Gdańsk, 29 stycznia 2024

Dr hab. inż. Rafał Leszczyna
Politechnika Gdańska
Wydział Zarządzania i Ekonomii
rle@zie.pg.edu.pl

Recenzja rozprawy doktorskiej

kpt. mgr. inż. Mariusza Cezarego Szarka

pt. „Budowanie mechanizmu obrony przed dedykowanymi kampaniami phishingowymi"

Promotor dr hab. inż. Ryszard Antkiewicz, prof. Wojskowej Akademii Technicznej

1. Wprowadzenie

Niniejsza recenzja rozprawy doktorskiej, której autorem jest kpt. mgr. inż. Mariusz Cezary Szarek, została wykonana w oparciu o uchwałę Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja Wojskowej Akademii Technicznej im. Jarosława Dąbrowskiego (nr 55/RDN ITiT/2023) z dnia 12 grudnia 2023 r. podpisaną przez Przewodniczącego Rady dr hab. inż. Zbigniewa Tarapatę, prof. WAT. Uchwała ta wskazuje mnie jako recenzenta w komisji doktorskiej w postępowaniu w sprawie nadania stopnia doktora kpt. mgr. inż. Mariuszowi Cezaremu Szarkowi.

Promotorem niniejszej rozprawy jest dr hab. inż. Ryszard Antkiewicz, prof. Wojskowej Akademii Technicznej w Warszawie.

2. Problem naukowy (teza) rozprawy

Phishing to cyberatak mający na celu wyłudzenie od użytkownika wrażliwych informacji, takich jak np. dane uwierzytelniające, dane osobowe, czy dane kart kredytowych, z wykorzystaniem inżynierii społecznej i oszustwa. Wg raportu FBI[1], jest

[1] Federal Bureau of Investigation, 2022 Internet Crime Report, https://www.ic3.gov/Media/PDF/AnnualReport/2022_IC3Report.pdf

to najczęstsze przestępstwo cyfrowe a także główny wektor początkowy innych ataków, takich jak np. *ransomware*. Centrum Zgłoszeń Przestępstw Internetowych (ang. *Internet Crime Complaint Center – IC3*) otrzymało największą liczbę zgłoszeń tego ataku, a związane straty finansowe szacowane są na ponad 52 miliony dolarów. ENISA[2], przywołując statystki APWG, wskazuje, że liczba ataków phishingowych wykazuje trend wzrostowy i notuje rekordy. Co więcej, coraz popularniejsze są ukierunkowane ataki phishingowe, takie jak *whaling* (celujące w zarząd przedsiębiorstw) czy *spear-phishing* (wymierzone w starannie wybrane osoby lub grupy osób). Grupa analityków zagrożeń Google zaobserwowała liczne kampanie spear-phishingowe ukierunkowane na państwa NATO, które najprawdopodobniej mają na celu uzyskanie informacji o wsparciu militarnym dla Ukrainy.

Skuteczne wykrywanie ataków phishingowych jest więc kluczowe zarówno w kontekście gospodarczym, społecznym, czy geopolitycznym. Powstało wiele komercyjnych narzędzi a literatura naukowa opisuje dziesiątki rozwiązań. Stosowane techniki obejmują m.in. czarne i białe listy, ocenę zawartości i strony wizualnej wiadomości, logikę rozmytą, czy uczenie maszynowe.

Analiza metod detekcji przeprowadzona przez Doktoranta wykazała, że cechują się one niedostatecznym dostosowaniem do ewoluujących sposobów atakowania oraz infrastruktury wykorzystywanej do przeprowadzenia ataku (1). Ponadto nie są w stanie rozpoznać, nowych, nieznanych wcześniej typów ataków (2) oraz niedostatecznie uwzględniają fakt stosowania legalnych metod i zaufanych źródeł przez cyberprzestępców (3). Przeanalizowane przez Doktoranta metody niewystarczająco też zapobiegają spersonalizowanym (5) oraz wielopoziomowym atakom (6). Wiążą się również z dużymi nakładami administracyjnymi i koniecznością ciągłego pozyskiwania wiedzy o nowych technikach (4).

W tym kontekście Doktorant zdecydował się na przeprowadzenie badań mających na celu opracowanie nowej metody wykrywania phishingu, która będzie pozbawiona wad i słabości zidentyfikowanych metod. Metoda ta powinna integrować (1) analizę porównawczą wiadomości e-mail, (2) analizę pól wiadomości, (3) powtarzalność schematu, (4) analizę treści wiadomości, (5) identyfikację szantażu, wymuszenia okupu, (6) analizę reputacji domen oraz (7) wykorzystywać metody data mining i uczenia maszynowego do wykrywania nieznanych zależności i schematów.

