

СРЕДА ДЛЯ ОРГАНИЗАЦИИ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ АЛГОРИТМОВ КОМПЛЕКСИРОВАНИЯ ИНТЕРВАЛОВ АГРЕГИРОВАНИЕМ ПРЕДПОЧТЕНИЙ

Шабрамов В.С.¹, Муравьев С.В.²

¹НИ ТПУ, ИШИТР, аспирант гр. А4-31, vss62@tpu.ru

²НИ ТПУ, ИШИТР, д.т.н., профессор, muravyov@tpu.ru

Аннотация

В статье рассматриваются основные принципы построения программной среды для организации экспериментальных исследований алгоритмов развиваемого научным коллективом метода комплексирования интервалов агрегированием предпочтений (IF&PA). Описаны основные требования к разрабатываемому на языке C++ программному обеспечению. Представлена функциональная часть программной среды, позволяющая проводить эксперименты с разными параметрами IF&PA, анализировать результаты исследований и формировать отчёты.

Ключевые слова: программная среда, язык программирования C++, комплексирование интервалов, агрегирование предпочтений.

Введение

Одной из ключевых задач в области обработки данных является нахождение точных и робастных оценок значений, основанных на анализе статистических выборок. Метод комплексирования интервалов агрегированием предпочтений (IF&PA) [1–3] представляет собой мощный инструмент для получения таких оценок. Метод был впервые предложен в 2011 и в настоящее время находится в стадии интенсивного развития с точки зрения как расширения сферы применения, так и совершенствования, и разработки новых алгоритмов. В этом контексте необходимо проведение большого объема экспериментальных исследований входящих в состав IF&PA алгоритмов, что, в свою очередь, требует наличия специализированной вычислительной среды, обеспечивающей удобство процедуры тестирования и анализа экспериментальных данных. Такая среда должна включать в себя модули для генерации входных данных, выполнения алгоритмов, оценивания показателей качества и визуализации полученных результатов. Разработка и оптимизация такой среды позволяет автоматизировать и ускорить процесс экспериментальных исследований алгоритмов и выявить их наиболее эффективные модификации.

Функции программы

Программа представляет собой специализированную среду для проведения экспериментальных исследований алгоритмов, входящих в состав метода IF&PA. Она должна функционировать удаленно и запускаться как веб-приложение, доступное посредством произвольного интернет-браузера. Программа должна предоставлять пользователю следующие функции:

- создание новых экспериментов и управление ими;
- добавление и настройка экспериментов с выбором различных алгоритмов;
- выполнение вычислений с возможностью использования многопоточных и параллельных технологий;
- анализ полученных результатов с созданием таблиц, графиков и других средств визуализации;
- формирование отчётов по экспериментам в заданном формате (Excel, PDF или CSV).

Программа ориентирована на исследователей и специалистов в области математического моделирования, анализа данных и искусственного интеллекта

Требования к программе

Исходя из назначения программы, были сформированы требования к конечной версии разрабатываемого продукта. Пользователю должны быть доступны следующие возможности.

При создании и управлении экспериментами:

- создание нового эксперимента, который открывается в рабочей зоне;
- открытие ранее созданного эксперимента;
- переключение между экспериментами;
- в эксперименте можно выполнять алгоритмы, просматривать их данные и управлять параметрами;
- генерация тестовых данных с выбором типа вероятностного распределения и объема данных;
- настройка эксперимента (например, для расчётов использовать графический процессор).

При добавлении и настройке экспериментов:

- выбор папки и/или конкретного файла с данными для эксперимента;
- выбор алгоритма из имеющейся библиотеки (например, алгоритм агрегирования предпочтений по правилу Кемени или алгоритм взвешенной медианы);
- ввод параметров алгоритмов (например, число m интервалов или число n альтернатив в алгоритмах агрегирования предпочтений по правилам Кемени и Борда);
- присвоение имени эксперименту;
- выбор способа организации параллельности вычислений.

