Р Е Ц Е Н З И Я

на дисертационен труд за получаване на научната и образователна степен „доктор“ в област на висшето образование 4. Природни науки, информатика и информатика, професионално направление: 4.6. Информатика и компютърни науки, научна специалност 01.01.12 „Информатика“

Автор на дисертационния труд: Георги Великов Иванов
Тема на дисертационния труд: „New Heuristic Methods for Generation of Bijective S-boxes with Good Cryptographic Properties“ („Нови евристични методи за генериране на биективни заместителни таблици с добри криптографски свойства“)
Научни консултанти: доц. дмн Емил Колев, доц. д-р Светла Никова
Рецензент: проф. д-р Красимир Неделчев Манев

Тази рецензия е написана и представена на основание на Заповед № 282 от 07.10.2016 г. на директора на ИМИ, БАН, както и на решение на назначеното със Заповедта Научно жури по процедурата (Протокол 1 от 14.10.2016 г.). и е изготвена във основа на ЗРАСРБ, Правилника за прилагане на ЗРАСРБ, Правилника за развитие на академичния състав на БАН и на ИМИ при БАН.

1. Общо описание на дисертацията другите материали

Дисертационният труд е с обем 129 страници (45 реда и 90 знака на ред) и се състои от 114 страници основен текст, 11 страници приложения и 4 страници цитирана в текста литература (общо 94 заглавия).
Списъкът с ключови за работата думи и резюмето в началото е обичайно при публикуването на научна статия, но ни се струва ненужно за дисертационен труд. Следват Съдържание, Списък на таблиците (16 таблици) и Списък на публикациите по дисертацията с 4 заглавия. Останалият текст е разделена на 6 глави, първата от които е Въведение, а последната Заключение. Приложението е оформено като седма глава. Като неотменна част от работата разглеждаме също и представените Автореферат, копия на публикациите по дисертацията и списъка с цитирания на публикациите.
Оформлението на работата е добро. Английският език, доколкото можем да преценим, също е добър, но основанията работата да бъде написана на английски не са ни понятни (можем да посочим поне три възражения).

2. Актуалност на разработения в дисертационния труд проблем

Тематиката на която е посветен дисертационният труд е много актуална. В условията на повсеместно използване на компютри и компютърни мрежи, в тях се съхраняват и обменят огромни количества данни, включително и такива с поверителен характер (от гледна точка на личната, фирмената или

националната сигурност). Защитата на тези данни от нежелан достъп трябва да бъде ежедневна и адекватна, предвид непрекъснатия интерес на заинтересовани лица и организации да получат достъп до защитените данни. Затова разработването на нови и по-надеждни техники за защита на данните е от особена важност. Една от съвременните техники за недеждно криптиране на данните е включването на биективни *заместващи* (предпочитам да използвам тази дума вместо авторовото *заместителни*) таблици (накратко БЗТ, в английската терминология – S-boxes) в класическите блокови и поточни шифри. БЗТ са непрост, дискретен математически обект, и намирането на нови таблици с полезни за криптографията свойства не е тривиално.

3. Познава ли дисертантът съвременното състояние на областта?

От обширното представяне във втора глава на свързаната със задачата теория – понятиен апарат и математически резултати, имащи отношение към поставената задача, както и от анализа на проблема и съществуващите възможности за решаването му в трета глава, можем да заключим че докторантът познава отлично научната област на дисертационния труд. Това ни заключение се потвърждава и от списъка на използваната литература, която наброява близо 100 заглавия. От тези стотина публикации, повече от 70 са излезли от печат след 1990, а 33 – след 2000 година. Затова смятаме, че използваната литература отразява съвременното състояние на областта.

4. Съдържателен анализ на дисертационния труд

Първа глава на дисертацията (Въведение), започва с кратка Мотивация и формулиране на целта на работата – разработване на ефективен евристичен инструментариум за генериране на големи множества от БЗТ с размери до 16x16, притежаващи криптографските свойства, необходими за вграждане в блокови и поточни шифри. За постигане на целта са формулирани 5 насоки на изследване: изясняване на ***взаимната зависимост*** на криптографските свойствана булевите функции и заместващите таблици; да се потърсят ***комбинации от съществуващи евристични техники***, които могат да дадат желаните резултати; да се построят ***нови евристични алгоритми*** за пораждане на големи множества добри заместващи таблици; да се изследват ***свойствата на получените алгоритми***, като сложност по време и памет и качество на получаваните резултати; да се използват получените резултати при ***разработване на нови блокови шифри***.

