Kraków, 29ty czerwca, 2024r.

prof. dr hab. inż. Piotr Faliszewski
Wydział Informatyki
Akademia Górniczo-Hutnicza
im. Stanisława Staszica w Krakowie
faliszew@agh.edu.pl

Recenzja rozprawy doktorskiej mgr. Tomasza Lizureja

Zgodnie z zaproszeniem Biura Rad Naukowych Uniwersytetu Warszawskiego, przedstawiam poniżej moją recenzję rozprawy doktorskiej Tomasza Lizureja. Tytuł rozprawy to "On Security of Systems Built on Blockchains" i została ona przygotowana przez autora pod kierunkiem prof. Stefana Dziembowskiego (promotora) oraz dr. Tomasz Michalaka (promotora pomocniczego). Rozprawa mieści się w dziedzinie nauk ścisłych i przyrodniczych, w dyscyplinie informatyka.

Wstęp

Rozprawa dotyczy bezpieczeństwa aplikacji opartych o technologię blockchain. W szczególności Tomasz Lizurej zajmuje się następującymi kwestiami:

1. Rozważa systemy oceny wiarygodności uczestników transakcji. W tym celu rozważa algorytm FGA (*ang.* Fairness-Goodness Algorithm), charakteryzuje go aksjomatycznie, oraz analizuje jak łatwe jest wpływanie na jego wyniki.

2. Analizuje pojęcie *dowodu indywidualnej wiedzy* (*ang.* proof of individual knowledge, PIK), które np. pozwala przekonać organizatora elektronicznych wyborów toczących się na blockchainie, że dana osoba faktycznie posiada całość informacji ją identyfikujących, a nie wynajęła ich komuś, kto chce nieuczciwie zagłosować za nią.

3. Tworzy kompilator obwodów cyfrowych, który potrafi zamienić dany obwód na równoważny, który jednak jest uzupełniony o fragment weryfikujący poprawność jego działania.

O ile prace te faktycznie dotyczą kwestii bezpieczeństwa aplikacji opartych o technologię blockchain, to są od siebie dość dalekie tematycznie i rozprawa nie prezentuje spójnego

wywodu prowadzącego do jednego celu. Zamiast tego otrzymujemy trzy sklejone artykuły, połączone wstępem i podsumowaniem. Z jednej strony jest to rozczarowujące, a z drugiej nie przesłania podstawowego celu studiów doktorskich—wykształcenia nowego naukowca, zdolnego samodzielnie zrealizować mniejszy lub większy projekt badawczy.

Doktorat zotał przygotowany w języku angielskim i pod względem językowym jest napisany bardzo dobrze. Niestety w wielu miejscach widoczny jest albo pośpiech, ale niestaranność doktoranta. Na przykład rysunek 2.12 (Figure 2.12) pojawia się na stronie 22, ale odniesienie do niego jest dużo dalej, na stronie 29. Nic nie stało na przeszkodzie, żeby ten rysunek był umieszczony w rozprawie tam, gdzie jest omawiany. Co gorsza, przedstawione na nim wykresy mają opisy sporządzone na tyle małą czcionką, że trudne je odczytać. Przedstawione trzy wykresy nie zawierają także informacji jakich zbiorów danych dotyczą (a raczej, nie mają tej informacji danej jasno, bo nazwy zbiorów danych są w podpisie). Zaskakuje także, że w podsumowaniu doktorant pisze, że wprowadził pojęcie *Secret Sharing with Snitching* ale te słowa nie pojawiają się gdzie indziej w rozprawie (zakładam, że kiedyś były i zostały usunięte, ale doktorant nie uaktualnił podsumowania—co właśnie odbieram jako niestaranność). Podobnie irytuje pewna nonszalancja w bibliografii (np. pozycja [74] jest przedstawiona jako pochodząca z konferencji "Advances in Cryptology - EUROCRYPT" a pozycja [75] jako opublikowana w "EUROCRYPT"; biorąc pod uwagę, że obie pozycje odnoszą się do tej samej pracy, to oczekiwałbym tu jednak pewnej spójności). Zaskakujące jest także, że definicje numerowane są względem sekcji (np. mamy definicję 4.5.1) ale twierdzenia są numerowane globalnie (np. mamy twierdzenie 21). Kilka innych przykładów niestaranności zamieszczę w dalszej części recenzji.

