dr hab. Marek Klonowski, prof. nadzw. PWr
Wydział Podstawowych Problemów Techniki
Katedra Informatyki
++48 692 842 272
Marek.Klonowski@pwr.edu.pl

Wrocław, 29.04.2017

Recenzja rozprawy doktorskiej Pana Przemysława Sokołowskiego

Niniejszym wnioskuję o

1. **przyjęcie rozprawy i dopuszczenie Pana Przemysława Sokołowskiego do dalszych etapów przewodu doktorskiego;**
2. **wyróżnienie rozprawy.**

Tematyka i struktura pracy

W rozprawie Autor zajmuje się głównie bezpieczeństwem kryptograficznych funkcji haszujących, które są podstawą niemal wszystkich powszechnie używanych protokołów kryptograficznych. Wyniki dotyczą przede wszystkim ataków[1] opartych o rozwinięcia kryptoanalizy różnicowej - przedstawione są słabości niektórych algorytmów, ze szczególnym uwzględnieniem kilku kandydatów na standard SHA-3 w zakończonym już konkursie NIST. Poza atakami na konkretne schematy, w rozprawie znajdziemy też istotne wyniki o charakterze ogólnym np. rozwinięcie metod wywodzących się z kryptoanalizy różnicowej (kryptoanalizy rotacyjnej).

Rozpraw składa się z siedmiu rozdziałów oraz apendyksu.

Rozdział pierwszy to wprowadzenie, które zawiera intuicje i nieformalne opisy dotyczące podstawowych pojęć związanych z funkcjami haszującymi, ich taksonomię, zastosowania oraz wybrane mechanizmy ataków na nie. Opisane zostały też metody wypracowywania standardów dotyczących funkcji haszujących (w tym konkurs NIST z roku 2008, który stanowił pośrednią motywację dla wielu badań prowadzonych przez Autora). W rozdziale tym przedstawiona została w czytelny sposób także struktura pracy.

[1] Nieco nadużywając terminologii będę używał określenia *atak*, chociaż bardziej właściwie byłoby *wykazanie słabości*.

Rozdział drugi formalizuje podstawowe pojęcia wspomniane na początku rozprawy. Dobre wyrażenie robi bardzo rzetelny opis generycznych konstrukcji funkcji haszujących (rozdział 2.2). Szczegółowo opisane są też różne tryby wykorzystywania protokołów szyfrowania do konstrukcji takich funkcji.

Opisy, chociaż w niektórych przypadkach nie w pełni formalne, są precyzyjne i zupełnie satysfakcjonujące. Nie odbiegają od standardów stosowanych w klasycznej literaturze przedmiotu. Rozdział napisany jest czytelnie i w sposób wyczerpujący. Jedyne czego mógłby więcej oczekiwać czytelnik to nieco szerszy opis tzw. konstrukcji gąbkowej (*sponge construction*).

Rozdział trzeci oparty został o publikację z konferencji ICISC 2011 i zawiera oryginalne rozważania, których współautorem jest Doktorant. Jest to bardzo dobra międzynarodowa konferencja z zakresu bezpieczeństwa informacji o uznanej i ustabilizowanej renomie. Rozdział przedstawia różne słabości szyfrów blokowych wykorzystywanych jako funkcje haszujące. Przedstawione zostają zarówno ataki na konkretne protokoły, jak też konstrukcje ogólne, które mogą zostać użyte do szerokiej klasy szyfrów blokowych. Autor w sposób systematyczny i klarowny przedstawia uzyskane rezultaty uwzględniając bardzo liczne modele działania adwersarza i tryby analizowanych protokołów (rozdział 3.1).

Uwagę zwraca lemat 3.1, o ogólnym znaczeniu dla analizy bezpieczeństwa w modelu ROM. W rozdziale 3.3 wykazane są słabości zarówno konkretnych, stosunkowo nowych protokołów jak i generycznego schematu Feistela.

