Посредством приложения 2 выполнен анализ 2500 фотографий по таким тегам, как «автомобиль», «собака», «цветок». В табл.V показаны расчеты $S(\alpha_i)$, $H(\alpha_i)$ и π_i для проведения нечеткого анализа приложения 2 [11].

Обозначим минимальные значения функций оценки и точности соответственно через $S_{\min}(\alpha_i)$ и $H_{\min}(\alpha_i)$, а максимальное значение π_i через $\pi_{i\max}$. Тогда исходя из архитектурных решений Intel по концепции big. LITTLE нечеткий анализ выполнения приложения 2 на Р-ядрах или Е-ядрах может производиться по алгоритму 2:

Шаг 1: Выполнение алгоритма 1.

Шаг 2: Анализ ПНЧ α_i , соответствующего максимальному значению $S_{\max}(\alpha_i)$: если $\mu_i^2 - \nu_i^2 \rightarrow 1$, то приложение имеет максимальную нечеткую загрузку на Р-ядрах с НТ-технологией; при отключении НТ-технологии – на Р-ядрах; в противном случае переход к шагу 3.

Шаг 3: Определение $S_{\min}(\alpha_i)$, $H_{\min}(\alpha_i)$, $\pi_{i\max}$ по соотношениям (17) – (21), (24) – (26).

Шаг 4: Анализ ПНЧ α_i , соответствующего минимальному значению $S_{\min}(\alpha_i)$: $\mu_i^2 - \nu_i^2 \rightarrow 0$, при этом $S_{\min}(\alpha_i) \rightarrow 0,5$; приложение назначается на энергоэффективные Е-ядра.

ТАБЛИЦА IV. ВЫПОЛНЕНИЕ ПРИЛОЖЕНИЯ 2

Ядра	Реальное время выполнения, с.	Загрузка ядер, %	Ускорение, количество раз
1	2	3	4
Е-ядра	286.8	Base	1
8Р-ядра без НТ-технологии	228.6	-20.3%	1.25
8Р-ядра с НТ-технологией	28.1	-90.2%	10.2
Core i9 - 12900К	48.4	-83.12%	5.9

ТАБЛИЦА V. Нечеткий анализ приложения 2

Ядра	α_i	$H(\alpha_i)$	$S(\alpha_i)$	π_i
1	2	3	4	5
Е-ядра	α_1 1.0, 0.0	1	1	0
8Р-ядра без НТ - технологии	α_2 0.2, 0.8	0.7	0.2	0.55
8Р-ядра с НТ-технологией	α_3 0.9, 0.1	0.82	0.9	0.42
Core i9-12900К	α_4 0.83, 0.17	0.72	0.83	0.53

Найденные по алгоритму 2 для приложения 2 $S_{\max}(\alpha_3)$, $H_{\max}(\alpha_3)$, $\pi_{3\min}$ совместно указывают на максимальную нечеткую загрузку Р-ядер с НТ-технологией ($0.9^2 - 0.1^2 = 0.8 \rightarrow 1$), что соответствует данным табл. III.

Приложение 3 реализует алгоритм сжатия данных WinRAR. В табл. VI представлены результаты выполнения этого приложения [11]. В табл.VII показаны расчеты $S(\alpha_i)$, $H(\alpha_i)$ и π_i для проведения нечеткого анализа приложения 3.

ТАБЛИЦА VI. ВЫПОЛНЕНИЕ ПРИЛОЖЕНИЯ 3

Ядра	Реальное время выполнения, с.	Загрузка ядер, %	Ускорение, количество Раз
1	2	3	4
Е-ядра	56.27	Base	1
8Р-ядра без НТ-технологии	27.38	-51.34 %	2.06
Core i9-12900К	20.65	-63.3%	2.7

В соответствии с шагом 2 алгоритма 2 для максимального ПНЧ α_3 определяется разность: $\mu_3^2 - \nu_3^2 = 0.63^2 - 0.37^2 = 0.26$. Согласно значению этой разности производится переход к шагу 3. На этом шаге определяются $S_{\min}(\alpha_2), H_{\min}(\alpha_2)$ и $\pi_{2\max}$. Выполним анализ $\alpha_2 : \mu_2^2 - \nu_2^2 = 0.51^2 - 0.49^2 = 0.02 \rightarrow 0$. Следовательно, приложение 3 необходимо назначить на Е-ядра.

