

СТАНОВИЩЕ

по дисертационен труд за присъждане

на образователна и научна степен „доктор“

Тема: „Оптимизация и паралелизация на алгоритми, свързани с теория на кодирането“

Автор: Мария Руменова Пашинска-Гаджева,

Секция „Математически Основи на Информатиката“ (МОИ),

Институт по Математика и Информатика (ИМИ), Българска Академия на
Науките (БАН)

Научен ръководител: проф. дмн Илия Георгиев Буюклиев

Научна област: 4. Природни науки, математика и информатика

Професионално направление: 4.6. Информатика и компютърни науки

Докторска програма: Информатика

Изготвил становището: доц. д-р Юри Любчов Борисов – ИМИ - БАН

Със заповед № 456/03.12.2024 г. на Директора на Института по
математика и информатика (ИМИ) при БАН съм назначен за председател на

Научното жури по процедурата и на първото заседание на това жури съм избран за изготвянето на настоящото становище. Предоставени са ми всички материали в съответствие с изискванията на нормативните документи, които са редовни и съответстват на ЗРАСРБ. Нямам информация за нарушения по процедурата, не ми е известно в дисертацията да има плагиатство.

1. Данни за дисертанта

Дисертантът Мария Руменова Пашинска-Гаджева е получила бакалавърска степен от ВТУ „Св.Св. Кирил и Методий“ в образователната програма „Компютърни науки“ през 2019 г. и магистърска степен от ВТУ „Св.Св. Кирил и Методий“ в образователната програма „Математически структури в информационната сигурност“ през 2020 г.. В приложената към документацията автобиография, за съжаление, няма данни за участието ѝ в научни проекти. В списъка с участия в научни форуми са вписани 9 доклада, от които шест на международни конференции, два на националните семинари по теория на кодиране „Акад. Стефан Додунеков“ и един на семинара по високопроизводителни пресмятания.

2. Данни за докторантурата

Докторантът е зачислен в редовна форма на обучение от 01.01.2021 г. с период на обучение 3 г. и отчислен с право на защита с решение на НС на ИМИ (Протокол № 14/21.12.2023г.). Предварителното обсъждане на дисертацията, на което присъствах, се състоя на 23.11.2024 г. на разширено заседание на

секция МОИ в Заседателната зала на х-л „Севастократор“, Арбанаси. Със заповед на директора на ИМИ е определено научно жури и дата на защитата 25.02.2025г.. Считам, че процедурата е редовна и няма нарушения.

3. Данни за дисертацията и автореферата

Представения за рецензия дисертационен труд е написан на български език и се състои от: увод (8 стр.), основен текст (75 стр.) от 5 глави, разделени на секции. В дисертацията са включени списъци: на научно-приложните приноси, на публикациите (шест на брой) и на изнесените доклади (девет на брой). Макар че по ЗРАСРБ не се изискват цитирания, към дисертацията е приложен и списък забелязаните 2 цитирания. Използваната литература се състои от 83 заглавия, а таблиците и фигурите са съответно десет и пет на брой. Трудът отговаря на изискванията на Закона за развитие на академичния състав в РБ (ЗРАСБ) и правилника за приложение на ЗРАСРБ (ППЗРАСРБ), както и на Правилника за условията и реда за придобиване на научни степени и за заемане на академични длъжности в БАН (ПБАН). Авторефератът (в обем от 30 стр.) отразява адекватно основните идеи и съществените крайни резултати, които са описани в дисертационния труд.

4. Общо описание на дисертацията

В глава 1 на дисертацията са представени някои основни понятия и дефиниции, свързани с теория на кодирането, както и методологиите за оптимизация и паралелизация на алгоритми. Дефинирани се основните характеристики на

линейните кодове, описани се техните тегловни характеристики и значимостта им. Представени са самодопълнителните кодове и техните свойства. В раздела, посветен на оптимизацията и паралелизацията на алгоритми, са описани основните техники за паралелизация на алгоритми, които включват както разделяне на задачите, така и разделяне на данните. Подробно е разгледана векторизацията като метод за паралелизация, използващ разширени регистри и съответни инструкции на централните процесори.

Глава 2 е посветена на разработката и оптимизацията на алгоритми за намиране на тегловния спектър на линейни кодове. Представените алгоритми за намиране на тегловния спектър прилагат метод за генериране на следваща кодова дума чрез линейна комбинация на редовете на пораждащата матрица, използвайки единствено събиране на вектори. Описани се различни реализации на функции за събиране на вектори над крайно поле и намиране на теглото на вектор над крайно поле, използвайки разширени регистри и специализирани инструкции на централните процесори за x86 архитектура. В раздела, посветен на експерименталните резултати, са представени данни, които демонстрират ефективността на разработените алгоритми. Също така е направен анализ на ефективността на различни компилатори при прилагането на векторизация в предложените алгоритми.

Глава 3 разглежда специфичните особености при векторизирането на алгоритми с помощта на разширените инструкции на централните процесори.

