Warszawa, 14 września 2020r.

dr hab. Mirosław Kurkowski, prof. UKSW, prof. WSPol

- **Instytut Informatyki**
  Uniwersytet kard. St. Wyszyńskiego w Warszawie
- **Zakład Cyberbezpieczeństwa**
  Wyższa Szkoła Policji w Szczytnie

Recenzja rozprawy doktorskiej mgra inż. Pawła Augustynowicza

Efektywne badanie nieprzywiedlności wielomianów binarnych o szczególnych postaciach

Promotor: dr hab. inż. Andrzej Paszkiewicz, prof. WAT

Niniejsza recenzja została sporządzona na prośbę Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja wyrażonej w odpowiednim piśmie przez dra hab. inż. Zbigniewa Piotrowskiego, prof. WAT (uchwała 22/RDN ITiT/2020). Opiniowana rozprawa dotyczy metod efektywnego badania nieprzywiedlności wybranych postaci wielomianów nad ciałem GF(2) mających zastosowanie w kryptograficznych systemach zabezpieczeń. Przewód doktorski prowadzony jest zgodnie z tzw. „starą procedurą" w dziedzinie nauk technicznych, w dyscyplinie naukowej Informatyka.

Wprowadzenie

Konieczność zapewniania poufności przesyłanych danych jest niekwestionowaną cechą wielu kluczowych systemów komunikacyjnych, zarówno społecznych, jak i urzędowych, czy militarnych. Dane te przesyła się najczęściej przez sieci komputerowe lub inne wyspecjalizowane systemy łączności, a następnie przechowuje się w odpowiednio zabezpieczonych bazach danych. W wielu przypadkach całość tych informacji musi być odpowiednio chroniona przed niepowołanym dostępem. W tej sytuacji jedynym odpowiednim środkiem, który może zapewnić potrzebną ochronę wrażliwych danych jest zastosowanie odpowiednich mechanizmów i algorytmów kryptograficznych.

Rozważane w rozprawie wielomiany binarne o szczególnych postaciach znajdują zastosowanie w konstrukcji lekkich, ale silnych kryptograficznie szyfrów strumieniowych. Dlatego biorąc pod uwagę obecny stan nauki w rozważanym zakresie oraz najnowsze potrzeby i trendy badawcze można z całą pewnością stwierdzić, że badanie własności wybranych typów wielomianów nad GF(2), a zwłaszcza ich nieprzywiedlności oraz efektywnych algorytmów odpowiadających na związane z tymi zakresami pytania jest

całkowicie uzasadnione i ważne biorąc pod uwagę bieżące problemy związane z rozwojem systemów zabezpieczeń w sieciach komputerowych i innych systemach łączności.

Zawartość rozprawy

Recenzowana rozprawa liczy 87 stron, nie licząc Spisu Treści, dodatków oraz Bibliografii. W mojej opinii układ rozprawy nie budzi zastrzeżeń i jest odpowiedni. Zacytowana w pracy literatura przedmiotu liczy 82 pozycje i biorąc pod uwagę potrzeby zawartych w rozprawie rozważań jest moim zdaniem dobrana wystarczająco i odpowiednio. Wspólnym mianownikiem łączącym zasadnicze tematy rozprawy jest zastosowanie obliczeń wielkoskalowych i współczesnych architektur obliczeniowych do poszukiwań wielomianów mających szczególne własności, ważne z punktu widzenia projektowania struktur algebraicznych mających praktyczne zastosowania w realnych systemach zabezpieczeń.

We Wstępie do rozprawy autor określił Cel pracy i hipotezy badawcze:

Celem niniejszej pracy jest systematyczne i empiryczne sprawdzenie wybranych własności trójmianów, wielomianów osadowych oraz rejestrów przesuwnych o nieliniowej funkcji sprzężenia zwrotnego, określanych dalej jako rejestry NLFSR (ang. Nonlinear Feedback Shift Register).

Doktorant określił również trzy Tezy badawcze:

1. Prowadzenie rozległych eksperymentów obliczeniowych pozwala na weryfikację prawdziwości znanych hipotez i sformułowanie nowych hipotez oraz twierdzeń.

