

РЕЦЕНЗИЯ

от проф. д.н. инж. Иван Ганчев Гарванов

Университет по библиотекознание и информационни технологии

на дисертационен труд за присъждане на образователната и научна степен „доктор“
по: област на висше образование 4. Природни науки, математика и информатика
професионално направление 4.6 Информатика и компютърни науки
докторска програма „Информатика“

Автор: Николай Георгиев Ханджийски

Тема: Итеративен парсиращ алгоритъм с приложение в профилирането на парсери

Научен ръководител: проф. д-р Елена Петрова Сомова - Пловдивски университет
„Паисий Хилендарски“

1. Общо описание на представените материали

Със заповед № РД-21-654 от 21.03.2024 г. на Ректора на Пловдивския университет „Паисий Хилендарски“ съм определен за член на научното жури за осигуряване на процедура за защита на дисертационен труд на тема „Итеративен парсиращ алгоритъм с приложение в профилирането на парсери“ за придобиване на образователната и научна степен „доктор“ в област на висше образование 4. „Природни науки, математика и информатика“, професионално направление 4.6 „Информатика и компютърни науки“, докторска програма „Информатика“. Автор на дисертационния труд е Николай Георгиев Ханджийски – докторант в редовна форма на обучение към катедра „Компютърна информатика“ с научен ръководител проф. д-р Елена Петрова Сомова от Пловдивски университет „Паисий Хилендарски“.

Представеният от Николай Георгиев Ханджийски комплект материали на електронен носител е в съответствие с Чл.36 (1) от Правилника за развитие на академичния състав на ПУ, включва следните документи:

- молба до Ректора на ПУ за разкриване на процедурата за защита на дисертационен труд;
- автобиография в европейски формат;
- протокол от катедрения съвет, свързан с докладване на готовността за откриване на процедурата и с предварително обсъждане на дисертационния труд;
- дисертационен труд на български език;

- автореферат на български и английски език;
- списък на научните публикации по темата на дисертацията;
- копия на научните публикации;
- списък на забелязани цитирания;
- декларация за оригиналност и достоверност на приложените документи;

Докторантът е приложил шест броя публикации, всички индексирани в Scopus и/или WoS., пет от тях притежават SJR, а три са с IF.

2. Кратки биографични данни за докторанта

Докторант Николай Ханджийски придобива бакалавърска степен по „Информатика“ и магистърска степен по „Софтуерни технологии“ в Пловдивски университет „Паисий Хилендарски“ през 2010 и 2012 г. По време на следването си специализира 5 месеца в Университет Йоханес Кеплер, Линц, Австрия през 2009 г. по програма Цепус и пет месеца в Университет Кобленц-Ландау, Кобленц, Германия през 2010 г. по програма Еразъм.

През целия си творчески път докторанта е свързан с информационните технологии като последователно във времето работи като: програмист, хоноруван асистент по SQL; научен асистент, управител на фирма и др. От януари 2018 г. до момента работи във фирма Софтуерни Системи – България като програмист отговарящ за софтуерната поддръжка.

3. Актуалност на тематиката и целесъобразност на поставените цели и задачи

Темата на дисертационният труд е изключително актуална и изследва въпроси свързани с проектирането на машина за парсиране на данни. Ползите от тази парсираща машина са, че ще обединява различни методи и алгоритми за анализ.

Цел на дисертационното изследване е да се изследват, предложат, проектират, разработят, експериментално приложат и апробират средства (машини, алгоритми, модели, езици и инструменти), които са подходящи за линейно транслиране на данни на базата на някои многозначни безконтекстни граматика.

За постигане на поставената цел са формулирани следните задачи:

Да се направи задълбочен обзор и анализ на теории, подходи, методи, алгоритми, модели, архитектури, машини и системи свързани с транслирането на данни;

Да се проектира архитектура на парсираща машина, която да обединява различни подходи и алгоритми за лексически и синтактичен анализи с някои добавени нови възможности;

Да се дефинират нов вид безконтекстни граматика съобразени с предложената архитектура на парсираща машина;

Да се проектира алгоритъм за парсиране, вграден в парсиращата машина, който да може да парсира данни на базата на новия вид граматика;

Да се проектира и реализира прототип на инструмент за измерване и сравняване на използваните ресурси от различни парсери;

Да се проведат експерименти с помощта на създадения инструмент.

4. Познаване на проблема

Докторанта познава много добре тематиката на изследвания проблем. Направил е задълбочен обзор и анализ на 191 литературни източника по проблема и е дефинирал коректно целта и задачите на неговото изследване. В последователното развитие на дисертационният труд – от увода, през четирите глави до заключението, личи доброто познаване на научния проблем – както в теоретичен, така и в практически план. Благодарение на уменията на докторанта и неговият научен ръководител са получени значими научни и научно приложни резултати които са апробирани в шест научни публикации.

5. Методика на изследването

Избраната методика на изследване е позволила на докторанта да постигне поставената от него цел както и да получи значими научни и научно приложни резултати. Приложени са адекватни методи за изследване, основаващи се най-вече върху синтез и анализ на теоретични постановки и резултати от реални изследвания. Избраният подход на изследване на формулираните цел и задачи на дисертационният труд е основан на анализиране на алгоритмите за извличане на необходима информация от някои данни и по-специално от текстови данни.

6. Характеристика и оценка на дисертационния труд

Предоставеният ми за рецензиране дисертационен труд се състои от 174 страници, оформени в увод, четири глави, изводи, приноси и приложения и използвана литература.

