

УТВЕРЖДАЮ:



Заместитель директора по
научно-организационной работе

Федерального исследовательского центра
«Красноярский научный центр СО РАН»

д-р хим. наук

Чесноков Николай Васильевич

21 марта 2017 г.

ОТЗЫВ

ведущей организации о диссертации

*Разработка алгоритмов выполнения молекулярного докинга с
использованием графических процессоров,*

представленной Фарковым Михаилом Александровичем

на соискание учёной степени кандидата технических наук

по специальности 05.13.17 - теоретические основы информатики

Диссертационная работа посвящена построению специального инструмента для решения трудоемкой оптимизационной задачи, а именно поиску (эволюционно) оптимальной функции энергии для связи между белками либо белками и лигандами. Данная задача весьма актуальна поскольку является основой для вычислений в широком классе вычислительных экспериментов в различных задачах молекулярной динамики: от фармакокинетики до протеомики. Тем самым, актуальность сформулированной цели исследования и поставленных для её достижения задач не вызывает сомнения.

В сколько-нибудь значимых с точки зрения приложений задачах характерная размерность соответствующей задачи оптимизации может быть

очень большой (число переменных может быть $\sim 10^3$ и выше). Столь высокая размерность пространства, а также сугубо нелинейный (хотя в большинстве моделей и самый простой из нелинейных - квадратичный) характер связей делает проблемы вычислений весьма нетривиальными и составляющими самостоятельную научную задачу.

Научная и практическая значимость диссертационной работы заключается в развитии программных средств, существенно улучшающих вычислительную сторону данной проблемы; несмотря на развитие вычислительной техники, прямые (методом грубой силы) подходы не являются удовлетворительными для сколько-нибудь ценных с практической точки зрения приложений этих методов.

Практическая значимость работы подтверждается актами о внедрении. Результаты исследования использовались ООО «Мобилфон» для решения задачи оптимизации распределения вычислений в многосерверной вычислительной среде, а также использовались ООО «ФанНано» для решения задачи расчёта кинетики реакции и транспорта на границе раздела «металлический микросетчатый электрод – органическая электрохромная композиция».

Разработанные алгоритмы могут быть использованы для выполнения биоинформационных исследований взаимодействия химических соединений. Помимо этого, предложенный алгоритм реализации метода дифференциальной эволюции может быть использован для ускорения решения широкого класса прикладных задач, требующих проведения большого количества процедур численной оптимизации.

Диссертационная работа состоит из Введения, четырёх Глав, Заключения и приложения. Первая глава содержит обзор существующих методов решения таких задач оптимизации; следует отметить, что обзор написан хорошо и полно: он является исчерпывающим как для понимания дальнейшего изложения материала, так и для обоснования выбора целей и

задач диссертационного исследования. Вторая глава посвящена описанию реализации метода дифференциальной эволюции для гетерогенного вычислительного устройства. В этой главе также приведены исчерпывающие сведения как по самому методу, так и по особенностям его реализации на графических процессорах, что, собственно, и является основной задачей диссертационного исследования.

Третья глава целиком посвящена описанию решения прикладной — с точки зрения диссертационной работы — задачи: собственно, расчётам молекулярного докинга. Молекулярный докинг — это процесс узнавания и удержания сравнительно большими биологическими макромолекулами (белками, в рамках настоящей диссертационной работы) других макромолекул либо малых молекул, играющих важную регуляторную роль в биологических процессах. Как обычно, (физические) модели докинга вызывают лёгкую оторопь у стороннего читателя, но за эти модели соискатель ответственности не несёт и, более того, их широкое распространение в практике моделирования взаимодействия делает их неизбежным элементом любого исследования в данной области.

Четвёртая глава посвящена описанию результатов тестирования разработанных алгоритмов, реализованных на различных аппаратных системах. Результаты этой главы показывают, что поставленная цель диссертационного исследования в общем и целом достигнута. Эта же глава (в совокупности с третьей) и порождают замечание. В работе ничего не говорится об устойчивости вычислений и полученных результатов. При этом устойчивость следует понимать (и анализировать) в разных смыслах, в частности

1) устойчивость к накоплению ошибок вычисления; возможно, эта проблема как-то решена в исходных алгоритмах, однако устойчивость этого типа при переносе алгоритмов в среду гетерогенных вычислителей как минимум заслуживает упоминания. Я не нашёл такого упоминания в тексте

диссертации;

2) «эволюционная» устойчивость: что, если в процессе расчётов один из вычислителей в ансамбле выдал неверные результаты и передал их другим устройствам? Будет данная ошибка нарастать или будет исключаться собственно эволюционным алгоритмом — собственно, в этом и состоит суть генетических алгоритмов: отбирать «хорошее» и выбраковывать «плохое». Однако уровень критически допустимых ошибок может сильно варьировать в зависимости от ситуации и требует оценок;

3) наконец, какова устойчивость в собственном смысле? Что будет с результатами вычислений, если мы возмутим исходные данные? А вдруг да пролезет через все кордоны и препоны какой-нибудь неожиданный синергизм и две (три, пять, достаточное количество) возмущённых данных дадут потерю устойчивости — при том, что по отдельности они такого эффекта могут и не давать.

Указанные недостатки не снижают высокой оценки работы. Полученные результаты являются новыми и достоверными, что позволяет их рассматривать как решение актуальной научной задачи. Основные результаты опубликованы в научных журналах и трудах конференций. Автореферат полно отражает содержание диссертации. Диссертационная работа М. А. Фаркова *Разработка алгоритмов выполнения молекулярного докинга с использованием графических процессоров* имеет внутренне единство и является завершённой научно-квалификационной работой, в которой на основании выполненных автором исследований содержится решение оптимизационной задачи поиска функции энергии для связи между белками и лигандами с применением современных высокопроизводительных вычислительных систем, в состав которых входят графические процессоры. Задача имеет важное значение для развития методов биоинформатики и вычислительного моделирования.

Диссертация соответствует критериям, установленным п.9

«Положения о порядке присуждения учёных степеней» постановления Правительства Российской Федерации от 24.09.2013 г. №842», а её автор Фарков Михаил Александрович достоин присуждения учёной степени кандидата технических наук.

Диссертационная работа была обсуждена и получила положительную оценку на заседании отдела вычислительной математики Института вычислительного моделирования СО РАН 21 02 2017 г., протокол № 2/2017.

Отзыв составил ведущий научный сотрудник отдела вычислительной математики Института вычислительного моделирования СО РАН - обособленного подразделения ФИЦ КНЦ СО РАН, д.ф.-м.н. Михаил Георгиевич Садовский.

Ведущий научный сотрудник отдела вычислительной математики
ИВМ СО РАН – обособленного подразделения ФИЦ КНЦ СО РАН
д.ф.-м.н.



Михаил Георгиевич Садовский

Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр «Красноярский научный центр Сибирского отделения Российской академии наук» (ФИЦ КНЦ СО РАН)»

Web-сайт: <http://icm.kscran.ru>. Адрес: 660036, г. Красноярск, Академгородок, д.50, стр. 44. Телефон: +7(391)243 27 56

Адрес электронной почты: sek@icm.krasn.ru