2. Możliwym jest skonstruowanie efektywnych algorytmów badania nieprzywiedlności wielomianów binarnych nad GF(2), uwzględniających specyfikę współczesnych systemów obliczeniowych, w tym procesorów umożliwiających przetwarzanie wektorowe oraz koprocesorów graficznych, jak również prowadzenie przy ich wykorzystaniu zaawansowanych eksperymentów obliczeniowych.

3. Środowisko obliczeniowe kart graficznych pozwala na efektywne badanie pełności okresu rejestrów przesuwnych o szczególnych, nieliniowych postaciach funkcji sprzężenia zwrotnego.

W skład rozprawy poza Wstępem wchodzi pięć rozdziałów.

Rozdział pierwszy zawiera wprowadzenie do rozważanej w pracy tematyki. Sformułowano tutaj w sposób ogólny cele naukowe opracowania nowych, efektywnych metod poszukiwania generatorów pewnych struktur algebraicznych mających praktyczne zastosowania w systemach zabezpieczeń, zakres rozprawy oraz kierujące autorem motywacje badawcze.

Kolejno w rozdziale drugim zaprezentowano bardzo szczegółowo podstawowe definicje i twierdzenia związane z poruszaną w nim tematyką. Przedstawiono wybrane elementy historii rozwoju algorytmów badania nieprzywiedlności wielomianów binarnych, w tym kilka algorytmów badania nieprzywiedlności (algorytmy Ben-Ora, Rabina, ich połączenie) oraz mnożenia wielomianów (Karatsuby i oparty na szybkiej transformacie Fouriera algorytm Schonhagiego). Szczególnie precyzyjnie przedstawiono podstawy matematyczne, na których algorytmy te bazują, operacje w nich wykonywane oraz analizę ich złożoności. Zauważyć należy, że najpierw autor omówił konstrukcje algorytmów o mniejszej złożoności obliczeniowej, by następnie dostosowywać istniejące rozwiązania do wymagań wdrażanych implementacji oraz architektury zastosowanych w badaniach eksperymentalnych jednostek obliczeniowych. Kolejno mgr inż. Augustynowicz opisuje autorską modyfikację algorytmu Ben-Ora pozwalającą na bardziej efektywne badanie nieprzywiedlności wielomianów binarnych. Zmiany uwzględniają proponowane środowiska obliczeniowe współczesnych jednostek procesorowych oraz koprocesorów graficznych. Pod koniec rozdziału zaprezentowano wyniki wielu serii badań eksperymentalnych oraz wyciągnięto odpowiednie wnioski.

W rozdziale trzecim doktorant przedstawił i szczegółowo omówił problematykę badania własności wielomianów osadowych, a szczególnie hipotezę mówiącą o stopniu wewnętrznym wielomianu osadowego. Mówi ona, że *dla każdego stopnia n istnieje binarny wielomian osadowy, którego stopień wewnętrzny nie przekracza wartości* $\log_2 n+3$. Prowadzone dalej badania są kontynuacją prac prof. Paszkiewicza, który pierwszy podał kontrprzykład obalający tę hipotezę. Prezentowany w dalszych częściach omawianego rozdziału rozległy eksperyment obliczeniowy miał na celu wyznaczenie wszystkich wielomianów osadowych do stopnia $n = 513 \cdot 10^3$. Jest to obecnie wynik znacznie przewyższający wszystkie znane do tej pory eksperymenty obliczeniowe w tym zakresie. W ramach eksperymentów zaprojektowano i zaimplementowano własną bibliotekę badania nierozkładalności, która jest dedykowana na koprocesory graficzne (maksymalizacja wykorzystania zasobów). Kolejno zastosowano dalszy rozwój algorytmów badania nieprzywiedlności poprzez zaproponowanie własnej modyfikacji algorytmu dostosowanej do współczesnych architektur obliczeniowych i wymagań eksperymentu obliczeniowego. Warto zaznaczyć, że autor zastosował wymuszanie numeracji kolejnych wielomianów zgodnie z porządkiem leksykograficznym dla uzyskania możliwości odtworzenia przebiegu obliczeń na wypadek ich niezaplanowanego przerwania.