O ile w całości rozdział jest dobrze napisany, miałem problemy ze zrozumieniem 3.3.2 oraz ostatniego akapitu podrozdziału 3.3.3 (dotyczącego uogólnienia ataku na szyfr Feistela). Bardzo dobre wrażenie robi wykazanie ogromnych różnic poziomu bezpieczeństwa pomiędzy modelami ze znanym i wybieranym kluczem.

Autor podkreśla, chyba słusznie, nowatorski charakter zaprezentowanego ataku na protokół `SAFER++`.

Rozdział czwarty prezentuje wyniki z konferencji FSE 2012. Zaznaczmy, że FSE jest jedną z najbardziej liczących się światowych konferencji dotyczących szyfrów symetrycznych, gdzie zaprezentowane zostały dotąd liczne, bardzo ważne wyniki, które przyczyniły się do ukształtowania współczesnej kryptografii - zarówno pod względem teoretycznym, jak i praktycznym.

W rozdziale Autor zajmuje się algorytmem blokowym `IDEA`, który mimo swoich niemal 30 lat, nadal jest używany. Przedmiotem głównej analizy bezpieczeństwa jest wykorzystanie algorytmu `IDEA` jako funkcji haszującej (z ustalonym, znanym kluczem). Należy podkreślić, że wyniki są bardzo liczne i dotyczą wielu trybów wykorzystania algorytmu `IDEA`. Siła ataków jest różna - od pokazania całkowitej kompromitacji schematu przy pewnych kluczach (podanie przykładu kolizji *expressis verbis*) do wskazania słabości przy ograniczonej liczbie rund, nie stanowiących jednak oczywistego zagrożenia przy realistycznych założeniach dotyczących działania adwersarza.

Ogólnie, Autor w rozdziale przedstawił bardzo wiele ciekawych pomysłów. Dodajmy też, że były rozważane różne typy kolizji. W rezultacie przebadanych zostało bardzo wiele problemów odpowiadających rzeczywistym scenariuszom. Innymi słowy, badania były szerokie i dobrze umotywowane.

Rozdział piąty przedstawia szereg rozszerzeń tzw. analizy rotacyjnej aby następnie zastosować je do wykazania słabości wybranych protokołów, które były rozpatrywane we wspomnianym konkursie NIST. Warto zaznaczyć, że analizowane protokoły to funkcje, które znalazły się w drugiej rundzie konkursu, a więc blisko finału.

Rozdział ten wydaje się technicznie najbardziej zaawansowany. Co więcej, zawiera on bardzo liczne wyniki, z których niemal każdy jest znaczący. Podkreślić trzeba, że ataki na kolejne protokoły nie są schematyczne a wymagają pomysłowości.

Zrozumienie tego rozdziału jest w mojej ocenie trudniejsze niż pozostałych i wymaga szczególnie intensywnego korzystania z innych źródeł. Jest to jednak akceptowalne, bo przy takiej liczbie zaawansowanych technicznie rezultatów trudno oczekiwać kompletności wszystkich rozważań. W szczególności dokładnego prezentowania atakowanych protokołów.

Rozdział szósty przedstawia wyniki z konferencji FSE 2015, które dotyczą kryptoanalizy rotacyjnej. Autor wskazuje błędy metodologiczne w założeniach poczynionych we wcześniejszej pracy z CT-RSA 2014 (częściowo współautorów Doktoranta); przedstawia skorygowane mechanizmy, które potem wykorzystuje do wykazania słabości dwóch protokołów kryptograficznych. Sama kryptoanaliza tych dwóch protokołów (`BLAKE2` oraz `Skein`) jest nietrywialna i nie ogranicza się jedynie do zastosowania schematu z pierwszej części rozdziału. Zauważmy, że drugi z tych protokołów to finalista konkursu NIST.

Ostatni rozdział stanowi krótkie podsumowanie rozprawy. Przedstawia także otwarte problemy oraz konsekwencje wyników rozprawy dla bezpieczeństwa wybranych protokołów kryptograficznych.

