Wrocław 08.09.2022r.

Dr hab. inż. Grzegorz Kołaczek, prof. uczelni
Katedra Informatyk i Inżynierii Systemów
Wydział Informatyki i Telekomunikacji
Politechnika Wrocławska
ul. Wybrzeże Wyspiańskiego 27
50-370 Wrocław

Recenzja rozprawy doktorskiej

Autor:

mgr Sylwia Stachowiak

Tytuł:

„SAT-kryptoanaliza wybranych algorytmów kryptografii symetrycznej”

Promotor:

dr hab. Mirosław Kurkowski,
prof. Uniwersytetu Kardynała Stefana Wyszyńskiego,
Wyższa Szkoła Policji w Szczytnie

Promotor pomocniczy:

dr hab. Konrad Furmańczyk, prof. Uczelni.
Szkoła Główna Gospodarstwa Wiejskiego

Niniejsza opinia została przygotowana na prośbę dr hab. Ryszarda Kozery, prof. SGGW, Przewodniczącego Rady Dyscypliny Informatyka Techniczna i Telekomunikacja z dnia 13.07.2022r.

1. Problem badawczy i jego znaczenie

Rozprawa doktorska mgr Sylwii Stachowiak dotyczy problematyki kryptoanalizy współczesnych algorytmów symetrycznych poprzez ich reprezentację w postaci kodowania boolowskiego oraz wykonanie analizy z użyciem SAT-solverów. Tym samym tematyka badawcza podjęta w rozprawie wpisuje się w zakres badań dotyczących szeroko rozumianego bezpieczeństwa danych przewarzanych, przesyłanych i przechowywanych w systemach informatycznych. Możliwość zapewnienia danym bezpieczeństwa, stanowi zarówno jeden z podstawowych, jak i też krytycznych warunków dla działania wielu systemów informatycznych wspierających realizację procesów biznesowych w ramach różnych działów współczesnej gospodarki. Funkcjonowanie wielu usług, poprzez usługi finansowe, a skończywszy na popularnych komunikatorach, wymaga dostarczenia systemu

gwarantującego bezpieczne przetwarzanie i przesyłanie danych. Świadczenie tego typu usług nie byłoby możliwe, gdybyśmy nie dysponowali możliwością zapewnienia bezpiecznych kanałów komunikacyjnych pomiędzy klientem i usługodawcą, a także możliwością zweryfikowania tożsamości komunikujących się podmiotów.

Jednym z podstawowych elementów umożliwiających budowanie bezpiecznych systemów informatycznych są symetryczne algorytmy szyfrujące. Dlatego też, możliwość oceny bezpieczeństwa wykorzystywanego algorytmu stanowi jedno z podstawowych i tym samym istotnych zadań współczesnej kryptologii.

W pracy przedstawiono metodę i kryptoanalizę wybranych współczesnych symetrycznych algorytmów szyfrujących, takich jak DES, Salsa20 oraz AES, zatem należy uznać, iż tematyka rozprawy wpisuje się w istotny i aktualny nurt prac badawczych z dziedziny zagadnień przynależnych do informatyki technicznej.

2. Kompozycja rozprawy

Recenzowana rozprawa liczy 183 strony. Została ona podzielona na sześć rozdziałów oraz dwa dodatki (A i B), w których zamieszczono wyniki statystyczne eksperymentu polegającego na wielokrotnym łamaniu szyfru DES dla różnych SAT-solverów i różnych wektorów danych wejściowych oraz tabelę z przykładem ilustrującym podstawienie S-Box w algorytmie AES. W pracy zamieszczono również spis rysunków, spis tabel oraz bibliografię złożoną z 145 pozycji literaturowych.

Układ i zawartość rozdziałów merytorycznych jest zasadniczo poprawna. W pierwszym rozdziale zamieszczono wprowadzenie, w którym został w sposób syntetyczny przedstawiony cel badań oraz została sformułowana główna teza badawcza rozprawy. Ponadto wymienione zostały główne rezultaty naukowe rozprawy oraz scharakteryzowano strukturę dalszej części rozprawy.

