R E V I E W

on a Dissertation

for obtaining the educational and scientific degree "Doctor"

Research area: 4. Natural Sciences, Mathematics and Informatics,

Professional field: 4.6. Informatics and Computer Science

Author: Miroslav Tsvetkov Markov

Title: EFFICIENT ALGORITHMS WITH APPLICATION TO PUBLIC KEY CRYPTOGRAPHY AND CODING THEORY

Referee: Prof. Iliya Bouyukliev
Institute of Mathematics and Informatics,
Bulgarian Academy of Sciences

I am a member of the scientific panel for this procedure according to order No. 208/18.07.2024 of the Director of Institute of Mathematics and Informatics, Bulgarian Academy of Sciences. As a member of the scientific panel, I have received all the administrative and scientific documents required by the Act on the Development of the Academic Staff in the Republic of Bulgaria (ADASRB), the Rules for its implementation and the Rules on the terms and conditions for awarding of academic degrees and occupying of academic positions at IMI and BAS.

1. Personal data.

Miroslav Markov graduated with a master's degree in electronics and automation from Technical University, Sofia, in 1996, and a bachelor's degree in mathematics and informatics from Sofia University in 2006. From 1992 to 2008, he led teams to create software for the needs of the Ministry of Interior. After that, he continued as the chief expert in the same field at the State Agency for National Security. Since August 2022, he has been a mathematician in Section “Mathematical Foundations of Informatics” of the Institute of Mathematics and Informatics. Since 2019, he has been enrolled in full-time doctoral studies at IMI. Since then,

he has been actively involved in the scientific forums organized by the section and the institute. The resume he has submitted shows many additional courses and specializations related to computer systems and architectures, operating systems, and programming languages.

2. Relevance of the topic.

The dissertation of Miroslav Markov is devoted to research in the field of cryptography and algebraic coding theory and is related to two specific problems. The first one to determine the cardinality of the set of points on each elliptic curve of two special families, and the second one is to compute the weight distribution of the Reed-Muller code $\mathcal{R}(4,9)$.

The importance and relevance of the first problem is determined by the fact that elliptic curves are increasingly used in many areas of mathematics and especially cryptography. As an example, we can point out the elliptic curve-based digital signature algorithm ECDSA. A key aspect in designing a cryptosystem based on an elliptic curve is finding the number of points on the chosen curve, as it plays a central role in determining the security, key size, and performance of the cryptosystem.

The study of weight invariants such as covering radius and weight distribution of Reed-Muller codes is directly related to the study of cryptographic properties of Boolean functions and vectorial Boolean functions, and the vectorial Boolean functions are the main nonlinear components of any block cipher. I can define this problem as particularly important for the work, because it, in addition to being interesting and current, is at the limit of what is possible to solve with the available theoretical and computational resources.

3. General characterization of the dissertation.

The dissertation consists of 74 pages and is structured in three chapters, two appendices and references of 59 titles. The introduction provides motivation and historical notes for the problems under consideration. The first two chapters deal with elliptic curves. The third chapter is devoted to the calculation of the weight distribution of the Reed-Muller code $\mathcal{R}(4,9)$. Because the preconceptions and statements of the two problems are different, the author begins each chapter with an introduction in which he presents the necessary known facts.

Chapter 1 is devoted to the elliptic curves defined by $y^2 = x^3 + a\ (mod\ p), a \neq 0$. The first part presents an introduction to the problem, listing some of the used methods, as well as information on their complexity. In the second part, the necessary definitions and preliminary statements are given. Then the developed method and the used algorithms are presented. The chapter concludes with examples of using the algorithms on specific elliptic curves and evaluating the complexity. The results of this research is published in the paper „An Efficient Approach to Point-Counting on Elliptic Curves from a Prominent Family over the Prime Field $\mathbb{F}_p$" in the scientific journal *Mathematics*, and is presented as a talk and then published with the title „An Approach for Computing the Number of Points on Elliptic Curve $y^2 = x^3 + a\ (mod\ p)$ via Explicit Formula for That Number Modulo p" in: *Ninth International Workshop on Signal Design and its Applications in Communications* (IWSDA) (2019).