Teza pracy została sformułowana następująco: *„połączenie (*wskazanych wyżej*) sposobów detekcji pozwoli na wykrywanie wcześniej niestosowanych schematów ataku, używanie nieznanych do tej pory wzorców – co eliminuje wady klasycznych metod*

---

[2] ENISA, ENISA Threat Landscape 2023, https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023

wykrywania". Przy czym „klasycznymi metodami wykrywania" zostały określone w rozprawie „metody, który nie zawierają klasyfikacji, elementów uczenia maszynowego".

Tak postawiona teza rozprawy została zasadniczo jasno sformułowana przez Doktoranta. Pewna niejasność w możliwości interpretacji celów i tezy pracy została opisana w rozdziale 3 niniejszej recenzji.

W swoich badaniach Doktorant podjął aktualną i ważną problematykę o istotnym znaczeniu społecznym, ekonomicznym czy geopolitycznym.

3. Zawartość rozprawy

Przedstawiona do recenzji rozprawa składa się z nienumerowanego wprowadzenia, pięciu głównych rozdziałów, nienumerowanego zakończenia oraz bibliografii obejmującej 100 pozycji. Treść uzupełniają dodatkowo wykaz używanych skrótów, spis ilustracji i tabel oraz 4 dodatki zawierające kolejno: wyniki analizy próbek złośliwego oprogramowania stanowiącego załączniki do próbek badawczych, komponenty środowiska Python wymagane do pracy opracowanego narzędzia, opis przykładowego ataku smishing oraz wartości macierzy pomyłek analizowanych klasyfikatorów. Całość obejmuje 274 strony.

Pierwszy rozdział poświęcony jest atakom phishingowym. Przedstawiono w nim główne trendy wykorzystania technologii informacyjnych i związane z nim narażenie na cyberataki, wskazując, że phishing należy do najczęściej identyfikowanych i stosowanych metod ataku. Opisano modele ataku „Cyber Kill Chain", „Phishing Kill Chain" oraz MITRE ATT&CK. Dość szczegółowo przedstawiono różne typy ataków phishingowych a w tym m.in. popularne e-mail phishing i smishing, czy rzadsze whaling i spear-phishing. Doktorant zidentyfikował tu aż 14 kategorii ataków phishingowych. W kolejnych podrozdziałach opisane zostały główne techniki implementacji phishingu oraz sposoby ukrywania faktu ataku (zaciemniania), odrębne miejsce poświęcając inżynierii społecznej (ang. social engineering), która odgrywa kluczową rolę w phishingu. Przedstawiono też kolejne etapy phishingu oraz kryteria klasyfikacji i oceny skuteczności tego ataku. Poruszono też temat podnoszenia świadomości użytkowników, jako istotnego instrumentu w walce z atakami phishingowymi. Rozdział pierwszy zamyka krótkie podsumowanie podejmowanego problemu badawczego ze wskazaniem celu badań i sformułowaniem tezy rozprawy. Ten teoretyczny rozdział pracy stosunkowo szeroko oraz wieloaspektowo przedstawia tematykę badawczą podjętą przez Doktoranta. Wskazuje na dobre rozpoznanie obszaru przez Doktoranta, w tym identyfikację odpowiedniej literatury. Z drugiej strony można mieć pewne zastrzeżenia dotyczące spójności oraz struktury tego rozdziału. Stosowanie numeracji podrozdziałów czwartego poziomu nie ułatwia czytania i celem jego uniknięcia Doktorant mógł rozważyć np. podzielenie rozdziału pierwszego na kilka równorzędnych części. Także treści początku rozdziału i podrozdziału I.2, czy opisy modeli ataków i zawartości podrozdziałów I.4.1 i I.3.4 posiadają cechy wspólne, które sugerowałyby umieszczenie ich w powiązanych sekcjach.

Rozdział drugi przedstawia główne techniki i metody detekcji phishingu rozpoznane przez Doktoranta w literaturze. Opisano tu wykrywanie w oparciu o listy wykluczeń/dopuszczeń i reguły detekcji, a także techniki z wykorzystaniem algorytmów genetycznych i uczenia maszynowego. Na końcu rozdziału umieszczono listę wad i słabości przedstawionych metod. Zgodnie z celem pracy badawczej Doktoranta, cech tych powinna być pozbawiona metoda opracowana przez niego w toku badań. W kontekście przedstawionych w rozprawie metod bazujących na uczeniu maszynowym, wątpliwość wzbudza wskazana przez Doktoranta druga wada metod detekcji phishingu tj. „Brak rozpoznania nowego, nieznanego wcześniej typu ataku lub niewystępującej techniki." Kwestia ta wymagałaby szerszego wyjaśnienia w rozprawie. Pojawia się też pewna niejasność dotycząca tego, czy metoda zaproponowana przez Doktoranta, ma eliminować słabości wyłącznie klasycznych, czyli nie wykorzystujących uczenia maszynowego, metod detekcji (zgodnie z tezą rozprawy), czy wszystkie wady wskazane w podrozdziale II.7.