Rozdział czwarty kontynuuje rozważania dotyczące systemów obliczeniowych oraz wybranych postaci wielomianów w kontekście efektywności obliczeń w implementacjach programistycznych i sprzętowych. Tym razem autor rozważa ciekawe własności trójmianów,

nawiązując do prac znanego projektu „The Great Polynomial Hunt" poszukującego wielomianów pierwotnych i nierozkładalnych, zwłaszcza o stopniu będącym liczbą Mersenna. Autor podobnie jak w poprzednim rozdziale rozszerza wyniki prof. Paszkiewicza, który zajmował się dotąd trójmianami postaci $x^{2\cdot 3^l} + x^{3^l} + 1$, dla $l \in \{1,2,\dots,12\}$, na liczbę $l = 13$. W prowadzonych badaniach zostały nie tylko wykonane obliczenia dla $l = 13$, ale również potwierdzone wszystkie dotychczasowe wyniki dla $l < 13$. Prowadząc wyżej opisane rozważania autor najpierw wprowadził odpowiednie definicje oraz twierdzenia, po czym wprowadził i opisał odpowiedni algorytm. Przytoczone są znane i ważne problemy związane z trójmianami nierozkładalnymi:

- *Jaka jest graniczna wartość ułamka wszystkich liczb naturalnych, dla których istnieje trójmian stopnia n nierozkładalny nad ciałem binarnym?*
- *Jaka jest średnia liczba trójmianów przypadających na stopień?*
- *Czy można znaleźć ciąg ośmiu kolejnych liczb naturalnych n, dla których nie istnieje trójmian nierozkładalny stopnia n?*
- *Czy dla kolejnych liczb parzystych k = 0, 2, 4, 6,... ułamek stopni, dla których istnieje dokładnie k trójmianów nierozkładalnych ustala sie wokół określonych liczb rzeczywistych?*
- *Czy istnieją dowolnie długie ciągi kolejnych stopni n, dla których nie istnieje binarny trójmian nieprzywiedlny?*

Doktorant przedstawia następnie olbrzymią ilość prac eksperymentalnych badających wyżej wymienione własności dla pewnych klas wielomianów komentując uzyskane wyniki oraz wyprowadzając odpowiednie wnioski. Należy podkreślić, że ogromna liczba prac obliczeniowych opisanych w tym rozdziale jest co najmniej imponująca.

W rozdziale piątym mgr inż. Augustynowicz zajął się rejestrami przesuwnymi o funkcji sprzężenia zwrotnego szczególnej postaci i ich własnościami w kontekście efektywności ich generowania. Jak w poprzednich rozdziałach autor doskonale wprowadził czytelnika w podstawy teoretyczne związane z tematem. Opisane zostały metody poszukiwania kwadratowych m-sekwencji, w tym z wykorzystaniem matryc programowalnych bramek logicznych FPGA. Wymienno również metodę Szmidta wykorzystującą logarytmy Zecha. Zbadano problem pełności okresu nieliniowego rejestru przesuwnego i podano odpowiedni algorytm. Opisano następnie implementację wykonaną na jednostki wieloprocesorowych oraz koprocesorach graficznych. Kolejno w pracy pokazano wyniki eksperymentalne przedstawiające wydajność zaprojektowanych algorytmów na różnych platformach. W ramach przeprowadzonego kompleksowego i całościowego eksperymentu

obliczeniowego znaleziono wszystkie rejestry przesuwne o funkcji sprzężenia zwrotnego wybranej postaci dla długości rejestrów $n \leq 30$.

Opinia merytoryczna rozprawy

Jak napisałem wcześniej, układ rozprawy moim zdaniem jest odpowiedni i nie budzi zastrzeżeń. Wstępne części rozprawy wprowadzające podstawowe pojęcia i definicje struktur potrzebne do dalszych rozważań moim zdaniem zostały opracowane bardzo dobrze. Uważam, że czytelnik jest dobrze zaznajomiony z podstawami teoretycznymi oraz obecnym stanem wiedzy w rozważanej tematyce, aby dalej zrozumiale śledzić zawarte w rozprawie treści. Należy po raz kolejny podkreślić, że wszystkie rozdziały zawierają bardzo dobrze przemyślane i opracowane wprowadzenie teoretyczne w rozważaną problematykę.

W pracy zaproponowano i praktycznie zaimplementowano rodzinę efektywnych algorytmów badania nieprzywiedlności wielomianów nad ciałem binarnym oraz poszukiwania rejestrów NLFSR o maksymalnym okresie. Algorytmy te uwzględniają możliwości i ograniczenia współczesnych architektur obliczeniowych, w tym szczególna uwagę położono na wykorzystanie najnowszych instrukcji udostępnianych zarówno przez jednostki procesorowe CPU (ang. Central Processing Unit) jak i graficzne moduły obliczeniowe GPU (ang. Graphical Processing Unit).