ТАБЛИЦА VII. НЕЧЕТКИЙ АНАЛИЗ ПРИЛОЖЕНИЯ 3

Ядра	α_i	$H(\alpha_i)$	$S(\alpha_i)$	π_i
1	2	3	4	5
Е-ядра	α_1 1.0, 0.0	1	1	0
8P-ядра без HT-технологии	α_2 0.51, 0.49	0.5	0.51	0.71
Core i9-12900K	α_3 0.63, 0.37	0.53	0.63	0.69

Нечеткий анализ загрузки ядер Intel Core i9-12900K показывает, что приложения, связанные с большим объемом вычислений в таких областях, как искусственный интеллект, машинное обучение, необходимо назначать на высокопроизводительные P-ядра. В то же время, приложения, не связанные со значительными вычислениями, следует назначать на энергоэффективные Е-ядра. Сжатие данных является примером такого приложения. Обозначим нечеткую загрузку процессора через FP . Тогда по теории ПНЧ FP можно представить следующим образом: $FP = \{\alpha_i, S(\alpha_i), H(\alpha_i), \pi_i\}$.

IV. ЗАКЛЮЧЕНИЕ

Согласно теории ПНЧ функции оценки $S(\alpha_i)$ и точности $H(\alpha_i)$, а также пифагоров нечеткий индекс π_i позволяют сравнить выполнение приложений на ядрах процессоров, выполненных по технологии big.LITTLE. На примере архитектурного решения Intel Core i9-12900K показан алгоритм нечеткого анализа загрузки процессора для P-ядер и Е-ядер.

СПИСОК ЛИТЕРАТУРЫ

- [1] Zadeh L.A. Fuzzy sets // Information and Control. 1965. Т.8. №3. С. 338 – 353.
- [2] Atanassov K.T. Intuitionistic Fuzzy Sets // VII ITKR's Session, Sofia, Central Science and Technical Library, Bulgarian Academy of Sciences 1697/84. 1983.
- [3] Yager R.R. Pythagorean fuzzy subsets // Proceedings of joint IFSA World Congress and NAFIPS Annual Meeting, Edmonton. Canada. 2013. С. 57-61.
- [4] Zhang X., Xu Z. Extension of TOPSIS to multiple criteria decision making with Pythagorean fuzzy sets // International Journal of Intelligent Systems. 2014. Т.29. №12. С. 1061-1078.
- [5] Kachurin Yu.Yu., Kryukov A.V., Kananykhin O.A., Fedorin A.V. Calculation of contrast for computer simulated resolution chart image // Journal of Physics: Conference Series. 2021. Т. 2127. № 1.
- [6] Dimitrienko Yu., Gubareva E., Tumanov A., Modeling of aeroelastic composite plates vibrations based on asymptotic theory // E3S Web of Conferences. 2023. Т. 376. <https://doi.org/10.1051/e3sconf/202337601039>
- [7] Dimitrienko Yu., Zakharov A., Koryakov M. Simulation of energetic composite materials combustion // E3S Web of Conferences. 2023. Т. 376. <https://doi.org/10.1051/e3sconf/20233760101>
- [8] Kuzenov V.V., Ryzhkov S.V., Varaksin A.Yu. Simulation of Parameters of Plasma Dynamics of a Magneto Plasma Compressor // Applied Sciences (Switzerland). 2023. Т. 13. № 9. <http://dx.doi.org/10.3390/app13095538>
- [9] Blinov A.S., Malastowski N.S., Myagkov L.L. Numerical simulation of transient heat transfer of a catalyst substrate // AIP Conference Proceedings. 2023. Т.2700. <http://dx.doi.org/10.1063/5.0124919>
- [10] Prinslow G. Overview of Performance Measurement and Analytical Modeling Techniques for Multi – core Processors <https://www.cse.wustl.edu/~jain/cse567-11/ftp/multicore/index.html>
- [11] Intel Core i9-12900K E-Cores Only Performance Review. 2021. №19. <https://www.techpowerup.com/review/intel-core-i9-12900k-e-cores-only-performance>