При анализе результатов и формировании отчётов интерфейс с пользователем должен допускать:

- выбор эксперимента и просмотр его результатов;
- доступ к графикам и диаграммам для построения интервалов, визуализации данных, таблиц, профилей предпочтения, и других аналитических сведений;
- удобное отслеживание хода выполнения расчётов и архивирование полученных результатов с возможностью длительного хранения.
- редактирование формы отчетов и создание отчётов по проведенным экспериментам.

Технические требования

Исходя из поставленных требований к разработке программы, были выдвинуты технические требования. Программная среда должна быть реализована на языке программирования C++ с использованием платформы Qt для создания графического пользовательского интерфейса. Основными используемыми технологическими компонентами являются следующие:

- **CUDA** (Compute Unified Device Architecture) – программно-аппаратная архитектура параллельных вычислений на основе графических процессоров фирмы Nvidia [4] для ускорения вычислений сложных алгоритмов.
- **OpenMP** (Open Multi-Processing) – открытый стандарт многопоточного программирования для ускорения вычислительных задач на центральном процессоре [5].
- **SIMD** (Single Instruction, Multiple Data) – принцип организации вычислений, обеспечивающий параллельную обработку данных [6] и
- **Qt** – каркас для разработки кроссплатформенного пользовательского интерфейса [7].

Программа должна поддерживать работу на двух популярных операционных системах и проверять наличие аппаратной поддержки SIMD и CUDA для максимального быстродействия.

Описанные выше требования предъявляются к доступной пользователю внешней части (frontend) разрабатываемого веб-приложения. Внутренняя часть (backend) приложения, функционирующая на сервере и предназначенная для выполнения расчётов с помощью

исследуемых алгоритмов и обработки полученных результатов, представляет собой пополняемую библиотеку модулей. Модули реализуют методы, алгоритмы или процедуры, подлежащие исследованию с различными целями. Примерный список реализованных к настоящему времени модулей имеет следующий вид:

- RandProMatr – генерирует случайные ранжирования для заданных m (число ранжирований) и n (число альтернатив).
- RandInt – генерирует случайные интервалы для заданных m (число интервалов) и n (число альтернатив = число дискретных значений).
- Int2Rank – преобразует интервалы в ранжирования.
- ShapeMatr – формирует матрицу профиля из ранжирований.
- RecursAll – находит все ранжирования консенсуса (РК) по правилу Кемени.
- Convolution – находит свертку всех РК, найденных по правилу Кемени.
- Borda – находит РК по правилу Борда.
- IF&PA-B – процедура IF&PA в версии *base* на основе правила Кемени и свертки.
- IF&PA-A – процедура IF&PA в версии *advanced* на основе правила Кемени, свертки и самоуточнения.
- IF&PA-L – процедура IF&PA в версии *light* на основе правила Борда.
- IF&PA-LA – процедура IF&PA в версии *light advanced* на основе правила Борда и самоуточнения.
- WeightMed – вычисление взвешенной медианы.
- L2NormSim – метод наименьших квадратов для одномерного случая.
- L1NormSim – метод наименьших модулей в форме прямого спуска для одномерного случая.
- IF&PA-SR – простая линейная регрессия на основе IF&PA и т. д.

Все упомянутые выше модули должны поддерживать работу как в многопоточном режиме, так и в однопоточном.