Wykazano zalety zaimplementowanych rozwiązań, w tym ich wszechstronność i skalowalność. Mogą być one dostosowane do nieomalże każdej platformy obliczeniowej, stopnia wielomianu lub jego postaci.

Podsumowując szczegółowo wyniki zawarte w rozprawie należy podkreślić, że:

- przeprowadzone przez doktoranta wyniki eksperymentalne pozwoliły na znalezienie wielu kontrprzykładów do znanej hipotezy dotyczącej stopnia wewnętrznego wielomianu osadowego. W badaniach tych potwierdzono dotychczasowe wyniki oraz przesunięto wyżej granicę poznania tego fragmentu matematycznej rzeczywistości.
- uzyskane dane empiryczne pozwoliły na obserwacje istotnych, z punktu widzenia zastosowań kryptograficznych, zależności dotyczących funkcji średniego stopnia wewnętrznego wielomianu osadowego. Wnioski te pozwoliły na zaproponowanie nowej metody poszukiwania wielomianu o stopniu wewnętrznym bliskim optymalnego, która w praktyce okazuje sie wydajniejsza od tradycyjnego procesu enumeracji leksykograficznej.

- systematyczne i wyczerpujące prowadzenie badań pozwoliło na zebranie następujących zbiorów danych:

  1. najmłodszych leksykograficznie wielomianów osadowych o stopniu mniejszym, bądź równym wartości 513000,

  2. trójmianów nieprzywiedlnych o stopniu mniejszym, bądź równym wartości 513000;

  3. trójmianów nierozkładalnych towarzyszących wielomianów postaci x2·3l + x3l + 1 dla l 2 {1, 2, . . . 13};

  4. rejestrów przesuwnych o szczególnej, nieliniowej postaci funkcji sprzężenia zwrotnego i maksymalnym okresie do stopnia 30 włącznie.

- wyniki przeprowadzonych badań eksperymentalnych poddano gruntownej analizie, która pozwoliła na weryfikacje niektórych spośród znanych hipotez dotyczących rozważanych w rozprawie klas wielomianów, poczynienie interesujących obserwacji, jak i wyznaczenie nowych, nieskończonych klas trójmianów nierozkładalnych.

Podsumowując tę część recenzji stwierdzam jednoznacznie, że moim zdaniem mgr inż. Paweł Augustynowicz zrealizował postawione sobie Cele badawcze i wykazał prawdziwość postawionych na początku rozprawy tez.

Uwagi polemiczne i krytyczne oraz elementy dyskusyjne

Przedstawione niżej uwagi nie zmniejszają moim zdaniem wartości naukowej rozprawy i nie mają wpływu na pozytywną opinię pracy jako całości. Zamieszczone uwagi mogą też stanowić pole do dalszych badań.

Można mieć wrażenie, że niedosyt budzi brak wyjaśnienia, czy możliwe jest uniknięcie pewnej anomalii związanej z realizacją algorytmu poszukiwania rejestrów przesuwnych o szczególnej postaci funkcji sprzężenia zwrotnego wykorzystującą układy graficzne. Okazało się bowiem, że uśredniony czas badania jednego pełnego okresu przez jeden wątek dla rozmiarów rejestrów mniejszych niż 30 rośnie w postępie liniowym wraz z rozmiarem rejestru. Inna jest sytuacja dla rejestrów dłuższych niż 30. Następuje wtedy nagły spadek wydajności powodując tym samym gwałtowny wzrost czasu badania. Zjawisko to jest to o tyle ciekawe, że po pierwsze podobny spadek wydajności nie jest obserwowany w przypadku jednostek procesorowych, po drugie zaś w przypadku nowszych rodzin układów graficznych wpływ rozmiaru przetwarzanego słowa na czas jego przetwarzania jest coraz mniej zauważalny. Nie jest jasne, czy można dokonać pewnej lekkiej reorganizacji całości algorytmu, która pozwoliłaby na zbadanie przy pomocy kart graficznych rejestrów o długościach większych niż 30?