Praca Pana Tomasza Lizureja opiera się na trzech publikacjach przedstawionych na konferencjach AAAI-2023, CRYPTO-2023, oraz TCC-2023:

> T. Lizurej, T. Michalak, and S. Dziembowski, On Manipulating Weight Predictions in Signed Weighted Networks, AAAI-2023.
>
> S. Dziembowski, S. Faust, and T. Lizurej, Individual Cryptography, CRYPTO-2023.
>
> M. A. Baig, S. Chakraborty, S. Dziembowski, M. Gałązka, T. Lizurej, and K. Pietrzak, Efficiently Testable Circuits without Conductivity, TCC-2023

Na tle wielu doktoratów jest to liczba niewielka, ale bibliometria nie stanowi kryterium oceny. Podkreślić należy, że są to czołowe konferencje oceniane na poziomie A* (pierwsze dwie) lub A (ostatnia). Natomiast wielkim zaskoczeniem była dla mnie informacja, że główny wkład techniczny w pracy "Individual Cryptography" pochodził od pozostałych autorów a wkad doktoranta dotyczył jedynie formalnego zapisu dowodów (tak rozumiem informację umieszczoną na stronie 7). W moim odczuciu nie uzasadnia to umieszczenia wyników z tej pracy w rozprawie. Skoro te wyniki jednak w rozprawie się znalazły, to zakładam że jest to kwestia niezgrabnego sformułowania w tekście rozprawy i doktorant jednak aktywnie uczestniczył w pracach nad wynikami technicznymi.

Na główną treść rozprawy składają się rozdziały 2-4, bezpośrednio powiązane z trzema opublikowanymi przez doktoranta pracami, których zawartość omówię poniżej (doktorant w sumie opublikował sześć prac na konferencjach; mam tu na myśli te trzy, które stanowią części rozprawy).

Manipulacja wynikami algorytmu FGA

W rozdziale 2 doktorant zajmuje się algorytmem FGA, który otrzymuje na wejściu ważony graf G, gdzie wierzchołki to uczestnicy danej aplikacji blockchainowej a krawędzie—etykietowane wartościami z przedziału od -1 do 1—informują jak uczestnicy nawzajem się oceniają (choć nie każdy uczestnik musi oceniać każdego innego; robi to wyłącznie wtedy, gdy weszli w jakąś formę interakcji). Algorytm FGA przypisuje każdemu uczestnikowi dwie wartości, *fairness*, opisującą jak uczciwy jest dany uczestnik w ocenie innych, oraz *goodness*, opisującą jak wysoko jest sam oceniany.

Doktorant przedstawił aksjomatyzację algorytmu FGA, czyli udowodnił że jeśli wartości *fairness* i *goodness* spełniają pewne dość naturalne własności, to musiały być obliczone zgodnie z algorytmem FGA. Jest to niewątpliwie ciekawy wynik, który pozwala lepiej zrozumieć algorytm FGA, choć trochę rozczarowuje fakt, że aksjomatów jest aż jedenaści (choć część z nich jest dość podobna, tylko dotyczą albo wartości fairness, albo goodness). Dowód charakteryzacji jest dość krótki, ale elegancki.

Doktorant rozważa także trudność obliczeniową wpływu na wynik algorytmu FGA przez dodawanie nowych krawędzi lub zmianę wag niektórych istniejących. W tej części doktorant pokazuje, że jego problemy są NP-zupełne oraz W[2]-trudne dla parametryzacji rozmiarem ataku. Dowody i wyniki są oczekiwane i standardowe, co jednak nie jest wadą samą w sobie—należy takie wyniki uzyskiwać by było widomo, że dalsza analiza bezpieczeństwa systemu musi się opierać na heurystykach (i/lub algorytmach wykładniczych) a nie na dokładnych algorytmach wielomianowych.

Na koniec doktorant przedstawia wyniki eksperymentów, które wskazują, że ataki na algorytm FGA mogą nie być zbyt skuteczne w praktyce. Spodziewałem się, że wyniki będą nieprzekonujące, ale tu doktorant zaskoczył mnie pozytywne. Wykorzystuje rozsądne zbiory danych i uzyskane wyniki wskazują, że ataki "nie wprost" (tj. ataki, których celem jest wpłynięcie na dany węzeł, ale bez dodawania interakcji z nim) nie są skuteczne. Doktorant udowodnił także wyniki dotyczące wpływu na wyniki FGA w sieciach, gdzie każdy wierzchołek ma co najmniej zadany stopień. Wyniki są ciekawe pod względem teoretycznym, ale nie było dla mnie w pełni przekonujące, czy mają wartość praktyczną i czemu takie sieci są ciekawe.

Rozdział 2 został napisany dość solidnie, ale jednak zawiera szereg drobnych usterek. Poniżej przedstawiam kilka przykładów:

Str. 8, "... an edge weight prediction method for signed weighted networks (by [84])."

Traktowanie cytowań jako rzeczowniki określające cytowaną pracę jest błędem stylistycznym (choć bardzo powszechnym; zapewne kiedyś stanie się to normą). Błąd ten powtarza się w całej rozprawie.

Str. 10, pierwsza linijka sekcji 2.2: Doktorant wprowadza graf $G = (V, E, W)$, ale następnie zamiast funkcji W używa funkcji ω.

Podpis rysunku 2.8 zawiera pojedynczą liczbę przeniesioną do drugiej linii, co jest nieestetyczne.

Str. 25, trzy linijki nad definicją 2.5.1 mamy "set set".

Tw. 9: Byłoby dobrze napisać wprost czy zmiana wartości jest addytywna czy multiplikatywna.