In Chapter 2 the authors study the family $\mathcal{D}_p$ of elliptic curves of the form:

$$\mathcal{D}_p = \{D_a : y^2 = x^3 + ax\ (mod\ p), a \neq 0\}.$$

The aim is to obtain a general approach to compute the order (the number of the points) of a curve from this family depending on *a* and *p*. The computational aspects of counting the points of curves of this family are presented in a third part. After calculating the complexity, the author concludes that the proposed approach significantly improves the best solution known so far. The results are published in „Point-Counting on Elliptic Curves Belonging to One Prominent Family", *Algebraic and Combinatorial Coding Theory* (ACCT) (2020).

Chapter 3 is devoted to computing the weight distribution of the Reed-Muller code $\mathcal{R}(4,9)$. There are several approaches for solving the weight distribution problem for an arbitrary linear code, related to exhaustive search, fast transformations, etc. But in this particular case the parameters of the code, especially dimension 256, make the use of general algorithms unthinkable. It is therefore crucial to exploit the structure and many of the properties of Reed-Muller codes based on their relation to Boolean functions. Common approaches in solving problems with high computational complexity is the use of theoretical solutions or reduction of calculations with theoretical reasoning.

The number of some of the weights is obtained using Gleason's theorem about the weight distribution of the binary doubly even self-dual codes, to which class the studied Reed-Muller code belongs. However, this turns out to be insufficient to finally solve the problem. In the next step, an approach based on Sarwate's method is used.

Computational work is divided into two main phases: a precomputing and actual computing. The pre-computing involves the following three tasks:

- Constitute and store the orbits of a partition of the cosets of the form $e + g + \mathcal{R}(2,7)$, where e and g are homogeneous polynomials on 7 binary variables of algebraic degree 4 and 3, respectively.
- Compute the weight enumerators of the cosets $e + g + \mathcal{R}(2,7)$ when g varies over a set of representatives of the orbits (the stabilizer St(e) of GA(7) partitions the cosets into disjoint orbits).
- Merge the orbits with identical weight enumerators to obtain the coarse partition $\Delta'(e)$, and make data arrangement permitting for given homogeneous polynomial f of degree 3 to look up the identifier of a block in $\Delta'(e)$, containing f (respectively, to have direct access to the common weight enumerator).

Very important for the obtained result are the following facts:

- A Reed-Muller code of higher order can be considered as obtained from a Reed-Muller code of lower order by a recurrence relation or $(u, u + v)$ construction.
- The considered linear and affine equivalences of cosets of a Reed-Muller code of order r, which are contained in a Reed-Muller codes of order $r + 1$ and $r + 2$.
- The weight enumerators of two affine equivalent cosets of a Reed-Muller code are identical.
- The use of the known relationship between the weight distributions of the Reed-Muller code of order r and the Reed-Muller codes of lower orders.

In the dissertation, in addition to the detailed description of the theoretical approach, a pseudo code of the used algorithms is also presented. Computational challenges are solved with the use of non-trivial ideas, serious algorithmic skills, capacity to use efficient software packages and multiprocessor computing systems. In one of the parts at the end of the chapter, there is a detailed assessment of the computational resources required and the available technique used. This is very indicative for problems of this type, because optimization of the algorithms is done until the required resource is no more than the available (with a reasonable computation time).

The dissertation is well balanced in content and follows the papers published during the doctoral program. The text is well laid out and shows a thorough command of mathematical terminology and skillful handling of computing resources. Each chapter begins with an introduction and ends with a conclusion. This facilitates the reader and enables him to follow the research without consulting other sources.

