

БИБЛИОТЕКА ДЛЯ ОБРАБОТКИ ДАННЫХ В ИНТЕРФЕЙСЕ МОЗГ-КОМПЬЮТЕР

Нуждин Юрий Олегович

Аспирант

НИЦ “Курчатовский институт”, Москва, Россия

E-mail: nuzhdin.urii@gmail.com

Создание интерфейсов мозг-компьютер требует написания алгоритмов классификации состояний мозга пользователя, регистрируемых различным оборудованием. При разработке алгоритмов обработки данных в интерфейсах мозг-компьютер требуется возможность проверять алгоритмы как на испытуемых во время эксперимента по управлению при помощи интерфейса (онлайн), так и на расчётных моделях по собранным данным (оффлайн). Возможность полностью повторить вычисления, произведённые во время эксперимента, необходима для эффективной отладки математических алгоритмов.

Обработка данных в режиме времени, приближенному к реальному (онлайн), требует, чтобы любые данные непрерывных измерений объединялись в небольшие блоки, над которыми совершаются преобразования по мере регистрации новых данных, далее будем называть такой режим обработки блочным. Выбор размера блока являет собой компромисс между допустимыми задержками, вызванными потребностью собрать блок, и накладными расходами на обработку блока, которые растут не медленнее, чем линейно при увеличении количества блоков.

При обработке данных оффлайн возможна оптимизация в виде склеивания мелких блоков в большой пакет данных и обработка в пакете, что в свою очередь позволяет воспользоваться различными оптимизациями алгоритмов для большого объёма данных. Далее будем именовать такой способ обработки данных пакетным.

Задача написания единой кодовой базы для онлайн и оффлайн обработки осложняется сильным снижением скорости библиотек и языков программирования для технических вычислений и статистической обработки при блочной обработке, из-за затрат на выделение памяти для промежуточных вычислений и накладных расходов на многочисленные вызовы обработки данных. Использовать медленную блочную обработку для оффлайн анализа представляется нецелесообразным, так как обычно требуется перебрать множество вариантов обработки для выбора оптимального. Написание же двух

версий алгоритмов - блочной и пакетной видится крайне нежелательным - кодовая база будет сильно различаться и возникает вероятность ошибок, приводящих к различиям в результатах вычислений в блочном и пакетном режиме.

Для устранения этих проблем была разработана библиотека на языке R, предлагающая инструменты для описания алгоритмов обработки в виде композиции функций, выполняющих операции над данными (трансформаций). Программисту требуется только описать эту композицию в функциональном стиле и воспользоваться одним из средств для запуска композиции в блочном или пакетном режиме. В пакетном режиме библиотека выполнит обработку над всеми данными сразу, что для R является оптимальным способом обработки большого массива данных. В блочном режиме данные будут обрабатываться поблочно.

Библиотека предлагает и инструменты для создания собственных трансформаций. Для этого требуется описать трансформацию как операцию преобразования блока данных фиксированного типа, но произвольного размера. Для обеспечения возможности обрабатывать блоки произвольного размера реализован механизм сохранения состояния между вызовами трансформации. Сохранение состояния позволяет выполнять накопление данных для обработки, тем самым обеспечивая возможность запускать произвольный алгоритм внутри трансформации. В библиотеку включены инструменты управления кольцевыми буферами и динамическими массивами, написанные на C++ и работающие быстрее, нежели реализации на языке R. Также программисту доступны инструменты для автоматического тестирования повторяемости результатов на произвольных блоках.

В настоящее время данная библиотека используется для разработки интерфейсов мозг-компьютер в НИЦ "Курчатовский институт" и МГУ им. М.В.Ломоносова.

Библиотека доступна в исходных кодах под лицензией GPL по адресу <https://github.com/tz-lom/Resonance-Rproj>

Работа выполнена при поддержке Российского научного фонда, грант 14-28-00234.