

Оптимизация параллельного алгоритма глобального выравнивания протяженных последовательностей

Панкратов Антон Николаевич^{1,2*}

pan@impb.ru

Тетуев Руслан Курманбиеевич¹

ruslan.tetuev@gmail.com

Пятков Максим Иванович¹

mpryatkov@gmail.com

¹Пущино, Институт математических проблем биологии РАН – филиал Института прикладной математики им.М.В.Келдыша РАН

²Пущино, Пущинский государственный естественно-научный институт

DOI: 10.30826/IDP201842

Цель этой работы заключается в сборе лучших практик и построении наиболее общей модели парного глобального выравнивания с возможностью выравнивания как можно более длинных последовательностей. В результате универсальный алгоритм глобального выравнивания построен на основе параллельной реализации FastLSA алгоритма Needleman-Wunsch'a с произвольной матрицей подстановок и алгоритмом Gotoh'a для аффинной системы штрафов за пробелы.

При построении алгоритма используется стратегия отказа от хранения промежуточных данных в пользу повторяющихся вычислений с открывющейся возможностью распараллеливания вычислений и минимизации использования памяти. Найдены аналитически оптимальные значения параметров распараллеливания метода. Алгоритм реализован на языке программирования Javascript как веб-приложение. Распределенные вычисления реализованы на языке программирования PHP и производятся на клиентах, которые используют данный веб-сервис.

Алгоритм может использоваться при выравнивании текстовых последовательностей в произвольном алфавите.

Работа поддержана грантом РФФИ № 16-01-00692.

- [1] Тетуев Р.К., Пятков М.И., Панкратов А.Н. Параллельный алгоритм глобального выравнивания протяжённых аминокислотных и нуклеотидных последовательностей // Мат. биол. и биоинф., 2017.12(1). — С. 137–150. http://www.matbio.org/2017/Tetuev_12_137.pdf.