Подробно са разгледани някои подкатегории от AVX512 инструкции за x86 архитектура и тяхното приложение за оптимизация на векторни операции. Описани са NEON инструкциите за ARM архитектура и техните специфики. Представени се алгоритми за събиране на вектори, използващи беззнакови типове данни и разширени инструкции със сатурация. Анализирани са резултатите от проведените експерименти, които демонстрират ефективността на оптимизациите с AVX512 и NEON инструкции.

Глава 4 се фокусира върху самодопълнителните кодове, достигащи границата на Грей-Ранкин, и техните свойства, както и върху методите за тяхното конструиране. Описани са подкодове с различни размерности и техните свойства. Представени са различни семейства от кодове с две и три тегла и връзките им със самодопълнителните кодове. Разгледани са методи за конструиране на кодове с по-висока размерност чрез използване на кодове с по-ниска размерност. Представени са резултатите от изчисленията, свързани с класификацията на самодопълнителните кодове.

В заключителната пета глава е изложено детайлното описание на разработената библиотека за намиране на тегловни инварианти на линейни кодове. Разгледани са подробно основните функционалности на библиотеката LinCodeWeightInv. Представени са както интерфейсът на библиотеката, така и модулът за тестване и верификация. Описани са използваните софтуерни инструменти при разработването и документацията на библиотеката, както и

структурата на файловете, съдържащи нейния програмен код, и методите за компилиране и инсталиране.

5. Научно-приложни приноси

Основните научно-приложни приноси на дисертацията са следните:

- Разработване на алгоритъм за генериране на непропорционални кодови думи работещ върху съставни полета само чрез събиране на два вектора.
- Разработване на алгоритъм за събиране на вектори чрез SSE, AVX и AVX512 инструкции при използване на по-битово представяне на елементите на полета GF(2), GF(4) и полетата с характеристика 3.
- Разработване на алгоритъм за събиране на вектори чрез SSE, AVX и AVX512 инструкции при байтово представяне за прости и съставни полета с до 64 елемента, и намиране на теглото на вектор чрез едно извикване на инструкцията *popcnt*.
- Анализ на ефективността на работата на различни компилатори с SSE и AVX инструкции при векторизация.
- Изучаване на характеристиките на наборите от инструкции AVX512 и NEON и анализиране на тяхната ефективност.
- Анализ на ефективността на различните видове инструкции в архитектури от вида x86.
- Разработване на алгоритъм за събиране на вектори чрез SSE, AVX и AVX512 инструкции при представяне на елементите на полета с до 128 елемента чрез цели числа без знак.
- Дефиниране на фамилии от дву- и три-тегловни кодове достигащи границата на Грей-Ранкин, и описание на връзките между отделните фамилии от тези видове.

- Разработване на конструкция за построяване на кодове от дадена фамилия с размерност k чрез кодове от съответна фамилия с размерност $k+2$.
- Разработване на математически софтуер (библиотека LinCodeWeightInv) за намиране на теглови инварианти на линейни кодове над полета с до 64 елемента, като се използват SSE4.1, AVX2, AVX512 инструкции за x86 архитектури и NEON инструкции за ARM архитектури.
- Представяне на особености и основните акценти при създаването на математически софтуер с отворен код.

Признавам посочените в дисертацията научно-приложни приноси.

6. Публикации и участия в научни форуми

В списъка с публикации по дисертацията има 6 статии. Четири от тези статии са в съавторство с научния ръководител проф. дмн Илия Буюклиев, а една от тях има за съавтор проф. дмн Стефка Буюклиева. Считам, че участието на дисертанта във всички публикации е равностойно с това на нейните съавтори. Броят на статиите отговаря на изискванията на ППЗРАСРБ.

Една от статиите е публикувана в списанието Mathematics 2022 с IF: 2.2 и SJR: 0.592 Q1, а друга е публикувана в серията Lecture Notes in Computer Science, Springer, vol. 13952, SJR: 0.606 (2023). Други три са включени в трудовете на международни научни конференции съответно:

- две в "Innovative STEM Education" IMI-BAS, 2021-2022;

- една в ICAI, Varna, Bulgaria 2022.

Шестата статия е подадена за рецензиране след втора ревизия в списание ACM Transactions on Mathematical Software.

7. Заключение

По мое мнение, представеният дисертационен труд „Оптимизация и паралелизация на алгоритми, свързани с теория на кодирането“ с автор Мария Руменова Пашинска-Гаджева съдържа научни и научно-приложни резултати, които представляват оригинален принос в софтуерното инженерство и информатиката. Дисертантът очевидно има задълбочени теоретични и практически знания в областта на оптимизирането на алгоритмите използвани в теория на кодирането, както и способност за самостоятелна научна и научно-приложна работа. С това становище считам, че тя отговаря напълно на изискванията установени от ЗРАСБ, както и на правилниците на БАН и на ИМИ, и убедено предлагам на уважаемото научно жури да гласува Мария Руменова Пашинска-Гаджева **ДА ПРИДОБИЕ** образователната и научна степен „доктор“ от в професионално направление 4.6. „Информатика и компютърни науки“.

22.01.2025 г.

Изготвил становището:

София

/доц. д-р Юри Борисов/