W rozdziale drugim zamieszczono obszerne wprowadzenie do zagadnień z dziedziny kryptologii oraz problemu SAT wraz z jego zastosowaniami. Autorka umieściła w tym rozdziale definicje podstawowych pojęć takich jak szyfrowanie, deszyfrowanie, system kryptograficzny. Umieszczono i przedstawiono charakterystykę podstawowych klas algorytmów szyfrujących poczynając od szyfrów klasycznych poprzez współczesne algorytmy symetryczne blokowe oraz strumieniowe i algorytmy asymetryczne. W dalszej części tego rozdziału przedstawiono taksonomię podstawowych metod kryptoanalitycznych oraz związane z kryptoanalizą zagadnienia z teorii złożoności obliczeniowej. Ostatnia część pierwszego rozdziału charakteryzuje problem spełnialności formuł boolowskich (SAT) wraz z podaniem rozbudowanego przykładu oraz opisuje rozwój SAT-solverów i przedstawia w jaki sposób SAT-solvery znajdowały zastosowanie w analizie algorytmów kryptograficznych.

Rozdział trzeci przedstawia zagadnienia dotyczące algorytmu DES oraz jego kryptoanalizy. W pierwszej części rozdziału (podrozdziały 3.1-3.3) w sposób wyczerpujący została przedstawiona charakterystyka i budowa algorytmu DES wraz z jego elementami składowymi. W kolejnej części tego rozdziału przedstawiona została metoda kodowania elementów algorytmu DES do postaci formuł logicznych, a następnie do postaci formatu DIMACS, który to format jest wymagany dla zbioru danych wejściowych przez wykorzystywane w badaniach

SAT-solvery. Ostatnia część tego rozdziału zawiera wyniki badań eksperymentalnych polegających na siłowym łamaniu kryptogramu z wykorzystaniem SAT-solverów dla kilku wariantów algorytmu DES. Badane warianty dotyczyły kryptoanalizy oryginalnej wersji DES oraz algorytmu DES ze zmodyfikowanym zestawem S-boxów, w tym z zaprojektowanymi przez Autorkę pracy S-boxami o charakterystyce liniowej. Zbadano również zależności pomiędzy liczbą i umiejscowieniem nieznanych bitów klucza a czasem kryptoanalizy.

Kolejny rozdział jest dedykowany kryptoanalizie algorytmu Salsa20. Podobnie jak w poprzednim rozdziale najpierw Autorka przedstawiła zasadę działania algorytmu, a następnie translację formuł logicznych do formatu DIMACS. Przedostatni podrozdział przedstawia wyniki badań eksperymentalnych, ataków przeprowadzonych na uproszczoną wersję algorytmu do dwóch i czterech rund przy wykorzystaniu Sat-solverów jedno oraz wielowątkowych i z różną liczbą znanych/nieznanych bitów klucza. Ostatnia część tego rozdziału zawiera podsumowanie przeprowadzonych prac.

Rozdział przedostatni zawiera specyfikację algorytmu AES wraz z pokazaniem sposobu kodowania elementarnych przekształceń algorytmu szyfrującego do postaci kodowania boolowskiego oraz przekształcenia do formatu DIMACS. Podobnie jak we wcześniejszym rozdziale umieszczono wyniki przeprowadzonych badań, które dotyczyły możliwości uzyskania poprawnej wartości klucza kryptograficznego dla uproszczonej do jednej rundy wersji algorytmu AES oraz różnej liczebności i pozycji nieznanych wartości bitów klucza. Rozdział ten również zawiera syntetyczne podsumowanie.

Ostatni, szósty rozdział przedłożonej rozprawy zawiera podsumowanie. Podsumowanie to, w znacznej części składa się z fragmentów zawierających podsumowania umieszczone we wcześniejszych rozdziałach pracy, a które to przedstawiały badania dedykowane algorytmom DES, Salsa20 oraz AES. W tym rozdziale również umieszczono krótki akapit dotyczący ewentualnych kierunków dalszych prac.

