

Квантильное преобразование в задачах структурной биоинформатики

Научный руководитель – Ершов Николай Михайлович

Полуян Сергей Владимирович

Аспирант

Государственный университет «Дубна», Институт системного анализа и управления,
Дубна, Россия

E-mail: svpoluyan@gmail.com

В настоящей работе рассматриваются две задачи структурной биоинформатики: задача предсказания пространственной структуры пептида по аминокислотной последовательности и задача локального пептид-белок докинга с известным линейным интерфейсом связывания и линейной структурой пептида. При этом предполагается, что оптимальное состояние пептида и комплекса пептид-белок находится в глобальном минимуме свободной энергии, поиск которого можно осуществить минимизацией энергии посредством глобальной оптимизации.

Задача поиска структуры пептида ставится в непрерывном пространстве степеней свободы и включают в себя торсионные углы главной и боковых цепей. Задача докинга формулируется как задача минимизации энергии связывания. К параметрам добавляется вектор смещения пептида, кватернион для вращения, а также углы боковых цепей белка в непосредственной близости к пептиду. Для представления структур в полноатомном разрешении и вычисления энергии использовалось силовое поле Rosetta [1].

Целью работы является исследование применения стохастических алгоритмов непрерывной оптимизации к поставленным задачам. В силу сложности задач [2] проблематично использовать алгоритмы без учета дополнительной информации. Можно значительно упростить пространство поиска используя низкоэнергетические фрагменты, а также статистически известные распределения для углов. Однако, предлагаемые в настоящее время подходы к глобальной оптимизации, в частности, эволюционные алгоритмы, требуют непрерывного пространства поиска без ограничений, кроме границ поиска для каждого параметра. В связи с этим возникает проблема использования дополнительной информации, которая может быть решена использованием многомерного эмпирического квантильного преобразования. Разработанный подход позволяет создавать из множества фрагментов непрерывное пространство поиска, сводя задачу предсказания и докинга в непрерывный единичный гиперкуб. В задаче докинга с помощью такого преобразования возможно удержание пептида в локальной окрестности места связывания.

Результаты численных экспериментов показывают, что с помощью эволюционных алгоритмов оптимизации с использованием предложенного подхода возможно предсказание пептидов длиной несколько десятков аминокислотных остатков, а также докинг коротких пептидов. Реализация квантильного преобразования и библиотека для предсказания структур пептидов доступна в открытом репозитории [3] сервиса GitHub.

Автор выражает благодарность научному руководителю Н.М. Ершову, к.ф.-м.н., с.н.с. ВМК МГУ им. Ломоносова, и коллективу кластера HybriLIT ОИЯИ [4].

Источники и литература

- 1) Alford R., et al. The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. 2017. DOI: <https://doi.org/10.1101/106054>

- 2) Ciemny M., et al. Protein-peptide docking: opportunities and challenges // Drug Discovery Today, 2018, 23(8), pp. 1530–1537. DOI: <https://doi.org/10.1016/j.drudis.2018.05.006>
- 3) GitHub repositories: <https://github.com/poluyan>
- 4) Heterogeneous Platform HybriLIT: <http://hlit.jinr.ru/en>