

Рецензия

от проф. д-р Радослав Йошинов

върху дисертационен труд за придобиване на научна степен „**доктор**“ с автор магистър Георги Симеонов Симеонов на тема „**МОДЕЛИРАНЕ И АНАЛИЗ НА ГОЛЕМИ ДАННИ ЗА ЕПИДЕМИЯТА ОТ COVID-19**“
с Научен ръководител: проф. д-р Петър Станчев Научен консултант: проф. д-р Огнян Кунчев
по професионално направление 4.6 „Информатика и компютърни науки“.

1. По процедурата

Със заповед на Директора на ИМИ-БАН бях назначен в научното жури на дисертационен труд за придобиване на научна степен „доктор“ с автор маг. Георги Симеонов Симеонов, а на първото заседание на научното жури бях избран за рецензент.

По въпросната процедура официално ми бе изпратена справка за минималните количествени показатели за придобиването на научната степен „Доктор“, съгласно Правилника на ИМИ-БАН за прилагане на ЗРАСРБ и Правилника към него, както и съответните показатели за маг. Георги Симеонов Симеонов, от които е видно, че той покрива минималните изисквания.

2. Кратки автобиографични данни

Средното си образование маг. Георги Симеонов Симеонов завършва с отличие във Високотехнологично училище "Електронни системи" под егидата на Технически университет София, специалност "Софтуерно развитие и информационни технологии" - II степен на професионална квалификация.

Висшето си образование започва в Технически университет София 2003 - 2005, Факултет по компютърни системи и контрол. Завършен първи курс с много добър успех. Предпочетен от докторанта е Факултет по математика и информатика, където се прехвърля за разширено обучение по софтуер, алгоритми и програмни езици от високо ниво.

Завършва Бакалавър 2005 - 2009 в Софийски Университет, Факултет по математика и информатика, специалност Информатика - завършил бакалавърска диплома с общ успеваемост много добър.

Магистърската си степен по "Сигурност на информацията в компютърните системи и мрежи" с отличен успех получава от Софийски Университет (2010 – 2011), Факултет по математика и информатика.

Взема участие в проекти на ИМИ-БАН, като специалист по внедряване на DMS, CMS, DAM системи и платформи за дългосрочно съхранение на данни, ETL (Екстракт, Трансформиране на натоварване), Метаданни, Извличане на данни, Извличане на информация, Търсене, Индексиране, Събиране на метаданни - OAI-PMH, уеб рамки и езици: C++, Python, PHP, Java / JSP.

Участие в европейски проекти за изграждане на научна мрежова инфраструктура с отворен достъп на електронни хранилища и архиви и по проекти на Инициативата за отворен достъп: OpenAIRE - Инфраструктурни изследвания с отворен достъп за Европа www.openaire.eu; · EuDML - Европейската библиотека за цифрова математика project.eudml.org; · DRIVER - Инфраструктура за цифрово хранилище
Член е на Съюза на математиците в България – SMB/UBM (www.math.bas.bg/smb), Алианс за научноизследователски данни – ПДП (www.rd-alliance.org)

3. Актуалност на дисертационния труд

Инфекциозните заболявания предизвикват милиони смъртни случаи всяка година и са основната причина за човешката смъртност. Математическото моделиране като приложен инструмент в сферата на общественото здраве се появява още в трудовете на Бернули около 1766 г

Необходимостта от централизиран контрол на епидемичните огнища може да се счита за основен двигател за развитието на съвременната епидемиология, по-специално за развитието на новата област използваща математическото моделиране като приложен инструмент, наречена Цифрова и Изчислителна Епидемиология (Digital and Computational Epidemiology), която да се използва и прилага широко в политиките от областта на общественото здраве.

През последното десетилетие благодарение на положените усилия в световен мащаб се ускорява развитието на глобална мрежа за наблюдение, предоставяща епидемиологични Големи данни, за борба с пандемии на възникващи и повторно възникващи инфекциозни заболявания.

