

Программное обеспечение для обработки данных энцефалографии

Оплачко Е.С., Рыкунов С.Д., Устинин М.Н.

ИМПБ РАН – филиал ИПМ им. М.В. Келдыша РАН

oplachkoe@gmail.com

Работа посвящена обзору программных продуктов, предназначенных для обработки данных энцефалографии. Концептуально данный набор программного обеспечения объединяет стек используемых технологий, таких как Matlab, а также подход к предоставлению инструментария пользователю, то есть необходимость установки на локальный компьютер. Также в работе описан облачный ресурс MathBrain, который предоставляет инструменты для обработки данных энцефалографии, используя модель «Приложение как Сервис». Ресурс предоставляет следующие методы анализа экспериментальных данных: прямое и обратное преобразования Фурье, анализ главных компонент, разложение на независимые компоненты, количественный анализ энцефалограмм, решение обратной задачи магнитоэнцефалографии на основе многоканальных спектральных данных.

Ключевые слова: энцефалография, облачные технологии.

Encephalography data handling software

Oplachko E.S., Rykunov S.D., Ustinin M.N.

IMPB RAS – Branch of KIAM RAS

The work is devoted to the review of the software products designed for the processing of encephalography data. Conceptually, this software has the same stack of technology, Matlab, and, the same approach of providing the toolset to users – with installation on local PC. Also, the work contains the description of cloud-based resource MathBrain, which works as a "Software as a Service" model. The resource provides access to the following analysis methods: direct and inverse Fourier transforms, Principal component analysis and Independent component analysis decompositions, quantitative analysis of encephalograms, magnetoencephalography inverse problem solution based on multichannel spectral data.

Key words: encephalography, cloud technologies.

1. Введение

Такие методы неинвазивного исследования, как электроэнцефалография (ЭЭГ) и магнитоэнцефалография (МЭГ), позволяют зафиксировать для дальнейшего анализа поля, вызываемые токами головного мозга. Результатом проведения указанных исследований являются большие объемы данных, к которым применяются различные математические методы для решения задачи определения источника, порождающего электрические токи, а также определения его характеристик. Для решения данной задачи используются возможности электронно-вычислительных машин больших мощностей, а также различное прикладное программное обеспечение (ПО). Все описанные ниже программные комплексы используют программное обеспечение MATLAB [1] как основу для разработки и хранения данных. Данный факт

обусловлен несколькими факторами, включая обилие возможностей построения графиков, обработки большого объема данных, простой собственный верхнеуровневый язык программирования и тесную интеграцию со многими скриптовыми языками программирования.

2. Программное обеспечение на основе Matlab, используемое для анализа электроэнцефалограмм и магнитоэнцефалограмм

Программный комплекс FieldTrip [2, 3] представляет собой набор инструментов для анализа МЭГ, ЭЭГ и других данных электроэнцефалограммы, которые доступны по открытому лицензионному соглашению GNU, т.е. может использоваться бесплатно. С технической точки зрения, программное обеспечение полностью разработано в среде MATLAB, языке программирования высокого уровня с

интерактивной средой для написания кода алгоритмов, анализа данных и визуализации, которая доступна для всех известных платформ. Стоит отметить, что MATLAB является широко известным инструментом для нейровизуализации, несмотря на достаточную дороговизну продукта. Набор инструментов FieldTrip содержит порядка 108 высокоуровневых и 858 низкоуровневых функций, которые суммарно составляют 103227 строк кода. Основной фокус делается на анализ данных неинвазивной и инвазивной электроэнцефалографии, включая записанные «спайки». С функциональной точки зрения набор утилит поддерживает обработку данных, анализ областей, относящихся к событиям, параметрический и не параметрический спектральный анализ, прямое и обратное моделирование источника, анализ связности, классификация, обработка данных в реальном времени, статистические выводы.