Krótki apendyks zawiera kilka uzupełnień treści zasadniczej części pracy. Przyznam, że ich wybór nie jest do końca zrozumiały.

Forma

Rozprawa jest napisana w języku angielskim - w mojej ocenie strona językowa jest bez zarzutu. Z trudem dostrzec można bardzo rzadkie błędy językowe. Drobne usterki edycyjne są bez znaczenia dla oceny rozprawy, czy nawet komfortu czytelnika. Ogólnie, trzeba stwierdzić, że rozprawa napisana jest w sposób niezwykle staranny.

Dysertacja jest przygotowana, można rzec, w „klasyczny" sposób - jest to szeroki opis wyników, wraz z obszernym tłem literaturowym, rzetelnym opisem powiązanych rezultatów oraz prezentacją (w większości) kompletnych rozumowań. Jest to o tyle istotne, że rozprawa ta swą formą wyróżnia się na tle obecnie spotykanych dysertacji będących w istocie bardzo skąpym „autoreferatem" z dołączonym zestawem publikacji[2]. Opis pełny, jaki zaprezentował Doktorant, wymagał oczywiście dużo większego wysiłku i czasu, co bardzo doceniam.

Struktura pracy jest dobra. Autor wprowadza w tematykę a jednocześnie unika powtarzania zbędnych, powszechnie znanych informacji. W niektórych częściach, opisanych

[2]Same publikacje zaś często mają niespójną notację a wyniki ze względu na limit stron, mają istotne luki w rozumowaniach i wady prezentacyjne.

niżej, rozumowania nie są w pełni kompletne a ich pełne przyswojenie wymaga sięgania do innych prac. Jest to jednak akceptowalne, biorąc pod uwagę ogrom wyników, liczbę protokołów do których Pan Sokołowski się odnosi oraz techniczne zaawansowanie części z nich.

Ogólnie, ocena strony formalnej rozprawy jest bardzo wysoka.

Ocena wyników rozprawy

Oceniając wyniki rozprawy trzeba zauważyć przede wszystkim, że jest ich bardzo dużo. Każdy z rozdziałów 3 – 6 prezentuje pewne ogólne rozwinięcia metod kryptoanalitycznych, które potem są stosowane do różnych protokołów.

W zdecydowanej większości zaprezentowane wyniki są istotnie różne. Czyli **nie** sprowadzają się do użycia znanego schematu wiele razy. Nawet w przypadku rodziny ataków pokazywanych w tym samym rozdziale przeciw różnym protokołom, często wymagane było wprowadzanie pomysłowych modyfikacji, dla osiągnięcia celu (np. w rozdziale 5).

Nie wszystkie definicje są w pełni formalne, jednak wszelkie rozumowania są precyzyjne. Co więcej, zaprezentowany poziom formalizmu jest w pełni zgodny z tym jaki stosuje się w badaniach kryptologicznych prowadzonych w najlepszych światowych ośrodkach. Bez wątpienia, klasyfikacja wyników do dziedziny *informatyka* w ramach **nauk matematycznych** jest prawidłowa. Dobrze, że Autor nie unika, poza formalnym opisem, częstego prezentowania intuicji stojących za omawianymi technikami.

Trudno wskazać, które wyniki są najistotniejsze. Za wielki atut rozprawy uważam to, że zostały w wyraźny sposób rozwinięte i skorygowane metody analizy mechanizmów ARX (rozdział 5 i 6), jak również wykazano słabości wielu konkretnych i powszechnie używanych algorytmów. Wydaje się, że zaprezentowane rezultaty mogą być wykorzystane do analizy bardzo szerokiej klasy szyfrów. Także tych skonstruowanych w przyszłości a nawet w jakiś sposób wpłynąć na sposób ich konstrukcji.

Nadmieńmy jeszcze, że wykazanie istotnych słabości konstrukcji będących finalistami (czy nawet uczestnikami drugiej rundy) konkursu NIST jest **znaczącym osiągnięciem**. Konkurs ten był z pewnością jednym z najważniejszych wydarzeń kryptografii światowej ostatnich lat a przez to stał się obiektem zainteresowania najlepszych badaczy na świecie.