В първа глава е направен задълбочен обзор и анализ на крайни автомати, стекови автомати, машини на Тюринг, алгоритми на Марков и др. Направеният обзор включва известни алгоритми за разпознаване и парсиране работещи на базата на безконтекстни граматика.

В глава втора са дефинирани различни видове синтактични дървета и командите за строеж, на базата на които дърветата могат да се създадат.

В трета глава е описан алгоритъма за Тунелно парсиране и е разгледан пример с дефинираните обекти, които се използват от алгоритъма, като е демонстрирано и примерно изпълнение на алгоритъма за избран низ от входни данни.

В четвърта глава се дискутира специално създаден за целта на дисертацията инструмент, наречен профилатор, с помощта на който могат да се генерират голям брой контекстно свободни граматика и входни данни и да се извършват експерименти с парсиращи машини, които са генерирани от различни парсер генератори на базата на тези граматика. Главата завършва с интерпретацията на резултатите от четири различни експеримента, които са проведени с помощта на профилатора.

Заключението обобщава получените научни резултати, прави изводи от изследването и дава насоки за бъдещи изследвания.

Резултатите са онагледени в 56 фигури и 8 таблици. Използвани са 191 литературни източници.

7. Приноси и значимост на разработката за науката и практиката

В дисертационния труд са посочени приноси, които бих класифицирал както с научен, така и с научно-приложен характер. В изследванията на Николай Ханджийски се потвърждават някои известни факти, обогатява се съществуваща научна област с нови знания и се предлага някои от новополучените научни резултати да се приложат в практиката.

От приложена декларация за оригиналност на представените резултати, като и от представените публикации по дисертационния труд може да се съди, че описаните резултати са оригинални и лично дело на докторанта. Протоколът от системата за антиплагиатство StrikePlagiarism показва по категоричен начин, че дисертацията е оригинална и автентична.

8. Преценка на публикациите по дисертационния труд

Докторанта Николай Ханджийски е апробирала части от дисертационния си труд в шест научни публикации на английски език. Шестте публикации са индексирани в световноизвестните бази от данни: 4 (четири) в Web of Science и 5 (пет) в Scopus. Пет от публикациите са в издания с SJR. Три от тях са с IF, но без квартали.

Съгласно минималните национални изисквания за получаване на ОНС „Доктор” по професионално направление 4.6. “Информатика и компютърни науки”, определени в чл. 2б,

ал. 2 и 3 на ЗРАСРБ и съответно по чл. 24, ал. 1 на Правилника за прилагане на ЗРАСРБ се изискват наличие на поне 30 точки по показателите от Група Г. Представените публикации по дисертационния труд формират обща сума от точките за показателите от Група Г равна на 168 точки, което превишава многократно изискуемия минимум от 30 точки. Всички публикации по конкурса са в съавторство с научния ръководител като във всички публикации той е първи автор.

Забелязано е едно цитиране на една от публикациите на докторанта.

9. Лично участие на докторанта

Личното участие на докторанта в проведеното дисертационно изследване е видимо както в дисертационният труд така и в приложените публикации. Прави много добро впечатление умението на докторанта да анализира резултати от други автори, което е видимо от теоретичните анализи и дискусии в първа глава на дисертационният труд. Високо оценявам и неговите възможности за работа с алгоритми и програми, както и изследователските му умения за анализ на получените резултати. От местата на публикуване на научни резултати, напълно убедено смятам, че резултатите са оригинални и притежават значим научен принос.

10. Автореферат

Представеният автореферат на български и английски език от докторанта Николай Ханджийски кореспондира с дисертационният труд и представя в синтезиран вид изследваните проблеми. Той е в достатъчен обем и отразява достоверно основните и най-съществените идеи, изводите и приносните моменти на дисертационния труд като съответства на изискванията на ЗРАСРБ и ППЗРАСРБ.

11. Критични забележки и препоръки

Част от представените резултати в дисертационния труд не са намерили отражение в публикациите по дисертацията и препоръката ми е те да бъдат публикувани.

12. Лични впечатления

Не познавам докторанта и нямам общи публикации с него, но от предоставените ми документи мога да изразя моето положително впечатление от получените научни резултати и от неговото кариерно развитие.

13. Препоръки за бъдещо използване на дисертационните приноси и резултати

Препоръчвам на докторанта в бъдеще да работи със студенти и докторанти за да предаде неговите знания и умения на младите и талантливи деца на България.

ЗАКЛЮЧЕНИЕ

Дисертационният труд *съдържа научни, научно-приложни и приложни резултати, които представляват оригинален принос в науката* и отговарят на всички изисквания на Закона за развитие на академичния състав в Република България (ЗРАСРБ), Правилника за прилагане на ЗРАСРБ и съответния Правилник на ПУ „Паисий Хилендарски“.

Дисертационният труд показва, че докторантът Николай Георгиев Ханджийски **притежава** задълбочени теоретични знания и професионални умения по научна специалност 4.6 „Информатика и компютърни науки“ като **демонстрира** качества и умения за самостоятелно провеждане на научно изследване.

Поради гореизложеното, убедено давам своята *положителна оценка* за проведеното изследване, представено от рецензираните по-горе дисертационен труд, автореферат, постигнати резултати и приноси, и *предлагам на почитаемото научно жури да присъди образователната и научна степен „доктор“* на Николай Георгиев Ханджийски в област на висше образование: 4. „Природни науки, математика и информатика“, професионално направление 4.6 „Информатика и компютърни науки“ докторска програма „Информатика“.

06.05.2024 г.

Рецензент:

Проф. д.н. Иван Гарванов