Визуализация чрез компютърна графика е съвкупност от методи и технически средства, използвани за създаване и обработка на графични изображения чрез компютър позволяваща творчески да се комбинират и

пълноценно да се прилагат модерните средства за графична обработка при решаването на практически задачи от различни приложни области.

Настоящата дисертация представлява принос към нововъзникващата област на Цифровата и Изчислителната Епидемиология, като може да бъде определена като моделиране на "Големи данни" на заразни болести, по-специално на COVID-19, чрез създаване и изучаване на нови математически модели за анализ на епидемични огнища, включително чрез компютърно моделиране и симулации в рамките на концепцията за Големи данни.

4. Степен на познаване състоянието на проблема и обща характеристика на труда

Обект на изследването е приложението на математически модели за анализ на разпространението на болестта Covid-19, като се търси общото при създаването на сложни обекти, чрез различни подходи, като експлицитно описание, фотограметрия или симулации.

Дисертационният труд съдържа 125 страници, от които 110 страници. Основният текст е систематизиран в четири глави и библиография към всяка отделна глава от общо 85 заглавия. Извън него са обособени допълнително 3 глави: Увод, Заключение, Приноси на дисертационното изследване, Справочен апарат, съдържащ: използвана и цитирана литература и Приложения. Работата съдържа 49 фигури.

Определени са насоки за бъдещи изследвания и развитие. Представен е списък с научни публикации по темата и забелязани цитирания.

В библиографията на дисертационния труд са цитирани 85 източника, включващи книги, научни статии и интернет публикации. На база обзорния анализ докторантът формулира целта и задачите на дисертационния труд.

5. Съответствие на предложената методика на изследване и поставените цел и задачи на дисертационния труд

Основната цел на настоящата дисертация е създаването на модели за анализ на разпространението на болестта Covid-19, както и да представи софтуерна разработка, като уеб-базирани онлайн инструменти за анализ в реално време с интерактивни визуализации. Дисертацията се състои от седем глави от които четири в основен текст. От тях: Първата глава е въведение в тематиката на математическите модели с отделения за епидемиологията. Втората глава описва краткосрочен модел. Третата глава е посветена на дългосрочен модел с ваксиниране и сезонен ефект. Четвъртата глава описва софтуерните платформи.

Задачите, които са решени за постигане на горната цел са:

1. Изграждане на модел за краткосрочна прогноза на епидемиологичните криви, модел TVBG-SEIR, базиран на сплайни с четири възли, представляващи коефициентите на "трансмисия (предаване)" и "премахване" на модела SEIR.

По тази задача е анализирана краткосрочната глобална еволюция на контролираните епидемии чрез въвеждане на мерки за ограничаване/отваряне от здравните власти. Моделът отчита включване на различни сценарии за прогнозиране с ограничаване на контактите – lockdown, вариращ във времето и позволяващ анализ не само на първичната епидемична вълна, но и на възникващата вторична вълна и следващи вълни.

2. Разработка на уеб инструмент за изграждане на сценарии за COVID19 (накратко SBT-COVID-19), използващ софтуерната архитектура на Jupyter Notebooks с Vokeh.

Основен принцип на работа е два програмни процеса, единият базиран на клиентска инфраструктура и другият на ядрото (Kernel). Инструментът може да се използва като софтуер за подпомагане на взимането на решения (Decision Support system - DSS) от (здравните) власти за изследване на различни сценарии и политики, като се контролира/променя мащабът на мерките за ограничения/задържане (домашна и социална изолация/карантина, ограничаване на пътувания и други), и се оценява тяхната ефективност.

Инструментът SBT-COVID-19 позволява да се прецени колко дълго трябва да се поддържат мерките за ограничаване.

3. Изграждане на модел за дългосрочна прогноза, базиран на ATVBGSEIR модела относно оценката на продължителността на епидемията от COVID-19 в дадена държава, отчитаща различни сценарии.