Pan Sokołowski wykazał też jak bardzo różnią się modele bezpieczeństwa tych samych mechanizmów przy różnym ich wykorzystaniu. Jest to temat szeroko omawiany w literaturze, jednak nie spotkałem się wcześniej z tak precyzyjnym ujęciem tego tematu.

Na krótkie omówienie zasługuje też metodologia. Kryptoanaliza wymaga bardzo głębokiego zrozumienia badanego problemu i ogromnej precyzji rozumowania, czyli „kontrolowania" pojedynczych bitów w wielorundowej procedurze wykorzystującej dziesiątki podprocedur. Specyfika badań kryptoanalitycznych wymaga dostosowania się do bardzo silnych ograniczeń w podejściu badawczym, które eliminują pewne narzędzia. Jako przykład można podać tu obliczanie wartości funkcji $N_i(i,t)$ (rozdział 5.1), gdzie wykonywanie żmudnych obliczeń wydaje się bezcelowe, jeśli można je bardzo dokładnie przybliżyć choćby za pomocą klasycznych metod kombinatoryki analitycznej. Trzeba jednak wziąć pod uwagę, że wymagana jest **bardzo precyzyjna** analiza dla bardzo małych

wartości argumentów, stąd stosowanie jakichkolwiek metod asymptotycznych jest bezskuteczne. Warto także wspomnieć, że Autor wykazał się dużą pomysłowością oraz umiejętnościami technicznymi stosując szeroki wachlarz metod typowych dla kryptoanalizy (w tym stosując eksperymenty numeryczne).

Warto nadmienić, że wyniki wchodzące w skład rozprawy były prezentowane na dwóch konferencjach FSE (*Fast Software Encryption*) oraz jednej ICISC (*Information Security and Cryptology*). Są to fora o bardzo dobrej międzynarodowej renomie, a samo FSE jest uważane z jedno z najważniejszych światowych wydarzeń dotyczących szyfrowania symetrycznego. Nie trzeba dodawać, że są to konferencje o bardzo dużej selektywności, gdzie recenzentami są czołowi kryptolodzy z wiodących ośrodków a recenzje wnikliwością nie ustępują dobrym czasopismom.

Dodajmy, że prace Doktoranta doczekały się przeszło 50 cytowań wg. GS (h-index 5), co jest wynikiem bardzo dobrym. Szczególnie, że bardziej wnikliwa ich analiza nie wskazuje na istnienie autocytowań a wśród cytujących znaleźć można pierwszoplanowe postacie światowej kryptografii [3].

Zaryzykuję stwierdzenie, że wyniki z rozprawy wpłynęły na dziedzinę i mają konsekwencje dla praktyki (konkurs NIST). Co więcej, znajdują się one w głównym i trudnym nurcie badań kryptograficznych.

Uwagi krytyczne

Nie widzę żadnych poważnych wad rozprawy, które zmuszałyby mnie prosić o jej korektę czy nawet o formalne ustosunkowanie się do wskazanych błędów. Oczekuję co najwyżej przygotowania do ewentualnej dyskusji na dalszych etapach przewodu.

Poniżej wskażę kilka uwag o niedużym znaczeniu dla oceny rozprawy.

1. W rozprawie nie ma formalnego zdefiniowania złożoności ataków zrandomizowanych. Definicja złożoności w jakiejś formie pojawia się dopiero w Lemacie 3.1. Sformułowanie *"complexity 1"* jest nieco niezręczne (s. 32). Ta złożoność jest zazwyczaj rozumiana jako wartość oczekiwana liczby operacji potrzebnych do skutecznego ataku. Chociaż raz jest to dokłada wartość („Brute Force Attack") a raz wielkość asymptotyczna („Birthday paradox") (rozdział 2.5.1).