4. Contributions and importance of the results obtained

I accept and approve the submitted contributions, indicated by the author. I would highlight the following contributions:

- According to the elliptic curves: New effective methods for calculating the order of an elliptic curve are developed and explicit formulas for this order are derived for curves of the following two families:

$$\mathcal{E}_p = \{E_a: y^2 = x^3 + a \ (mod\ p), a \neq 0\},$$

$$\mathcal{D}_p = \{D_a: y^2 = x^3 + ax \ (mod\ p), a \neq 0\}.$$

  An efficient method has been developed to simultaneously calculate the six possible orders of the curves from the family $\mathcal{E}_p$, for p ≡ 1 (mod 6), as well as the four possible orders for the curves from the family $\mathcal{D}_p$, for $p \equiv 1\ (mod\ 4)$. The complexity of this method is $\tilde{O}(log^2 p)$.

- According to the computing the weight distribution of the Reed-Muller code $\mathcal{R}(4,9)$:

– An efficient algorithm for computing the weight distribution of the Reed-Muller binary code $\mathcal{R}(4,9)$ is developed by combining Sarwate's approach with results on the classification of Boolean functions according to their affine equivalence published in 2008 and 2023.

– The weight distribution of the Reed-Muller binary code RM(4, 9), which has parameters [512,256,32] (length 512, dimension 256, and minimum distance 32) is computed.

5. Publications and citations

The PhD Thesis of Miroslav Markov is based on his research on elliptic curves and Reed-Muller codes, and is described in four publications – one in a scientific journal, and three in proceedings of international scientific conferences. All these papers are written in English. The results on the elliptic curves from the family $\mathcal{E}_p$ are published in the scientific journal *Mathematics*, which has JCR-IF from Web of Science and belongs to quartile Q1. It brings him 75 points in the table of scientometric indicators. All four publications are co-authored with his scientific supervisor Yuri Borisov. I believe that the contribution of Miroslav Markov in the joint publications is essential.

The results of this research have been presented at seven conferences in Bulgaria and abroad:

- Ninth International Workshop on Signal Design and its Applications in Communications (IWSDA), Dongguan, China, 20–24 October 2019;
- The international workshop on Algebraic and Combinatorial Coding Theory ACCT‘2020, online;
- The conference *High-Performance Computing for Mathematics and Applications*, held in Sofia in June 2023;
- The Thirteenth International Workshop on Coding and Cryptography (WCC), Perugia, Italy, 17-21 June 2024;
- Annual national seminars in coding theory - 2019, 2022 and 2023.

Miroslav Markov has provided a list of five citations, two of which are for his work on elliptic curves and three for the Reed-Muller code.

6. The author’s summary

The abstract correctly reflects the content of the thesis. I would say that its layout is a bit different from the standard one. At the beginning, a general description of the thesis is presented, followed by goals and objectives, research methodology, approval of the results and a list of publications on the dissertation. After that, the presentation of the individual chapters begins, respectively in 12, 6 and 18 pages. The abstract ends with an author summary, acknowledgments and a list of references. The author has tried to present his thesis quite fully in it, but I think it would be better if the first and third chapters were described more briefly, with an emphasis on the author's contributions.

7. Personal impressions

I have known Miroslav Markov since 2018 when we discussed his PhD enrollment at the Coding Theory Annual Workshop. Since then, he has participated in all editions of the workshop. There we have discussed various topics related to algorithms, software optimization, application of high-performance computing and use of supercomputer. My impressions are that Miroslav Markov is very professionally prepared both in the mathematical theories he needs and in the development of algorithms, programming and the use of high-performance computing technologies.

8. Conclusion

The presented dissertation satisfies all the criteria and indicators of the law and the regulations in Bulgaria. After I familiarized myself with the presented dissertation, the importance of the research and the scientific and applied contributions contained therein, I give an overall positive assessment to the applicant **Miroslav Tsvetkov Markov** to obtain the scientific degree "Doctor" in

Research area: 4. Natural Sciences, Mathematics and Informatics,

Professional field: 4.6 Informatics and Computer Science.

05.09.2024 Reviewer:

/Prof. Iliya Bouyukliev/