Включено е използването на ваксини в модела, които се извършват по определен план за ваксинация, предоставен на месечна база.

Отчетен в модела е и сезонния ефект. Алгоритъмът взема предвид основното ограничение на здравната система, което е броят на леглата за интензивно лечение (ICU), предназначени за пациенти с COVID-19.

4. Разработен е уеб базиран инструмент за сценарии за мерки на ограничаване/отваряне, базиран на софтуерните платформи Jupyter и Vokeh, използващ алгоритъма, имплементиращ методологията на модела ATVBG-SEIR.

Визуализирана е ефикасността на инструмента, направена върху данни от COVID-19 от Австрия, България, Германия, Италия, Великобритания и САЩ.

Разработката има интересни обобщения и насоки за бъдещо развитие.

За постигането на поставената цел и на свързаните с нея задачи, в изследването е приложена комплексна методология, която обхваща следните методи и подходи: приближени и евристични алгоритми, симулации, квантово-химични подходи, HPC изчисления; целочислена оптимизация; векторна и тензорна геометрия; обикновени и частни диференциални уравнения; и др.

Избраните методи съответстват на основната цел и задачи, поставени за решаване от докторанта.

6. Характеристика на естеството и оценка на достоверността на материала, върху който се градят приносите на дисертационния труд

Методите и моделите, които се създават и използват, отговарят на целевата задача. За някои са получени емпирични числени резултати, а за други са представени концепции за решения и са дадени насоки за бъдещи изследвания.

Не съм забелязал грешки нито в конкретните, нито в концептуалните модели. Намирам също, че предложените стратегии са добре обосновани.

7. Приноси на дисертационния труд

Научни приноси. Изграждане на нови модели от тип SEIR

Модели, които са с променливи във времето параметри, базирани на сплайни.

1. Първият модел TVBG-SEIR се основава на сплайн модели за параметрите на SEIR (на скоростта на предаване и премахване), генерира сценарии за краткосрочно прогнозиране с двумесечен времеви хоризонт; има набор от параметри (за отпускане и затягане на мерките за ограничаване на Covid-19), които могат да бъдат променяни, за да се получат различни сценарии.

2. Вторият модел ATVBG-SEIR се различава съществено от първия, генерира сценарии за дългосрочно прогнозиране и по този начин позволява генерирането на прогнози в хоризонта за около няколко години. Той разполага с набор от параметри, които са специфични за всяка страна: ескалация, продължителност на периодите на блокиране, процент интензивни отделения, лимит на интензивни отделения, два вида планове за ваксинация и няколко технически параметъра.

Приложни приноси. Разработени софтуерни решения:

Уеб базирани инструменти (софтуер като уеб услуга) базирани на Jupyter Notebook и Vokeh.

1. Първият инструмент се основава на модела TVBG-SEIR и реализира интерактивна средна стойност (визуализация на сценариите и кривите на

моделите SEIR) за генериране на краткосрочни прогнози/прогнози, чрез уеб контроли за интерактивното изменение на параметрите. Той се реализира посредством Jupyter Notebook и софтуерен пакет Vokeh за общ клиентски сървър.

2. Вторият разработен уеб-базиран инструмент използва модела ATVBG-SEIR и се занимава с генериране на дългосрочни сценарии/прогнози за прогнозиране. Тези сценарии включват персонализирани планове за ваксинация, както и подходящ модел на сезонния ефект. Има опция, позволяваща на потребителя да променя голямо разнообразие от модели (и параметри на модела) по интерактивен начин.

Рецензентът приема така описаните приноси, като препоръчва на авторът да се научи по-точно и систематизирано да излага своите постижения.

8. Степен на личното участие на дисертанта в приносите

За личното участие на докторанта съдя по публикационната дейност на докторанта отразена в публикуваните по дисертацията материали. Докторантката убедително представя постигнатите резултати, с много добра и задълбочена аргументация, както и използва професионално графично оформление на материалите.