2. W podrozdziale 5.3.2 niektóre fragmenty pozostają niejasne. Na przykład, pewne wartości prawdopodobieństwa zostały wyznaczone eksperymentalnie (co jest zrozumiałe i uzasadnione). Jednak nie bardzo wiem co znaczy, że zostały wyznaczone heurystycznie (s. 98) ?

3. Czasami Autor nie precyzuje czy mamy ustaloną wartość, czy wartość losową. Zazwyczaj jest to w pełni zrozumiałe z kontekstu, jednak momentami powoduje pewne niejasności. Np. w Lemacie 5.1 mamy *(...) Given n-bit words* $x_1, \ldots, x_k$ *and a positive integer* r *(...)*. Powinno być raczej: *(...) Given n-bit* ***radnom*** *words* $x_1, \ldots, x_k$ *and a* ***fixed*** *positive integer* r *(...)*. Podobnie w Lematach 2.1- 2.4 dobrze byłoby wskazać po czym liczone jest prawdopodobieństwo.

[3] Np. Barta Preneela

4. Na początku strony 96 nie jest dla mnie jasne, skąd mamy wartość 2^{-61} jako prawdopodobieństwa spełniania układu równań ? Nie jest to iloczyn prawdopodobieństw spełniania każdego z równań. Poza tym spełnialność kolejnych równań układu wydaje się być stochastycznie zależna (Te same fragmenty klucza są używane).

5. Poza pojedynczymi pozycjami, bibliografia kończy się na roku 2012. Jest to być może uzasadnione tym, że w pewnym sensie rozprawa koncentruje się na wynikach dotyczących konkursu NISTu (zakończonego właśnie w 2012). Szerszy opis nowszych wyników byłby jednak dalszym wzmocnieniem jakości rozprawy.

6. Mimo ogólnie bardzo dobrej prezentacji, pewne (nieliczne) fragmenty są opisane zbyt lakonicznie - np. w 4.5 trudno zrozumieć schemat ataków na niektóre protokoły.

Pozwolę sobie jeszcze na kilka uwag, już nawet nie o małym, ale wręcz zaniedbywalnym znaczeniu.

- s. 3, linia 9: w dwóch miejscach jest 0 a powinno być $\{0\}$
- s.19, ostatnia linia: jest m powinno być m_i
- s.37, linia 10 - r powinien być jako indeks dolny (w dwóch miejscach)
- s.94 wzór na W_j: powinno być $\star$ zamiast $*$.
- Stosowanie operatora $\lll$ jest czasami nieczytelne w ciągu innych operacji. Być może nawiasy poprawiłyby przekaz.
- Funkcja rot (s. 91) dla jasności powinna być zdefiniowana w 5.2.2.
- Mimo ogólnie bardzo wysokiej oceny prezentacji, pewne opisy są zbyt skrótowe np. 1.4.1 atak urodzinowy oraz *meet-in-the-middle*.
- W kilku miejscach zaokrąglenia w wykładnikach przy oszacowaniu prawdopodobieństw wydają się niekonsekwentne.
- W samej bibliografii brakuje w sporej części podania konferencji, na której dana praca była prezentowana. Sam numer tomu LNCS niewiele mówi i nie pozwala szybko umiejscowić konkretnych wyników.

Konkluzja

Nie mam najmniejszych wątpliwości, że rozprawa Pana Przemysława Sokołowskiego zawiera świetne wyniki i jest dobrze napisana. Nie mam też wątpliwości, że wszelkie ustawowe i zwyczajowe wymagania spełnione zostały z nadmiarem.

Wnioskuję o wyróżnienie rozprawy. Wyniki zaprezentowane w dysertacji są technicznie trudne i pomysłowe. Co więcej, jest ich dużo. Praca oparta jest na bardzo dobrych publikacjach, prezentowanych na czołowych światowych konferencjach o dużej

selektywności. W końcu, doktorat jest dobrze napisany. Ta praca mogłaby być podstawą do nadania stopnia w każdym z najlepszych światowych uniwersytetów, gdzie prowadzi się badania dotyczące kryptografii. Czego chcieć więcej ?