3. Oryginalne osiągnięcia

W recenzowanej pracy zawarto kilka wartościowych i oryginalnych koncepcji jak i też praktycznych rozwiązań zrealizowanych w postaci implementacji w formie kodu umożliwiającego analizę algorytmów z wykorzystaniem SAT-solverów. Uzyskane wyniki wzbogacają wiedzę w zakresie wiedzy na temat możliwości wykorzystania automatycznej analizy realizowanej z użyciem SAT-solverów w zagadnieniach dotyczących bezpieczeństwa współczesnych algorytmów kryptograficznych. W szczególności do najbardziej znaczących osiągnięć Autorki rozprawy należy zaliczyć:

1. Opracowanie bezpośredniego kodowania boolowskiego liniowych S-boxów dla algorytmu DES oraz kodowania dla algorytmów Salsa20 oraz AES
2. Przeprowadzenie SAT-kryptoanalizy dla zmodyfikowanych algorytmów Salsa20 oraz AES
3. Zademonstrowanie praktycznych możliwości kryptoanalizy współczesnych algorytmów szyfrujących z wykorzystaniem SAT-solverów.

Powyższe osiągnięcia należy uznać za oryginalne i znaczące dla dyscypliny naukowej informatyka techniczna i telekomunikacja. Na tej podstawie Recenzent opiera swoją ogólną pozytywną ocenę rozprawy oraz osiągnięcia te stanowią równocześnie potwierdzenie postawionej przez Autorkę głównej tezy pracy, czyli pokazują, iż zastosowanie bezpośredniego kodowania boolowskiego oraz SAT-solverów jest skuteczną metodą do badania własności symetrycznych algorytmów szyfrujących.

4. Uwagi krytyczne

Przedłożona do recenzji rozprawa jest pod względem merytorycznym i redakcyjnym w przeważającej części napisana poprawnie. W ocenie Recenzenta jednak Autorka nie ustrzegła się jednak również pewnych braków, niedociągnięć i nieścisłości. Poniżej zostały zamieszczone najważniejsze uwagi krytyczne do treści, jak i formy pracy.

a. Uwagi natury ogólnej

1. Analiza literatury. Każda praca badawcza jest pewną kontynuacją czy też nawiązaniem do wcześniejszych prac badawczych. Formułując problem badawczy dążymy do uzupełnienia pewnych luki w aktualnym stanie wiedzy. Dlatego też jednym z istotnych elementów każdej pracy badawczej jest rzetelna analiza aktualnego stanu wiedzy. Możliwość trafnej identyfikacji obszarów niewiedzy, bądź też zagadnień, które mają potencjał badawczy, gdyż aktualnie dostępne rozwiązania nie są optymalne, związana jest przede wszystkim z rzetelną analizą aktualnej literatury przedmiotu. Praca posiada obszerną bibliografię, która w znacznej mierze pokrywa obowiązujący stan wiedzy, nie mniej jednak, w przekonaniu Recenzenta ten element rozprawy mógłby być zrealizowany w sposób bardziej systematyczny. W szczególności, encyklopedyczne przywołanie historii i typów kryptoanalizy (rozdział 2.4), akapit odnośnie kryptoanalizy algorytmu DES (rozdział 3.6), analogiczny przegląd zagadnień dotyczących kryptoanalizy algorytmu Salsa20 na początku rozdziału 4.5 oraz algorytmu AES w rozdziale 5.5, dają jedynie możliwość względnego odniesienia poziomu skuteczności zaproponowanej metody względem innych technik kryptoanalitycznych, natomiast nie stanowią pełnego kontekstu rozwoju i sposobu wykorzystania SAT-solverów we współczesnej kryptoanalizie. W szczególności, wartościowe byłoby wykazanie (lub wskazanie źródeł, gdzie taki dowód przeprowadzono), że zaproponowane podejście „charakteryzuje się brakiem nadmiarowości liczby stosowanych do kodowania szyfru zmiennych” (str. 57), co stanowi istotny rezultat przedłożonego osiągnięcia naukowego i ma stanowić element odróżniający przeprowadzone badania od wcześniej przeprowadzonych prac.
2. Proporcje pomiędzy częścią przeglądową a badawczą. Znaczną część pracy stanowi charakterystyka ogólnych pojęć i opis pełnej specyfikacji wybranych algorytmów. W opinii Recenzenta, ze względu na swój podstawowy charakter te fragmenty pracy mogły zostać opracowane w sposób bardziej skrótowy. Nie mniej jednak, pewną zaletą takiego podejścia przyjętego przez Autorkę

rozprawy może być możliwość wykorzystania tekstu rozprawy np. jako podręcznika czy też skryptu akademickiego.