След това Въведението продължава с кратко изложение на получените резултати, което ми се струва излишно, тъй като почти буквално се повтаря в автореферата, където, според нас, му е мястото и, съкратено, в Заключението. В края на главата е представена структурата на работата и съдържанието на останалите глави.

Глава втора въвежда читателя в теорията на булевите функции (БФ), в частта ѝ която е важна за работата. При общ обем на съдържателната част от 94 страници (изключвайки Увода и Приложенията), обемът на въведението в тематиката от 55 страници ми се струва твърде голям. Не е трудно да се покаже, че не всички дефинирани понятия (97) и твърдения (73) се използват за целите на работата. Че образователната страна на докторантурата е успешна може да се докаже чрез изпита по специалността.

В **Глава трета** са представени основните способи за генериране на заместващи таблици: – псевдо-случайно генериране, генериране с алгебрични конструкции и евристично генериране. Сравнени са качествата на трите способа. Известно е, че с алгебрични конструкции (т.н. finite field inversion, FFI) са получени едни от най-добрите като характеристика БФ, но поради наличието на добра алгебрична структура, получените от тях ЗТ са податливи на алгебрични атаки. Освен това, с тази техника се получават единични екземпляри. Ето защо, за целите на работата е избрано евристичното генериране – техника при която се започва с едно множество от обекти и се пораждат итеративно (случайно или с помощта на подходящи евристично оправдани промени) близки до тях обекти – „наследници“, при които интересни характеристики се подобряват. Целта е, да се построят големи множества БЗТ с характеристики близки до тези получени с алгебрични конструкции. Представени са и основните евристични техники, използвани до момента за търсене на БЗТ с размери 8x8:

– *hill climbing*: представени са алгоритми за търсене на БФ и БЗТ с тази техника, при критерий за стъпка във „височина“ е използвана нелинейността на наследниците. Техниката е оценена като недостатъчно ефективна (нелинейност 98-100 и не голям брой таблици с размер 8x8);

– *simulated annealing*, вариант на *hill climbing*, при който се избягват локалните екстремуми. Представен е алгоритъм, който дава по-добри резултати от *hill climbing*, но все още не е достатъчно ефективен (нелинейност 102, не голям брой таблици);

– *генетични алгоритми*, при които за получаване на нова популация се правят кръстоски и мутации. Представен е алгоритъм, който от популация случайно генерирани БЗТ успява да построи таблици с нелинейност 104.

В главата са представени и т.н. имунни алгоритми – разновидност на генетичните. Тази техника, според докторанта, не е била използвана до момента за генериране на БЗТ.

Глава четвърта е основна за дисертацията. В нея са представени разработените от автора две фамилии евристични алгоритми за построяване на заместващи таблици.

Първата фамилия от 4 алгоритъма, наричана от автора *реверсивни генетични,* започват с популация от внимателно подбрани известни таблици (получени с FFI), с хубави стойности на някои от характеристиките, а целта е да се получат, от една страна, голямо количество таблици с може би по-лоши стойности на най-важните характеристики, но, от друга страна, с по-добра устойчивост на атаки.

Първият от алгоритмите (RGA1), според докторанта, се различава от класически генетичен алгоритъм по това че работи реверсивно (не можем да приемем това твърдение, тъй като не виждаме разликата). В този алгоритъм в създаването на следващо поколение участват само функциите, преминали някакъв праг на нелинейност.

Във втория (RGA2) като допълнителен кретерий за участие в създаване на следващотопоколение се използва ценова функция, описана в литературата, базирана на Уолш-Адамаровия спектър на таблицата. Разлика с RGA1 е, че вместо един начин на случайно „кръстосване“ се използват 5.

В RGA3 се добавя още едно условие за участие на таблица в създаване на следващо поколение, базирано на диференциално еднообразие. Колкото до варинта advRGA3, в него е добавено и изискването, една ли няколко компонентите на таблицата (от текста не става ясно колко точно, а псевдокодът е труден за разчитане) да са с нелинейност по-голяма от тази на началната популация, в опит да се атакуват неполучавани параметри.

