

РЕШЕНИЕ ЗАДАЧИ О ПОИСКЕ ГАМИЛЬТОНОВА ПУТИ В ГРАФЕ С ПОМОЩЬЮ ДНК-ВЫЧИСЛЕНИЙ

Сергеенко А. Н., студентка кафедры биофизики СПбПУ,
anna.sergeenko98@gmail.com

Аннотация

Рассматривается алгоритм решения задачи о коммивояжере, основанный на дезоксирибонуклеиновых-вычислениях (ДНК-вычислениях). Приводится сравнение его трудоемкости с трудоемкостью переборных алгоритмов

Введение

Задача о коммивояжере является классической задачей комбинаторной оптимизации. Она относится к числу NP-полных и трансвычислительных задач, то есть она не может быть решена методом перебора для числа городов, больших 66, за время, меньшее возраста Вселенной [1]. На текущий момент существуют несколько путей ее решения, таких как: простейший перебор, муравьиный алгоритм, генетический алгоритм, алгоритм динамического программирования и др., однако большинство из них имеют экспоненциально растущую сложность. В 1994 году [2] Эдлман показал, что существует еще один возможный алгоритм нахождения кратчайшего пути, основанный на ДНК-вычислениях, трудоемкость которого не растет экспоненциально в зависимости от объема входных данных.

Поиск новых способов решения задачи о коммивояжере и сравнение их трудоемкостей представляет собой как фундаментальный интерес, так как на их основе возможно построить принцип работы новых компьютеров [3], [4], так и практический: задачи о построении оптимальных схем движений, распознавании траекторий, образов решаются с помощью задачи о коммивояжере.

Постановка задачи

В этой работе теоретически исследуется метод решения задачи с помощью ДНК-вычислений, а также сравнивается его трудоемкость с трудоемкостью метода перебора.

Предположим, что требуется найти кратчайший маршрут из Санкт-Петербурга в Москву с учетом того, что необходимо побывать единожды во Владимире, Ярославле, Костроме, Суздале и Иванове, причем возможные варианты передвижения следующие: Санкт-Петербург-Москва, Санкт-Петербург-Ярославль, Санкт-Петербург-Кострома, Ярославль-Суздаль, Ярославль-Кострома, Суздаль-Иваново, Суздаль-Владимир, Владимир-Ярославль, Кострома-Ярославль, Кострома-Владимир, Иваново-Кострома, Иваново-Владимир, Иваново-Москва.

Ход решения

С помощью ДНК-вычислений можно построить следующее решение этой задачи: пусть каждому городу соответствует некоторая случайная последовательность 20 нуклеотидов. Тогда для возможных путей передвижения запишем последовательности, построенные по следующему принципу: от города отправления берется комплементарная ему вторая половина нуклеотидов, от города прибытия – комплементарная ему первая половина.

На следующем этапе необходимо добавить все последовательности нуклеотидов, соответствующие городам и путям передвижения в одну пробирку (эксперимент проводится в биофизической лаборатории). Благодаря комплементарности одновременно образуются двухцепочечные ДНК, которые будут означать всевозможные варианты маршрутов. Далее, с помощью полимеразной цепной реакции (ПЦР) необходимо увеличить количество ДНК-молекул с началом и концом в заданных городах (Санкт-Петербург, Москва). В ходе ПЦР помимо копирования нужных маршрутов будут также накапливаться маршруты, начинающиеся в Санкт-Петербурге, но не заканчивающиеся в Москве, а также наоборот. Стоит отметить, что скорость копирования таких молекул будет заметно ниже, так что после ПЦР мы получим преобладающее количество нужных нам ДНК-молекул. Потом необходимо отобрать молекулы, которым соответствует нужное нам количество городов, так как по условию задачи нужно посетить все города по одному разу. Для этого отсортируем ДНК молекулы по длине с помощью электрофореза. После следует этап проверки посещения всех городов. Для этого используется метод аффинной хроматографии: ДНК-молекулы проверяются на наличие городов поочередно. На выходе мы получаем только те молекулы, которые содержат нужные нам последовательности в правильном количестве и порядке. На последнем этапе остается только секвенировать ДНК-молекулы и сопоставить полученные по-

последовательности нуклеотидов с возможными путями. Эти последовательности являются решением поставленной задачи.

Сравнение трудоемкостей

Трудоемкость реализации методов не зависит от количества городов N и является константой. Однако сложность алгоритма ДНК-вычислений растет как $N+86400$ в связи с необходимостью сопоставления последовательности нуклеотидов городам и путям передвижения. Трудоемкость метода перебора в случае несимметричной задачи растет как $(N-1)!$. Выбор константы для расчета трудоемкости алгоритма ДНК-вычислений является оценочным и обусловлен нормировкой на метод перебора, так как для проведения ПЦР, электрофореза, аффинной хроматографии и секвенирования ДНК требуется определенное постоянное количество времени. На Рис. 1 представлено сравнение трудоемкостей для алгоритма, основанного на ДНК-вычислениях и алгоритма перебора. Наблюдается существенное увеличение сложности для решения при увеличении количества городов при переборе вариантов, в то время как время с помощью ДНК-вычислений растет линейно и является возможным даже при больших N .

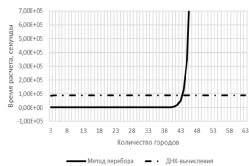


Рис. 1: Сравнение трудоемкостей метода перебора и алгоритма, основанного на ДНК-вычислениях

Заключение

Таким образом, применение алгоритма, основанного на ДНК-вычислениях, возможно для решения задачи о коммивояжере. Более того, этот алгоритм будет работать быстрее метода перебора для количества городов, больших 45.

Литература

- [1] Klir, George J. Facets of systems science. — Springer, 1991. — P. 121–128
- [2] L.M. Adleman. Molecular Computation of Solutions to Combinatorial Problems. — Science, New Series, Vol. 266, No. 5187 (Nov. 11, 1994) pp. 1021-102
- [3] Garg S, Shah S, Bui H, Song T, Mokhtar R, Reif J. Renewable Time-Responsive DNA Circuits. Small. 2018 Jul 18
- [4] Eshra, A.; Shah, S.; Song, T.; Reif, J. (2019). "Renewable DNA hairpin-based logic circuits". IEEE Transactions on Nanotechnology