Характерът на изследването предполага много добра и широка подготовка в областта на методи и средства за интегриране, обработка и моделиране на хетерогенни данни. Считаю, че докторантът се е справил успешно, като не поставям под съмнение личното му участие в разработването на дисертационния материал.

9. Преценка на публикациите по дисертационния труд

По темата на дисертационния труд са предоставени 5 публикации в реномирани международни издания. Изнесени са и 4 доклада по темата на научни сесии на ФМИ, както и на научни семинари на ИМИ.

Публикациите в реферирани издания са на английски език и в съавторство, което е приемливо, вземайки предвид мултидисциплинарния характер на изследванията.

Те са публикувани в признати международни издания. Едната е публикувана в доклади на БАН и има SGR. Втората статия е публикувана на международната конференция DIP2020. Тези публикации са индексирани в SCOPUS.

Рецензентът отчита че публикационната дейност по темата на дисертацията обхваща периода 2004-2021 година и е направена в реномирани международни научни издания.

Фактът, че всичките публикации са в съавторство, отразява мултидисциплинарността на разглежданата в тях проблематика.

Публикациите отразяват по-съществените резултати, постигнати в дисертационния труд. Докладвани са на достатъчно реномирани научни форуми, което приемам за апробация в научните среди.

10. Съответствие на автореферата с изискванията за изготвянето му и адекватност на отразяване на основните положения и приносите на дисертационния труд

Представеният проект за автореферат е в съответствие с правилника за изготвяне на авторефератите по дисертационните трудове, посочен в сайта на ИМИ-БАН. Отразява постигнатите резултати, както и приносите на автора. Графично е оформен много добре и включва необходимата информация, описваща в резюме дисертационния труд.

11. Мнения, препоръки и бележки

В дисертационния труд се разработва една много сложна, динамично развиваща се и перспективна област – създаването на модели за анализ на разпространението на болестта Covid-19, както и да представи софтуерна разработка, като уеб-базирани онлайн инструменти за анализ в реално време с интерактивни визуализации. Анализирани са методи и средства за визуализиране, интегриране, обработка и моделиране на хетерогенни данни, получени от изследователска дейност. Разработена е иновативна софтуерна среда, като е дадено и практическото ѝ приложение.

Това предполага достатъчно задълбочени знания, възможност за интерпретация и формулиране на стратегии за ефективно развитие на областта. Съдържателно и графично материалът е разработен много добре. Този материал представлява интерес за широк кръг читатели и ако след преработка се публикува ще има мултиапликативен ефект.

Препоръчам на магистър Георги Симеонов Симеонов да продължава активната си публикационна дейност в научни списания с импакт фактор.

Препоръчвам по-точна вербализация на неговите постижения - да се научи по-систематизирано да излага своите приноси.

Отправлям забележка, че не е направена адекватна номерация на библиографията, което затруднява реферирането към нея. Също така не е подредена по азбучен ред.

Дисертационния материал само би подобрил своето качество, ако се допълни с речник на използваните термини и съкращения.

Някои не съществени забележки съм отразил върху копие, което ми бе предоставено.

12. ЗАКЛЮЧЕНИЕ

Съдържанието и приносите на дисертационния труд на магистър Георги Симеонов Симеонов напълно покрива изискванията на Закона за развитие на академичния състав на Република България, на Правилника за неговото приложение и на Правилника за условията и реда за придобиване на научни степени в ИМИ-БАН. Извършена е значителна по обем и съдържание изследователска работа. Има достатъчен брой научно-приложни и приложни приноси. Представени са достатъчен брой публикации по дисертацията публикувани на престижни научни форуми. Безспорно е личното участие на автора в разработката и получените приноси. Това ми дава основание убедено да препоръчам на Уважаемото Научно жури да присъди на **магистър Георги Симеонов Симеонов** научната степен „**доктор**“ в професионално направление 4.6 „Информатика и компютърни науки“.

РЕЦЕНЗЕНТ:

/проф. д-р Р. Йошинов/

София, 10.06.2022 г.