Результаты разработки

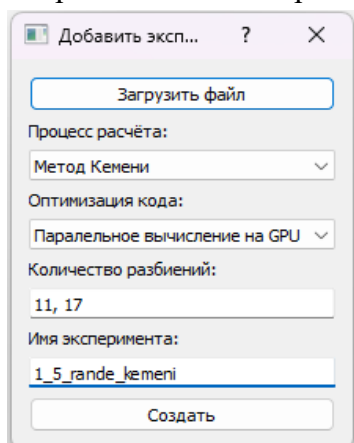
В соответствии с описанными выше требованиями была разработана программная среда, которая в настоящее время находится в стадии тестирования и отладки. С целью тестирования

в разработанной среде были проведены три эксперимента по сравнению оценок центра вероятностного распределения, найденных алгоритмами WeightMed (взвешенной медианы) и IF&PA-B (на основе правила Кемени и свертки). Для каждого эксперимента были сгенерированы по три набора данных со следующими настройками:

$n = 11$;
 $x_{\text{nom}} = 5, 10, 15$;
 $u_{\text{nom}} = 0,05x_{\text{nom}}$;
 число индивидуальных задач = 100;
 $m = 5, 15, 25, 35, \dots, 95$,

где x_{nom} – номинальное значение центра распределения,
 u_{nom} – его неопределенность.

Рис. 1. Пример меню для добавления нового эксперимента



Результаты тестирования показывают, что разработанная программная среда позволяет систематизировать, упростить и ускорить процесс организации и проведения экспериментальных исследований алгоритмов, а также формирования и визуализации их результатов. На рис. 1, 2 и 3 приведены примеры элементов пользовательского интерфейса программы.