Rozdział obejmuje szereg (około 20) pozycji rozpoznanej literatury. Tymczasem literatura dziedzinowa jest znacznie bogatsza. Dlatego podczas badań wskazane byłoby zastosowanie systematycznego podejścia analizy literatury (ang. *systematic literature review*[3]), które daje wysokie szanse pełnej identyfikacji najważniejszych alternatywnych pozycji. Brakuje tu też porównania metody opracowanej przez Doktoranta z większą liczbą technik wykorzystujących uczenie maszynowe. Dałoby one ważny punkt odniesienia dla otrzymanych wyników.

Kompletności z pewnością nie brakuje za to rozdziałowi trzeciemu rozprawy. Szczegółowo przedstawiono tu aż blisko 30 rozpoznanych przez Doktoranta wskaźników występowania ataku phishingowego. Opisom towarzyszą przykłady oraz formuły reguł wykrywania. Ta część pracy stanowi niewątpliwie istotny wkład badawczy Doktoranta i świadczy o dobrej orientacji w dziedzinie problemowej. Zidentyfikowane wskaźniki staną się podstawą do stworzenia wektora cech dla autorskiej metody wykrywania phishingu wykorzystującej uczenie maszynowe.

Metoda ta została opisana w rozdziale czwartym rozprawy. Wskazano tu ostateczną listę 19 cech służących do rozpoznawania wiadomości phishingowych oraz objaśniono poszczególne kroki metody. W rozdziale pojawiają się fragmenty kodu algorytmów oraz diagramy ilustrujące działanie jej komponentów. Na końcu rozdziału przedstawiono 5 typów klasyfikatorów będących kandydatami do wykorzystania w opracowywanej metodzie. Opisy te mają charakter teoretyczny i zastanawia dlaczego nie zostały

[3] Np. [1] Mehran Alidoost Nia, A. Ruiz-Martínez, Systematic literature review on the state of the art and future research work in anonymous communications systems, Computers & Electrical Engineering, Volume 69, 2018, Pages 497-520, ISSN 0045-7906, https://doi.org/10.1016/j.compeleceng.2017.11.027.
[2] Kitchenham, B., Brereton, P.: A systematic review of systematic review process research in software engineering. Information and Software Technology 55(12), 2049 (2013). https://doi.org/10.1016/j.infsof.2013.07.010

umieszczone w początkowej, teoretycznej części rozprawy. Ten fragment pracy nie uchronił się też przed błędem numerowania, nie jest on również odzwierciedlony w spisie treści. Generalnie, poprawa struktury oraz ponowna redakcja rozdziału czwartego zwiększyłaby jego czytelność.

Rozdział piąty poświęcony jest weryfikacji metody w dwóch głównych aspektach:

a) prawidłowości identyfikacji cech w wiadomościach, oraz
b) poprawności klasyfikacji.

Analizę wykrywania cech w wiadomościach przeprowadzono bazując na zbiorze 50 wiadomości. Dla tak dobranej próby algorytm wykazał dobre wartości metryk jakościowych (czułość, precyzja i dokładność) oscylujące wokół 99%. Ocenę klasyfikacji wiadomości wykorzystując uczenie maszynowe przeprowadzono w 6 iteracjach mających na celu m.in. „dostrojenie" algorytmu odczytu pól wiadomości i wykrywania cech, czy odpowiednie przygotowanie zbioru uczącego. W ostatniej iteracji wykorzystano 6 technik równoważenia zbiorów tj. SMOTE, SMOTE-ENN, ADASYN, Random Over Sampler, Random Under Sampler i Tomek Links. Techniki te zostały krótko opisane w rozdziale. Również w tym wypadku zastanawia dlaczego opisy, o charakterze teoretycznym nie zostały umieszczone w początkowej, teoretycznej części rozprawy. Jako metrykę oceny jakości klasyfikatora wybrano trafność (ang. *accuracy*). W toku kolejnych iteracji uzyskano wartość około 85% przy zbiorze uczącym zawierającym prawdopodobnie 981 wiadomości phishingowych, 378 wiadomości typu spam i 84 wiadomości normalnych (przed równoważeniem). Pojawiające się tutaj określenie „prawdopodobnie" wynika z faktu, że nie zostało to dostatecznie jasno wyjaśnione w rozdziale. Na przykład w podrozdziale V.2.1 wspomina się o 1168 wiadomościach phishingowych, by w V.2.3 wskazywać liczbę 966 zwiększaną do 981. Kwestia skupienia się na jednym wskaźniku jakości wymagałaby wyjaśnienia. Pewien niedostatek stanowi brak analiz efektywności i złożoności metody.