W pracy brakuje informacji, czy zaprezentowany algorytm testowania okresu sekwencji generowanej przez rejestr przesuwny o zadanej funkcji sprzężenia zwrotnego

można zaimplementować na układach reprogramowalnych matryc bramek logicznych FPGA oraz czy w tym przypadku prezentuje on lepszą, czy gorszą wydajność od rozwiązań zaprezentowanych między innymi w pracach Janusza Szmidta i Nikolaya Poluyanenko.

W recenzowanej rozprawie istnieje kilka aspektów otwartych, w ramach których można przeprowadzić szerszą dyskusję podczas publicznej obrony.

Po pierwsze jedną z kwestii dyskusyjnych jest możliwość dostosowania opracowanego algorytmu do badania nieprzywiedlności wielomianów nad innymi ciałami prostymi o wybranej postaci, na przykład dwumianów nad ciałem charakterystyki 3. Z opisu algorytmu oraz jego właściwości można odnieść wrażenie, iż takie modyfikacje mogły być możliwe. Wątek takiej możliwości rozszerzenia opracowanego rozwiązania nie został w rozprawie zawarty.

Kolejną ciekawą otwartą kwestią jest sprawa dotycząca ciągłego i szybkiego rozwoju środowisk kart graficznych, a co za tym idzie dalszym znacznym wzrostem wydajności prowadzonych na nich obliczeń. Zastosowana w badaniach opisanych w rozprawie seria kart graficznych Pascal nie jest już obecnie najnowszym rozwiązaniem w tym zakresie. Jest to zapewne spowodowane faktem, że eksperymenty zawarte w rozprawy były wykonywane na przestrzeni kilku lat. Pojawia się zatem pytanie, czy przy zastosowaniu nowszych kart na przykład z najnowszej architektury Ampere będzie możliwy do uzyskania jeszcze większy wzrost wydajności? Do jakich granic? Kolejne interesujące pytanie brzmi również, czy istnieje układ graficzny, który zaprezentuje lepsze wyniki wydajnościowe, niż jednostki procesorowe.

Uwagi redakcyjne

Jak każda praca naukowa również recenzowana rozprawa nie jest wolna od niedociągnięć, pomyłek, czy błędów natury redakcyjnej. Z przyjemnością muszę stwierdzić, że rozprawa mgra inż. Pawła Augustynowicza zawiera wyjątkowo mało takich pomyłek i jest chyba najlepiej zredagowaną rozprawą doktorską jaką do tej pory recenzowałem. Poniżej zamieszczam listę kilku przykładowych pomyłek/błędów:

- str. 13, wers 4 – *błędne oznaczenie składowych wielomianów sumy (f_j zamiast g_j)*,
- str. 13 – *...zachodzi dla jakiś...*
- str. 22 – *... składa sie z skalowalnej macierzy...*
- pusty fragment strony 58,
- str. 56 – *„problemu .”*
- str. 85 – *„...jest osiągnięciem wcześniej lepszym...”*.

Wniosek końcowy

Przedstawione w recenzowanej rozprawie doktorskiej rozważania związane z metodami efektywnego badania nieprzywiedlności wybranych postaci wielomianów nad ciałem GF(2) dotyczą bieżących, ważnych i interesujących problemów naukowych związanych z konstrukcją i metodami analizy współczesnych systemów kryptograficznych. Rozprawa doktorska mgra inż. Pawła Augustynowicza zawiera wiele oryginalnych oraz interesujących wyników. Należy podkreślić wykonanie olbrzymiej liczby badań eksperymentalnych oraz bardzo staranne i merytorycznie można powiedzieć, że wzorcowe przygotowanie rozprawy. Moje uwagi krytyczne zawarte w recenzji nie zmieniają pozytywnej opinii o rozprawie jako całości.

Biorąc pod uwagę wyniki naukowe przedstawione w recenzowanej rozprawie doktorskiej mgra inż. Pawła Augustynowicza stwierdzam, że moim zdaniem, praca ta spełnia wymagania stawiane rozprawom doktorskim przez obowiązującą aktualnie w Polsce Ustawę o Stopniach i Tytule Naukowym. Stawiam zatem wniosek o dopuszczenie mgra inż. Pawła Augustynowicza do dalszych etapów przewodu doktorskiego prowadzonego w dziedzinie nauk technicznych w dyscyplinie Informatyka przez Radę Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja Wojskowej Akademii Technicznej.