3. Sposób opracowania części dedykowanej wstępowi do kryptologii. W ocenie Recenzenta ta część rozprawy (podrozdziały 2.1-2.4) powinna zostać napisana w oparciu o wybraną, dobrze ugruntowaną taksonomię (algorytmy ze względu np. na typ funkcji szyfrującej: podstawieniowe, przestawieniowe, hybrydowe, ze względu na rozmiar wejścia: strumieniowe, blokowe, ze względu na typ klucza: symetyczne, asymetryczne, itp.) W zaproponowanym ujęciu brak uporządkowania prezentowanych treści. Również w poprawnej interpretacji tekstu nie pomaga zastosowany sposób formatowania tekstu, który powoduje, że można odnieść wrażenie np. że algorytmy z przesunięciem, algorytmy podstawieniowe, algorytmy afiniczne, algorytm Vigenere – stanowią różne klasy/typy algorytmów (rozdział 2.1), a jest to tylko lista arbitralnie wybranych przez Autorkę rozprawy elementów ze zbioru algorytmów klasycznych – zarówno klas metod (algorytmy podstawieniowe), jak i też przykładów konkretnych algorytmów (np. alg. Vigenera).
4. Precyzja języka.
    a. Autorce rozprawy nie udało się osiągnąć wysokiego poziomu precyzji opisu, np. w dosyć swobodny sposób, zamiennie w treści rozprawy stosowane są terminy „dane" oraz „informacja". Pojęcia te są zbliżone do siebie, ale na pewno nie są tożsame, zwłaszcza w kontekście dziedziny jaką jest bezpieczeństwo.
    b. Recenzent również ma pewne zastrzeżenia do zrównania co do znaczenia terminów „tajność" oraz „poufność" (rozdział 2.1.).
    c. Stosowanie terminu „ilość bitów" zamiast „liczba bitów"
    d. Stosowanie terminu „dodane bity", w odniesieniu do sytuacji, gdzie dokonano ustalenia wartości bitów na wybranych pozycjach klucza szyfrującego (analogicznie „usunięte bity")
5. Definicje wykorzystywanych w tekście pojęć. W tezie rozprawy jest mowa o „efektywności metody badania własności szyfrów" – natomiast w treści rozprawy nie zostało podane jaka jest interpretacja terminów „efektywność" oraz „własności szyfrów". Podobnie, lepsze dla możliwości jednoznacznej interpretacji tekstu pracy byłoby zdefiniowanie również kilku innych pojęć występujących w treści, a kluczowych dla zrealizowanych scenariuszy badawczych, takich jak:
    a. „granica obliczalności",
    b. „margines bezpieczeństwa",
    c. „stabilność pracy",

d. „bezpośrednie kodowanie boolowskie” - ten termin, ponieważ stanowi istotny element rozprawy oraz występuje wielokrotnie w przekroju całej pracy, powinien zostać wprowadzony o wiele wcześniej oraz w sposób bardziej widoczny i jednoznaczny niż to zostało uczynione na stronie 43

e. „bezpośrednia SAT-kryptoanaliza” (i czym się różni od SAT-kryptoanalizy)

