

РЕЦЕНЗИЯ

от: проф.д-р Никола Ив. Янев

Институт по математика и информатика,
Българска академия на науките, гр. София
ул. " Акад. Г. Бончев", 1113 София,
ел. поща: choby@math.bas.bg

Относно: дисертационен труд за получаване на образователната и научна степен „доктор”, на докторант Методи Георгиев Трайков на тема: **„Математически модели и алгоритми за предсказване на пространствената структура на протеини”**.

Област на висше образование - 4. Природни науки, математика и информатика. Професионално направление: 4.6. Информатика и компютърни науки

Научен ръководител: проф. д-р Петър Миланов

Представям настоящата РЕЦЕНЗИЯ, съгласно ЗРАСРБ, като член на научното жури по защитата на дисертационния труд на докторант Методи Георгиев Трайков, съгласно заповед № 1501/29.06.20017 г. на ректора на ЮЗУ „Неофит Рилски”, Благоевград.

Актуалност и цели на изследването

Една от загадките на природата, предизвикваща сериозен научен интерес е: какъв е механизма, който трансформира(свива) за броени секунди дадена протеинова нишка (първична структура) в in vitro среда в уникална тримерна (третична) конформация. В надпреварата за разгадаването са включени от терафлопови суперкомпютри , успешно имитиращи няколко микросекунди от процеса на свиване до обикновени лаптопи, които изпълняват алгоритми за решаване на задачи от комбинаторен характер. Тези задачи са свързани с т.н. NP модели, в основата на които са хидрофобните свойства на аминокоселините , формиращи първичната структура и хипотезата, че целта на свиването е конформация, максимално респектираща хидрофобността на аминокиселините от клас H. Освен редуцията от 20 вида аминокиселини до 2 вида H, P в тези модели се допуска дискретизация на конформационното

пространство чрез използване на различни решетки :квадратни, кубични, триъгълни, кристалоподобни и др., чрез които се редуцира и броя на конформациите до краен. Достъпността на формулировката на оптимизационните комбинаторни задачи и предизвикателната им сложност привличат сериозен брой участници за създаване на ефективни предимно евристични алгоритми. Последното достатъчно тясно определя тематиката на дисертацията в областта на биоинформатиката по-общо характеризирана като: интердисциплинарна област, развиваща методи и софтуерни средства за тълкуване на биологични данни.

Още по-конкретно резултатите са свързани със създаване на алгоритми (точни и евристични) за предсказване на тримерната структура на протеини с известна първична структура.

Дисертационният труд се състои от въведение, четири глави:

- (I) Въведение в проблема за нагъване на протеини, (II) Модели за нагъване на протеини, (III) Алгоритми за нагъване на протеини, (IV) Резултати и обсъждания

заклучение, изводи и цитирана литература. Общият му обем е 160 страници, от които 10 страници използвана литература. Включени са 80 фигури и 17 таблици. Цитираната литература обхваща 186 заглавия, от които 1 на кирилица и 185 на латиница. Структурата на дисертацията отговаря на общоприетите изисквания.

Резултати

Основните резултати (четвърта глава) са:

-сериозен обзор на алгоритмите и техниките за предсказване на естествената (native) структура на протеините (този резултат е достъпен само за българоезични читатели);

-точен и евристичен алгоритъм за намиране на оптимална структура върху двумерна решетка;
-точен и евристични алгоритми за намиране на оптимална структура върху тримерна решетка;
-основен принос в трансформиране на създадени от съавтори, целочислени модели във вход за търговски софтуер (CPLEX , GUROBI) за решаване на линейни целочислени задачи, а също и в нетривиалната задача за визуализация на решението.

Моделирането на конкретна оптимизационна задача като математическа задача за намиране на екстремум на линейна функция на n променливи е една от възможностите за получаване на решението на оптимизационната задача особено при наличие на софтуер за решаване на математическата задача. Точният алгоритъм, споменат в една от публикациите, освен като самостоятелно постижение за решаване на задачи с дължина на първичната протеинова структура до 50 аминокиселини е и основен компонент в евристиките, решаващи (приближено) произволно дълги редици. Оценката на значимостта на резултатите в случая е облекчена, защото се сравняват алгоритми по бързодействие и стойност на целева функция върху еталонни (benchmarks) задачи. Допълнителна характеристика е импакт фактора на списанията, скоростта на публикуване и в случая на journal of computational biology, анонсирането на една от публикациите като една от 10 –те с висок импакт.

За неудачно формулирани, считам следните твърдения:

- 1.(стр 5) недефиниран „проблем за нагъване на протеини „, е обявен за NP-пълн,
- 2.(стр 6,задача 5) „ Да се реализират компютърни симулации на процеса на нагъване на протеини с разработените алгоритми“ е или неясно поставена или не е изпълнена,
- 3.(стр 12) .“ Тези алгебрични ограничения позволяват разширяване на мощните методики на линейното програмиране, за да се достигне до разглеждане на комбинаторен оптимизационен проблем.“ – ограниченията за целочисленост не разширяват , а се преодоляват чрез линейното оптимизиране.

4.(раздел 4.1.1.1) “Математическият модел на проблема за нагъване на протеини в 2D е базиран на квадратна решетка с фиксиран размер N на страната:

- $N = n$;
- $N = 2\sqrt{n}$;
- $N = \frac{n}{2}$. “

Горните фиксации нямат нищо общо с математическия модел , а с негови алгоритмични реализации. Освен това самият модел е представен непълно. Необходим е коментар за различни възможности за свързване на целевата функция с допустимите точки.

Препоръчително е споменатите неточности да бъдат отстранени в евентуалните интернет варианти на дисертацията.

Получените резултати са публикувани в 2 статии в списания с импакт фактор, и 8 в сборници с абстракти от международни конференции. Всички публикации са на английски език.

Освен самостоятелната подготовка, квалификацията на докторанта е повишена и от участие в интензивни курсове:

- Едномесечна специализация по DAAD, 15.10-15.11.2014г. в Института по математика и информатика, БАН в София, при проф. д-р Никола Янев.
- DAAD Intensive course - New Aspects of the Time Frequency Analysis involving Frechet Frames, 28.09.16-03.10.16, Нови сад, Сърбия;
- Двумесечна специализация по Еразъм+, 01.03-31.04.2017г. в Университета в Рен 1, Рен, Франция, при проф. д-р Румен Андонов (IRISA);

Качества на автореферата. Авторефератът отразява точно и коректно съдържанието на дисертационния труд.

Заключение

Оценявам дисертационен труд положително и считам, че напълно отговаря на изискванията за присъждане на образователната и научна степен „доктор”, съгласно закона за развитие на академичния състав в Република България и на Правилника за развитие на Академичния състав на ЮЗУ „Неофит Рилски”.

Поради това предлагам на членовете на научното жури да дадат своя положителен вот за присъждане на образователната и научната степен “доктор” на асистент Методи Георгиев Трайков.

23.07.2015

Н.Янев