Программный комплекс ELAN [4] предоставляет широкий спектр предварительной обработки и анализа сигналов многих видов электрофизиологических данных. Типы данных, которые анализируются с помощью ELAN, включают ЭЭГ, МЭГ, а также инвазивные записи, такие как внутричерепное ЭЭГ и локальные полевые потенциалы (LFP) животных. В дополнение ко многим основным инструментам обработки сигналов, основной функционал ELAN включает топографическое отображение и частотно-временной анализ наряду с статистическими данными и интерактивными инструментами для визуализации и навигации в наборах данных. Весь этот функционал реализован на языке программирования C. Программный продукт ELAN доступен для 32 и 64-битных версий операционной системы Linux (Debian, Fedora, Ubuntu и CentOS). Операционные системы других производителей потребуют предварительной установки виртуальной машины для запуска Linux-подобных систем. Программный комплекс требует, как минимум, 128 мегабайт операционной памяти (однако, рекомендуется устанавливать 4 Gb RAM).

Программное обеспечение HADES [5, 6] представляет из себя программное обеспечение, работающее на основе приложения MATLAB, код которого находится в общем доступе. ПО имеет графический пользовательский интерфейс, также основанный на ПО MATLAB, и соответственно, требует для своей работы запущенного в это время ПО MATLAB. Программный комплекс HADES был написан и протестирован для версии MATLAB 7.9.0. Входные данные могут быть представлены в стандартном формате .mat или в виде текстового файла (кодировка ASCII). Также поддерживается формат .fif при подключении функций MNE.

Программное обеспечение Brainstorm [7] представляет из себя приложение с открытым кодом, основанное преимущественно на скриптах MATLAB. Интерфейс приложения написан с

использованием таких языков программирования, как Java/Swing, встроенных в среду разработки MATLAB. Использование такого набора программных продуктов обеспечивает решению Brainstorm полную независимость от операционной системы и другого ПО. Также стоит отметить, что Brainstorm не требует установленной версии MATLAB, что позволяет пользователю сэкономить на лицензии продукта. Все функции программного комплекса доступны через графический интерфейс (GUI) без какого-либо взаимодействия со средой MATLAB. Анализ в системе Brainstorm затрагивает работу с различными данными, такими как МЭГ, ЭЭГ, результаты магнитно-резонансной томографии.

Основная цель программного комплекса OpenMEEG [8, 9] является решение прямых задач, возникающих при обработке магнитной и электроэнцефалограммы. Программное обеспечение написано на языке программирования C/C++ с использованием библиотеки Intel MKL для ОС Windows и ATLAS (BLAS/LAPACK) для Linux, которые используются для достижения точных расчетов в линейной алгебре. Также при разработке системы была использована библиотека MATIO для обеспечения совместимости с ПО Matlab для операционной системы iOS. Исходный код программного продукта располагается на платформе INRIA GForge и доступен для всех пользователей через публичную версию системы.

Приложение EEGLAB [10] представляет собой кроссплатформенное программное обеспечение, основанное на программной среде MATLAB, предназначенное для обработки экспериментальных данных ЭЭГ с различным количеством каналов. Доступный для пользователя функционал включает обработку данных ЭЭГ с учетом каналов и сведений о событиях (с возможностью импорта), визуализацию. Логически разделяют два функциональных слоя EEGLAB: 1) верхний уровень состоит из функций, позволяющих пользователю взаимодействовать с ПО через графический интерфейс, что дает возможность избежать прямого использования ПО MATLAB, меню EEGLAB позволяет управлять используемой приложением памятью, что выгодно, когда имеются ограничения со стороны аппаратной части; 2) средний уровень позволяет настраивать обработку данных под конкретные нужды с помощью истории команд и интерактивных функций. Опытные пользователи MATLAB могут использовать структуры данных EEGLAB и автономные функции обработки сигналов, чтобы написать собственный пакет для анализа экспериментальных данных. ПО EEGLAB находится в свободном доступе на условиях публичной лицензии GNU для некоммерческого использования.