6. Prezentacja wyników prac badawczych. W treści rozprawy wielokrotnie przedstawiane są wyniki różnych eksperymentów (np. tabele z czasami obliczeń). W pracy naukowej jednak oprócz samych obserwacji istotna jest analiza wyników. W rozprawie brakuje bardziej szczegółowego podejścia analitycznego do prezentowanych wyników. Nie chodzi tutaj o przedstawienie np. rozszerzonej analizy statystycznej, bo ta w pewnym zakresie została dokonana, ale o stawianie pytań i dociekanie jaka jest istota natury rzeczy, czyli dlaczego otrzymano takie wyniki, a nie inne, czy są one zgodne z oczekiwaniami, czy też odbiegają od tego co wynika np. z samej natury algorytmu, czy też realizowanego scenariusza eksperymentu, itp. „Przechodzenie do porządku dziennego” nad prezentowanymi wynikami może budzić wątpliwości, zwłaszcza kiedy te wyniki nie są do końca jednoznaczne, czy też zgodne z oczekiwaniami, jak to ma miejsce w kilku miejscach ocenianej rozprawy, np. w tabeli 3.12 w kontekście czasów uzyskanych przez SAT-solver Plingeling, czy też tabele 4.3 i 4.4 – dlaczego czas przetwarzania dla SAT-solverów wielowątkowych jest przeważnie wyższy niż dla solverów jednowątkowych.

7. Ocena możliwości uzyskania pozytywnego rezultatu kryptoanalizy. Czy przyjęcie w rozprawie jako wyznacznika i progu bezpieczeństwa analizowanych algorytmów czasów uzyskanych podczas kryptoanalizy realizowanej przy pomocy pojedynczego systemu komputerowego klasy desktop, nie jest znaczącym uproszczeniem? Zwłaszcza w czasach gdy jest stosunkowo łatwy dostęp do znacznych mocy obliczeniowych, w tym również dla „przeciętnego użytkownika”, np. poprzez dostęp do usług w systemach chmurowych?

8. Dostępność kodu źródłowego. Obecnie powszechnie stosowaną praktyką jest udostępnianie kodu, zwłaszcza kodu związanego z realizowaną pracą badawczą. Autorka mogła również udostępnić wytworzony kod, na który wielokrotnie powołuje się w treści pracy, np. w postaci ogólnodostępnego repozytorium. Stanowić mogłoby to dodatkową cenną możliwość upowszechnienia i rozwoju przeprowadzonych prac badawczych.

9. Podsumowanie. W podsumowaniu zabrakło syntetycznego porównania zaproponowanego podejścia do kryptoanalizy algorytmów symetrycznych do innych, wcześniej przywołanych w treści rozprawy podejść. Na przykład, nie przedstawiono w rozprawie analizy porównującej zapotrzebowanie na

pamięć, czy też innych mierzalnych kryteriów oceny pracy algorytmu, dla zaproponowanej metody i metod alternatywnych.

10. Rozważane dalsze prace obejmują m.in. „optymalizację formuł kodowania", podczas gdy jako zaletę zaproponowanej metody podano „brak nadmiarowości", nie jest zatem jasne na czym taka optymalizacja miałaby polegać i czego dotyczyć.

b. Uwagi natury szczegółowej

1. Str. 12, 28,…,137 „w rozsądnym czasie" – w pracy o charakterze naukowym lepiej byłoby się posługiwać jednoznacznymi kryteriami oceny jakości
2. Niezręczność językowa: str. 19 „kładąc w powyższych formułach …"
3. Str. 24 „Tryb ten nie jest uważany za bezpieczny, dostarcza bowiem stronie atakującej wiele par postaci tekst jawny – szyfrogram, szyfrowanych tym samym kluczem co ułatwia kryptoanalizę. Inne tryby pracy eliminują to zagrożenie." – inne tryby też szyfrują bloki przy użyciu ustalonej jednej wartości klucza. Problem z trybem ECB polega na czymś innym.
4. Str. 29. „Projekty kryptograficzne mogą być, w ogólności, bezpieczne albo bezwarunkowo, albo warunkowo." – czy chodzi o kryptosystemy?
5. Str. 31. „Ogólnie, im większy rozmiar danych wejściowych, tym więcej zasobów (czasu, pamięci, procesorów) potrzeba do ich przetwarzania w celu rozwiązania problemu. Złożoność algorytmu w ogóle, w tym na przykład algorytmu kryptograficznego, jest funkcją rozmiaru danych wejściowych." – a czy nie może to być funkcja stała?
6. Str. 32. „Jednak stopień tego wielomianu jest zbyt duży, by dla liczb wielkości interesującej obecnie w praktycznych obliczeniach algorytm ten dawał odpowiedź w czasie praktycznie potrzebnym." – to zdanie powinno być inaczej sformułowane, aby mogło być dobrze zrozumiane.
7. Str. 41 „Stosowane tutaj techniki SAT okazały się wysoce efektywne mimo wysokich parametrów badanych systemów i modelujących ich pracę formuł" - co to są wysokie parametry?
8. Str. 43 „Zaletą naszej metody jest kodowanie każdego bitu bezpośrednio podczas pracy algorytmu, bez nadmiarowości pod względem rozmiaru formuły kodowania." – na czym polega nadmiarowość/brak nadmiarowości w tym kodowaniu? Czy jest w literaturze dokładna analiza zalet tego podejścia?
9. Str. 57 „W miejsce funkcji F wstawiono funkcję XOR." – czy to założenie jest obowiązujące dla całego przedstawionego dalej procesu kryptoanalizy?