Като цяло приемам йерархичния подход на създаване на фамилията, като за всеки следващ алгоритъм се търси подходящо модифициране така, че да се запазят добрите му качества и се потърси подобряване. Слабост на фамилията е, че доста зависи от избора на редица константи – и при кръстосването и при селекцията – за които не е съвсем ясно влиянието им върху резултатите.

При втората фамилия търсенето започва от произволна таблица като с всяко поколение се подобряват исканите свойства до получаване на таблица с добри свойства. Първият алгоритъм SIA1 е комбинация на hill climbing с cloning. Ценовата функция в hill climbing е ппроизведение на една известна от литературата и две, създадени от автора в резултат на многобройни експерименти. От описанието на алгоритъма следва, че той завършва с една таблица и не е ясно как в този случай се постига целта – генериране на голямо множество от таблици. Алгоритъмът SIA2 е модификация на SIA1, имаща за цел да създаде много специфични таблици – с размер 6x6 и такива, че не са афинно еквивалентни на известни от литературата таблици – не е посочено защо е рашавана такава специфична задача. За целта ценовата функция е заменена с по-подходяща, а като начални са използванr таблиците, които докторантът се е опитвал да избегне или техни афинни еквивалентни.

И при тези два алгоритъма има константи, от които зависи работата на алгоритъма, евристичното, случайно или базирано на експерименти, избиране на които може както да доведе до успех така и до недобри резултяати.

Глава 5 е посветена на експериментите, извършени с разработените от докторанта алгоритми. В случая на фамилията генетични алгоритми, са получени голям брой таблици с неголеми размери (8, 10), които са обилен материал за преценка на качествата на алгоритмите и най-вече за избора на управляващите алгоритмите константи – прагове, референтни стойности и брой итерации. Естествено, при нарастването на размера на таблиците, бързодействието на алгоритмите намалява, въздействието на евристично избраните константи се чувства все по-осезаемо и за постигане на очакваните резултати ще се налага допълнителни експерименти и създаване на по-добри модификации на алгоритмите.

В края на краищата ясно се вижда как в йерархията от алгоритми, на всяко следващо ниво, бързодействието намалява за сметка на подобряване на резултатите. Такова поведение е приемливо, като се има предвид, че процесът на търсене не е в реално време, а и с нарастване на броя на ядрата на съвременните компютърни конфигурации, бързодействието ще се подобрява. Въпреки някои относително негативни резултати от експериментите, сравнението с известни до моменти подходи показва, че получените от докторанта резултати са по-добри от доста от известните до момента и са сравними с най-добрите.

Резултатите при работата на имунните алгоритми също са добри. Макар че при стартиране от произволни начални таблици, трудно се достигат най-добрите параметри постигани от FFI-алгоритъма. Въпреки това, с тези алгоритми са получени комбинации от параметри при таблиците с малки размери, които не са били постигани с друг евристичен алгоритъм. Двата варианта не могат да се сравняват помежду си, защото имат различни цели. Резултатите показва, че пионерската идея на докторанта да разработи имунен алгоритъм за генериране на многобройни БЗТ е успешна и след достатъчно експерименти, за настройване на евристичните константи, може да бъдат получени и по-добри резултати. В сравнение с фамилията генетични алгоритми, имунните отстъпват по качество на резултатите, но това е естествено, тъй като опитът с използване на първите е доста по-голям. Сравнителният анализ на разработените алгоритми в края на главата е най-ценното в нея, според нас. Би ни се искало да видим и някаква оценка за сложността по време, но това при евристичните алгоритми е предизвикателство.

Глава 6 е заключителна за работата. Тя съдържа ненужно повторен, според нас, преглед на съдържанието на дисертацията, каквото срещнахме вече като

резюме и разширено в Първа глава и каквото има в достатъчно добър вид в Автореферата, както и списък с приносите на дисертацията според автора, мястото на които е тук, а ненужно е повторено в Първа глава.