3. Облачный ресурс MathBrain

Институт математических проблем биологии - филиал ИПМ им. М.В. Келдыша РАН разрабатывает проект Mathbrain, который представляет собой интернет-ресурс для анализа данных энцефалографии, доступ к инструменту осуществляется с помощью модели «Приложение как Сервис» (Software as a Service). Ресурс доступен по адресу www.mathbrain.ru. Система предоставляет пользователю инструмент для построения высокочастотного разложения в ряды Фурье полных спектров, а также разложения в ряды Фурье скользящим окном, восстановления временных рядов по заранее рассчитанному спектру и другие функции. С точки зрения технической реализации программный комплекс представлен набором скриптов Python, которые выполняют роль основного вычислительного механизма с использованием библиотек numpy, scipy и numexpr [11].

Архитектура решения построена на микросервисах, что дает возможность балансировать нагрузку, масштабировать решение, и, в случае пиковых нагрузок, использовать ресурсы публичных облаков. Данный подход позволяет разложить компоненты системы в отдельные контейнеры и настроить взаимосвязи между ними. При этом каждый из таких контейнеров может быть запущен на разных узлах, в зависимости от нагрузки и очереди задач. В качестве платформы была выбрана технология Docker Swarm [12, 13]. Все узлы кластера Docker являются bare-metal серверами под управлением операционной системы CentOS.

Архитектура ресурса MathBrain содержит три узла. Один из них является управляющим, остальные – узлы для выполнения задач пользователя. Управляющий узел содержит репозиторий «слепков», которые могут быть реализованы в контейнере, а также веб-интерфейс Portainer. Основная концепция облачного ресурса – это возможность быстрой масштабируемости. В такой ситуации для расширения локальных аппаратных ресурсов можно использовать публичные облачные мощности. Существует три основных поставщика публичных облаков: Google Cloud Platform [14], Amazon AWS и Microsoft Azure [15].

При создании гибридной инфраструктуры [16] (использование как локальных аппаратных мощностей, так и публичных ресурсов), создается один или несколько контейнеров в публичном облаке, после отработки задачи, такой контейнер удаляется.

С точки зрения интерфейса, пользователю доступны следующие основные функции:

- аппроксимация спектра с задаваемыми параметрами промежутка частот, который будет анализироваться;

- восстановление спектра с задаваемыми параметрами промежутка частот, которые будут анализироваться, частоты дискретизации и промежуток времени;

- решение обратной задачи МЭГ;

- решение задачи с помощью преобразования Карунева–Лозва;

- решение задачи с помощью анализа независимых компонент;

- количественный анализ ряда.

Наиболее сложным в реализации было решение обратной задачи МЭГ [17], т.к. оно требовало асинхронной обработки данных, а также построения нескольких графиков на основе сформированных в реальном времени JSON данных. Более того, было предъявлено требование интерактивности данных графиков, что является ресурсоемкой задачей.

4. Заключение

Программный комплекс Mathbrain дает возможность пользователю решать задачи, связанные с обработкой экспериментальных данных МЭГ, такие как:

- аппроксимация спектра с задаваемыми параметрами промежутка частот, который будет анализироваться;

- восстановление спектра с задаваемыми параметрами промежутка частот, которые будут анализироваться, частоты дискретизации и промежуток времени;

- решение обратной задачи МЭГ;

- решение задачи с помощью преобразования Карунева–Лозва;

- решение задачи с помощью анализа независимых компонент;

- количественный анализ ряда.

При реализации ресурса были учтены также задачи и ограничения, связанные с большим объемом обрабатываемых данных и, в некоторых случаях, с трудоемким вычислительным процессом. Таким образом, были использованы новейшие методы и программные продукты, которые позволили реализовать облачный ресурс, предоставляемый по модели Приложение как Сервис (Software as a Service, SaaS). Реализованная архитектура ресурса позволяет не задумываться над требуемыми для анализа данных аппаратными мощностями и пользоваться всеми преимуществами облачных технологий. К таким преимуществам относятся независимость от операционной системы и отсутствие необходимости установки программного обеспечения (платного и бесплатного) на компьютер.

5. Благодарности

Работа выполнена при поддержке Программы Президиума РАН № 27 «Фундаментальные проблемы решения сложных практических задач с помощью суперкомпьютеров» и грантов РФФИ 17-

29-02178, 18-29-02087, 16-07-00937, 16-07-01000, 17-07-00677, 17-07-00686.