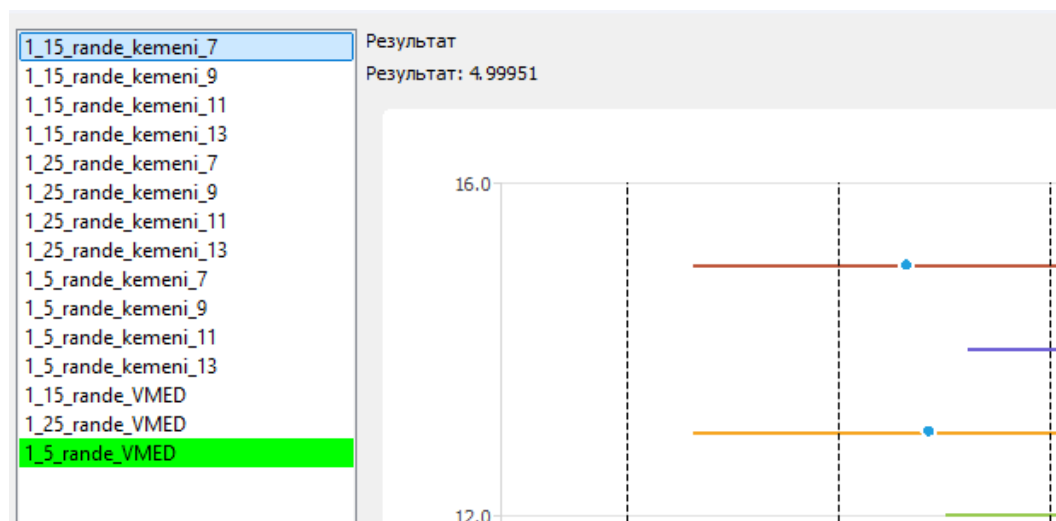


Рис. 2. Пример визуализации исходных данных эксперимента

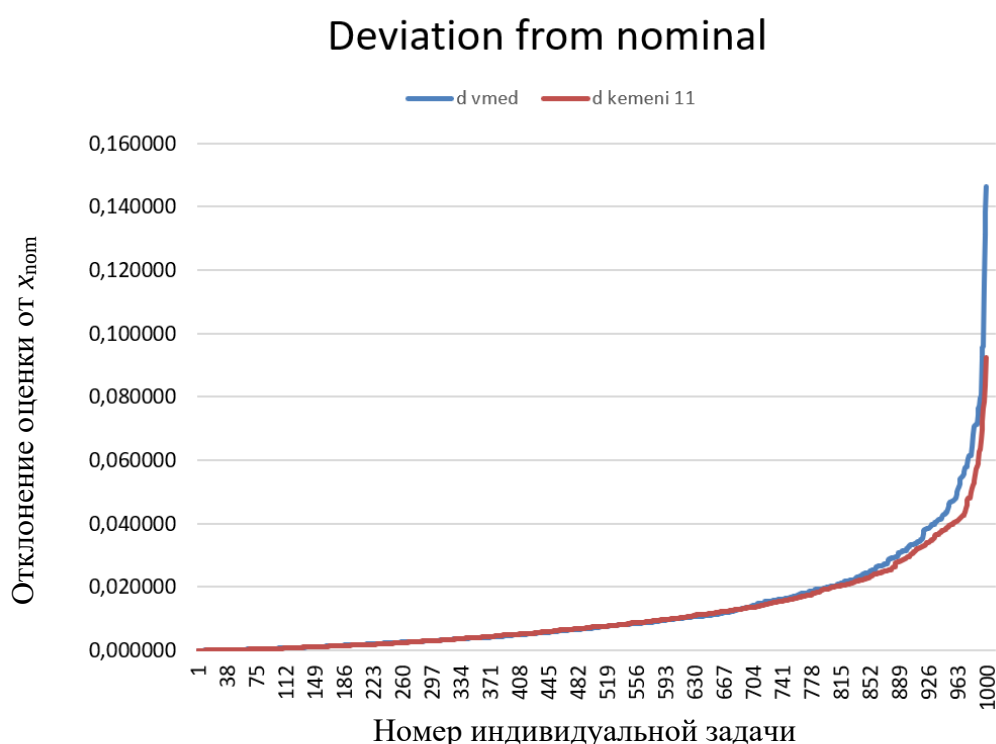


Рис. 3. Пример визуализации выходных данных эксперимента по сравнению оценок центра распределения, найденных алгоритмами *WeightMed* (кривая *dvmed*) и *IF&PA-B* (кривая *dkemeny11*)

Заключение

Разработана программная среда для организации экспериментальных исследований алгоритмов развиваемого научным коллективом метода комплексирования интервалов агрегированием предпочтений (IF&PA). Программная среда может функционировать удаленно и запускаться как веб-приложение, доступное через сеть интернет. В ходе работы были сформированы требования к frontend- и backend- частям веб-приложения, описаны используемые технологические компоненты. Для тестирования среды были проведены предварительные экспериментальные исследования разработанных алгоритмов, которые позволили выявить направления дальнейшего ее совершенствования. В частности,

предполагается разработка удобных средств обмена как входными, так и итоговыми данными между исследуемыми алгоритмами.

Список использованных источников

1. Muravyov S.V., Marinushkina I.A. Largest consistent subsets in interlaboratory comparisons: Preference aggregation approach // 14th Joint International IMEKO TC1, TC7, TC13 Symposium on Intelligent Quality Measurements - Theory, Education and Training (31 August – 2 September 2011, Ilmenau, Germany). – P. 69-73.
2. Muravyov S.V., Khudonogova L.I., Emelyanova E.Y. Interval data fusion with preference aggregation // Measurement. – 2018. – Vol. 116. – P. 621-630.
3. Muravyov S.V., Khudonogova L.I., Ho M.D., Analysis of heteroscedastic measurement data by the self-refining method of interval fusion with preference aggregation – IF&PA // Measurement. – 2021. – Vol. 183. – P. 109851.
4. CUDA Toolkit // NVIDIA Corporation. – 2025. – [Электронный ресурс]. – URL: developer.nvidia.com/cuda-toolkit (дата обращения: 23.11.2024).
5. OpenMP Application Programming Interface. Version 6.0 November 2024 [Электронный ресурс]. – URL: openmp.org/specifications/ (дата обращения: 24.11.2024).
7. Разновидности SIMD // Хабр. – [Электронный ресурс]. – URL: habr.com/ru/articles/441536/ (Дата обращения: 01.02.2024).
8. Qt Group Documentation // The Qt Company. – [Электронный ресурс]. – URL: doc.qt.io/ (Дата обращения: 05.03.2025).