5.Приноси на дисертационния труд

1. Докторантът се е запознал в подробности с криптографските свойства на БФ и БЗТ, както и със съществуващите техники за генериране на БЗТ, с което е изпълнена образователната цел на докторантурата.
2. Намерени са ефективни комбинации от евристични техники за търсене на БЗТ с колкото може по-добри свойства, както и пораждане на големи множества от БЗТ с предварително зададени криптографски свойства и са създадени 6 алгоритъма, в които известни техники са комбинирани с нови такива, предложени от автора.
3. Генерирани са много БЗТ с неголеми размери, които са по-добри от получените с някои други техники и сравними с най-добрите известни до момента. Желанието да се разрешат някои нерешени проблеми за съществуване на БЗТ с определени свойства, видимо не е изпълнимо с разработените от докторанта инструменти.
4. Направен е сравнителен анализ на създадените алгоритми, установени са предимствата им и някои недостатъци, които могат да бъдат обект на бъдещи изследвания. Приложени ли са резултатите в реални шифри не може да знаем.

6. Критични бележки

Забелязаните в текста неточности и неясноти са предоставени на докторанта в отделен документ. Съществени забележки по съдържателната част нямаме. Не харесваме стила на автора да повтаря ненужно, например дефиницията на едно свойство при споменаване на свойството, или пък неформалното описание на едно понятие непосредствено преди формалната дефиниция.

Съществена забележка имаме към твърдението на 18/19 стр. че алгоритъмът, който е вариант на Гаусова елиминация е непрактичен и приложим само за малки n, защото всъщност той е на порядък по-ефективен от умножението на вектор по матрица, заради свръх-симетричността на матрицата A_n на трансформацията на Уолш-Адамар – времева сложност $O(n2^n)$ в първия, срещу $O(2^{2n})$ за втория. И в по-общ план, препоръчваме на докторанта по-задълбочено да подходи към сложността на алгоритмите по време, заради забелязани дребни недостатъци от този вид в алгоритмите му.

7. Преценка на публикациите по дисертационния труд.

По дисертацията има 4 публикации. Работата под номер 3 е публикувана в реномираното списание Cryptography and Communications (последен импакт фактор 0.742). Работата под номер 3 е доклад на международна конференция

и е измежду избраните за публикуване (12 от 27) в тома на конференцията, издаден от Springer след сериозно рецензиране. Работата под номер 1 е публикувана в сборник доклади на конференция у нас, обявена като международна (макар че от тома на конференцията не личи това да е така), затова предполагаме, че работата е рецензирана. Работата под номер 2 е публикувана в електронен архив Cryptology ePrint Archive, статутът на който не ни е известен. Всички публикации отразяват резултати от дисертационния труд. Смятаме, че изискването на Правилника за три рецензирани публикации, поне едната от които е в списание е изпълнено. Не приемаме за добра практика повтарянето на части от една публикация в друга.

Самостоятелни публикации докторантът няма, но в три от четирите работи той е първи съавтор, а в четвъртата – втори. Това ни дава основание да заключим, че приносът му е значителен.

Независимо от неголемия брой публикации по дисертацията, известните до момента цитирания на са доста, 9, като5 от тях са за работата под номер 3, и по две за работите 1 и 4. За съжаление, цитатите са в статии, които не са ни достъпни и не можем да преценим значимостта на цитиранията.

В заключение можем да изкажем увереността си, че публикациите достатъчно добре отразяват съдържанието на дисертационния труд и са получили публичността и признанието, необходими за положителна оценка на дисертацията.

8. Авторефератът е изготвен в съответствие с изискванията на Правилника за условията и реда за придобиване на научни степени и за заемане на научни длъжности на ИМИ–БАН, като отразява изчерпателно и точно съдържанието на дисертационния труд и приносите.

9. Не познавам докторанта лично.

10. Заключение. Като имам предвид казаното по-горе, смятам че представеният дисертационен труд отговаря на изискванията на ЗРАСРБ, Правилника за прилагане на ЗРАСРБ, Правилника за развитие на академичния състав на БАН и на ИМИ при БАН, и му давам положителна оценка. Предлагам на уважаваното жури да присъди на Георги Великов Иванов научната и образователна степен "Доктор" в област на висшето образование 4. Природни науки, информатика и информатика, професионално направление: 4.6. Информатика и компютърни науки, научна специалност 01.01.12 „Информатика”

София, 09.12.2016 г. Рецензент:

проф. д-р Красимир Манев