6. Список литературы

1. Quarteroni A., Saleri F., Gervasio P. Scientific Computing with MATLAB and Octave: Third Edition. In: *Texts in Computational Science and Engineering*. Berlin, Heidelberg: Springer, 2014. P. 1–4.
2. Ostenveld R., Fries P. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*. 2011. V. 2011. doi: [10.1155/2011/156869](https://doi.org/10.1155/2011/156869).
3. Stolk A., Todorovic A. Online and offline tools for head movement compensation in MEG. *NeuroImage*. 2013. V. 68. P. 39–48.
4. Aguera P., Jerbi K. ELAN: A Software Package for Analysis and Visualization of MEG, EEG, and LFP Signals. *Computational Intelligence and Neuroscience*. 2011. V. 2011. doi: [10.1155/2011/158970](https://doi.org/10.1155/2011/158970).
5. Campi C., Pascarella A. Highly Automated Dipole Estimation (HADES). *Computational Intelligence and Neuroscience*. 2011. V. 2011. doi: [10.1155/2011/982185](https://doi.org/10.1155/2011/982185).
6. Hämmäläinen M. MNE Software Overview. *MGH/HMS/MIT Athinoula A. Martinos Center for Biomedical Imaging*. 2009. URL: <https://wiki.aalto.fi/download/attachments/40600812/MNE-manual-2.7.pdf> (дата обращения: 27.08.2018).
7. Francois T., Baillet S. Brainstorm: A User-Friendly Application for MEG/EEG Analysis. *Computational Intelligence and Neuroscience*. 2011. V. 2011. doi: [10.1155/2011/879716](https://doi.org/10.1155/2011/879716).
8. Gramfort A., Papadopoulo A., Olivi E. OpenMEEG: opensource software for quasistatic bioelectromagnetics. *BioMedical Engineering OnLine*. 2010. V. 9. Article No. 45. doi: [10.1186/1475-925X-9-45](https://doi.org/10.1186/1475-925X-9-45).
9. Gramfort A., Papadopoulo T. Forward Field Computation with OpenMEEG. *Computational Intelligence and Neuroscience*. 2011. V. 2011. doi: [10.1155/2011/923703](https://doi.org/10.1155/2011/923703).
10. Delorme A., Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*. 2004. V. 134. № 1. P. 9–21.
11. Shell S. *An introduction to Numpy and SciPy*. 2014. URL: <http://www.engr.ucsb.edu/~shell/che210d/numpy.pdf> (дата обращения: 27.08.2018).
12. Lee H.-C. *DCCN Docker Swarm Cluster Documentation*. 2017. URL: <https://media.readthedocs.org/pdf/dccn-docker-swarm/latest/dccn-docker-swarm.pdf> (дата обращения: 27.08.2018).
13. De Souza J. Is Docker container based virtualization useful for Scientific purposes at DESY? *University College London*. 2015. URL: <http://www.desy.de/f/students/2015/reports/JoshuaDeSouza.pdf.gz> (дата обращения: 27.08.2018).
14. Liu Z., Cho S. Characterizing Machines and Workloads on a Google Cluster. In: *Proceedings of the Eighth International Workshop on Scheduling and Resource Management for Parallel and Distributed Systems (SRMPDS'12)*. 2012. P. 397–403.
15. Gandhi V.A., Kumbharana C.K. Comparative study of Amazon EC2 and Microsoft Azure cloud architecture. *International Journal of Advanced Networking Applications (IJANA)*. 2010. P. 117–123.
16. Calatrava Arroyo A. High Performance Scientific Computing over Hybrid Cloud Platforms. *Universitat Politècnica de València*. 2016. doi: [10.4995/Thesis/10251/75265](https://doi.org/10.4995/Thesis/10251/75265).
17. Устинин М.Н., Сычев В.В., Линас Р.Р. Интегрированный пакет программ MEGMRIAп для анализа и моделирования данных магнитной энцефалографии. *Математическая биология и биоинформатика*. 2013. Т. 8. № 2. С. 691–707.