Str. 26: W pierwszej linijce dowodu Tw. 9 zdaje się pojawiać nowa notacja, $\mathcal{I}(t)$.

Str. 28: "quality" zamiast "equality".

Str. 29: "Alorithm" zamiast "Algorithm".

Powyższe usterki to oczywiście drobiazgi, na ogół nie warte wspominania. Przywołuję je jednak jako przykład pewnej niestaranności w pisaniu tekstu (jednak bez wpływu na merytoryczną wartość pracy).

Dowód indywidualnej wiedzy

W rozdziale 3 doktorant przestawia koncepcję dowodu wiedzy indywidualnej (*ang.* proof of individual knowledge). Pomysł polega na tym, by posiadacz danego sekretu był w stanie wykonać na nim obliczenia na tyle szybko, by było oczywiste, że obliczenie nie mogło być wykonane w systemie rozproszonym, którego uczestnicy nie ufają sobie nawzajem i którzy posiadają jedynie fragmenty tego sekretu. Jest to niewątpliwie najciekawszy rozdział rozprawy, choć nie jest jasny wkład doktoranta w jego techniczną część.

Moim największym zarzutem wobec rozdziału 3 jest sposób jego przedstawienia, co jednak może wynikać z faktu, że nie zajmuję się na codzień kryptografią. Trudno było mi zrozumieć przedstawiany wynik, gdy był opisany na coraz bardziej szczegółowym poziomie, ale bez jasnego przedstawienia jego całej struktury. Protokół dowodu wiedzy indywidualnej, jako całość, pojawia się dopiero na str. 45, podczas gdy sam rozdział zaczyna się na str. 33. Poza tym rozdział 3 zawiera też podobne usterki jak pozostały tekst (np. odniesienia do rysunków wiele stron od miejsc, gdzie są umieszczone, tytuł podrozdziału "Proof of Thm. 14" zamiast "Proof of Theorem 14" itp.).

Od strony merytorycznej wyniki wydają się bardzo ciekawe i dalece nietrywialne, ale jako że nie zajmuję się kryptografią, wartość mojej opinii w tym zakresie jest ograniczona.

Dowody wiedzy indywidualnej mogą być przydatne przy konstrukcji rozproszonych wyborów, gdyż pozwalają istotnie utrudnić proces sprzedawania głosów. Doktorant opisuje przykład jak taki system głosowania mógłby działać, ale ogranicza się do elementarnych wyborów, w których głosujemy albo na kandydata 0 albo na kandydata 1. Jestem ciekawy, czy podobną technikę możnaby użyć w przypadku, w którym głosy mają bardziej zawiłą strukturę (np., są porządkami nad zbiorem kandydatów; choć wówczas kupowanie głosów jest łatwe z innych powodów).

Kompilator obwodów cyfrowych

W rozdziale 4 doktorant zajmuje się kwestią testowania obwodów cyfrowych. Zadanie polega na tym, że mając dany obwód, chcemy go zmodyfikować tak, by na podstawie jego działania dla możliwie małego zbioru testowych wejść dało się ocenić, czy został zmodyfikowany. Kluczową nowością wprowadzaną przez doktoranta do literatury jest fakt, że jego algorytm nie wymaga założenia o przewodzeniu (*ang.* conductivity). Innymi słowy możliwe jest modyfikowanie wartości na połączeniu pewnej bramki logicznej z bramką A, bez jednoczesnego modyfikowana wartości na połączeniu tej samej bramki z bramką B.

Wyniki przedstawione w rozdziale 4 są ciekawe i zdecydowanie wymagały wysiłku i głębokiego zrozumienia problemu. Istotnym utrudnieniem w czytaniu tego rozdziału była dla mnie niejasna zależność między jego treścią a cytowanymi pracami (pozycje 8 i 9 w bibliografii). Dopiero w podsumowaniu (rozdział 5) znalazłem napisane wprost, w sposób nie budzący wątpliwości, że wyniki z rozdziału 5 poprawiają te wcześniejsze. Byłoby też dużo lepiej, gdyby doktorant wprost odnosił się do pracy "Efficiently Testable Circuits without Conductivity" [9] jako do pracy, której jest współautorem, a nie jak do pracy stworzonej niezależnie od niego (w tekście pracy czasem stosowane jest jedno podejście, a czasem drugie).

Podsumowanie

Z jednej strony, zapreznetowana rozprawa nie jest idealna. Przeszkadza brak spójności i bardziej przekonującego połączenia między rozdziałami. Waga uzyskanych wyników—nie licząc rozdziału 3, gdzie wkład jest niejasny—także wydaje się na granicy tego, czego oczekuje się od rozprawy doktorskiej. Mimo tych zastrzeżeń uznaję, że przedstawiona rozprawa uprawnia do nadania mgr. Tomaszowi Lizurejowi stopnia doktora w dziedzinie nauk ścisłych i przyrodniczych, w dyscyplinie informatyka.