10. Str. 63 „oraz sumę użytych do opisu bitów liczb całkowitych” – chodzi raczej o liczbę użytych liczb całkowitych.
11. Str. 65 „Od tej pory uważa się za bezpieczne używanie tylko modyfikacji algorytmu DES o nazwie Triple DES. Obecnie można również zastosować inne, mocne modyfikacje DES.” – obecnie nie zaleca się stosowania żadnych wariantów algorytmu DES.
12. Str. 67 podpisy rysunków 3.12 oraz 3.13 są identyczne, co może być mylące. W treści pracy można odszukać szerszy opis tego co jest na rysunkach, ale podpis jest po to, aby informować, a w tym przypadku tracimy jednoznaczność przekazu.
13. Str. 84 „Ostatnim przekształceniem, który jest użyty w szyfrze Salsa20, jest funkcja szyfrująca Salsa20” – literówka
14. Str. 146 „Planowane jest również użycie do obliczeń maszyn wieloprocesorowych oraz architektur dedykowanych.” – o jakie architektury chodzi i czy wykorzystany w eksperymentach system czterordzeniowy nie jest również przykładem współczesnej architektury wieloprocesorowej?
15. Dodatek B, mógłby z powodzeniem zostać zawarty w tabeli 5.2.

5. Konkluzja

Powyższe uwagi krytyczne nie umniejszają wartości merytorycznej pracy, która stanowi oryginalny wkład Autorki w zagadnienia związane z kryptoanalizą współczesnych algorytmów kryptograficznych. Autorka skupiła się na ocenie możliwości wykorzystania SAT-solverów do przeprowadzenia ataków z wybranym tekstem jawnym. Autorka wykazała się w recenzowanej rozprawie właściwie stosowanym i zaawansowanym aparatem matematycznym oraz bardzo dobrą znajomością aktualnej problematyki z zakresu analizy bezpieczeństwa współczesnych algorytmów kryptograficznych, w tym symetrycznych algorytmów szyfrujących. Dla przedstawionych zagadnień został sformułowany szereg interesujących i użytecznych metod analizy, w tym przedstawione zostały wyniki badań symulacyjnych uwzględniających różne warianty realizowanego ataku oraz różny sposób konfiguracji badanych algorytmów.

Recenzowana rozprawa przedstawia rozwiązanie ważnego i oryginalnego problemu wzbogacając wiedzę dotyczącą kryptoanalizy algorytmów szyfrujących.

Przedstawione w punkcie 4. niniejszej recenzji uwagi nie mają wpływu na jednoznaczne, pozytywną ocenę przedłożonej rozprawy.

Biorąc powyższe pod uwagę stwierdzam, że praca mgr Sylwii Stachowiak pt. „SAT-kryptoanaliza wybranych algorytmów kryptografii symetrycznej” spełnia wymagania stawiane rozprawom doktorskim w świetle stosownej ustawy stopniach naukowych i tytule naukowym. Wnoszę o jej przyjęcie i dopuszczenie do jej publicznej obrony.

Grzegorz Kołaczek