Rozdział zamyka część zatytułowana „Wnioski". Doktorant wskazuje tu dodatkowe obserwacje z eksperymentów. Należą do nich uzyskane wartości prawdopodobieństw wykrywanych cech, które mogą posłużyć do dalszych prac w kierunku zmniejszenia wektora uczącego. Jednocześnie zagadnienie błędnej klasyfikacji cechy świadczącej o możliwym szantażu oraz częściowo cechy związanej z błędami językowymi wykazuje związek z wcześniejszym podrozdziałem dotyczącym weryfikacji jakości odczytu pól wiadomości. W tym kontekście zastanawia dlaczego sama weryfikacja przebiegła dla stosunkowo małego zbioru wiadomości, gdy do dyspozycji była ich większa liczba, wykorzystana w drugiej części analiz. Podobnie jak w przypadku rozdziału czwartego, także i tutaj, poprawa struktury oraz ponowna redakcja rozdziału zwiększyłaby jego czytelność. Na przykład w podrozdziale V.2.3 pojawiają się kolejne nienumerowane sekcje. Jest tam również treść dotycząca rozpoznania nowych adresów IP oraz adresów e-mail identyfikowanych jako phishingowe, którą należałoby umieścić w odrębnej sekcji.

W Podsumowaniu wskazano najważniejsze elementy pracy badawczej oraz istotne osiągnięcia autorskie. Wskazano również kilka kierunków dalszych prac nad opracowanym przez Doktoranta rozwiązaniem.

Poza uwagami dotyczącymi struktury rozdziałów, rozprawa nie ustrzegła się przed drobnymi błędami językowymi (głównie deklinacyjnymi) i redakcyjnymi.

4. Oryginalny dorobek Doktoranta, jego znaczenie poznawcze oraz przydatność praktyczna dla nauki i techniki

Opracowana przez Doktoranta metoda wykrywania wiadomości phishingowych z wykorzystaniem uczenia maszynowego jest oryginalną propozycją w obszarze badań nad technikami wykrywania cyberataków. Stanowi alternatywę dla istniejących narzędzi i propozycji na etapie koncepcyjnym. Przedstawione przez Doktoranta analizy wskazują obiecującą skuteczność detekcji phishingu.

Ważnym osiągnięciem Doktoranta jest rozpoznanie, w ramach analizy literatury oraz obserwacji praktycznych, obszernego zbioru wskaźników występowania ataku phishingowego. Inne oryginalne osiągnięcia Doktoranta to:

- zaimplementowanie zaproponowanej metody detekcji phishingu i przeprowadzenie eksperymentów mających na celu weryfikację oraz poprawianie jej efektywności,
- zaprojektowanie i zaimplementowanie nowego algorytmu odczytującego pola nagłówka wiadomości e-mail zawierającego rozszerzania analizy treści i poprawności językowej,
- opracowanie komponentu uczenia maszynowego do analizy odnośników URL,
- analiza i rozpoznanie licznych technik realizacji phishingu oraz sposobów ukrywania faktu ataku,
- rozpoznanie nowych adresów IP oraz adresów e-mail identyfikowanych jako phishingowe.

Wszystkie te osiągnięcia posiadają istotne znaczenie poznawcze. Wykazują również potencjał do zastosowania praktycznego. W przypadku tego ostatniego konieczne byłyby m.in. dalsze prace związane z analizą efektywności, złożoności i skuteczności.

5. Wiedza Autora oraz znajomość współczesnej literatury z dyscypliny naukowej, której dotyczy rozprawa

Jak już wspomniano w rozdziale 3 niniejszej recenzji, treści rozprawy dotyczące ataków phishingowych (kategorii, technik, etapów etc.) świadczą o dobrej znajomości obszaru badawczego przez Doktoranta, w tym identyfikację odpowiedniej literatury. Dobre rozpoznanie poruszanej tematyki przez Doktoranta, nawet jeszcze bardziej potwierdza rozdział trzeci, w którym przedstawiono blisko 30 rozpoznanych przez Doktoranta wskaźników występowania ataku phishingowego.

Pewien niedosyt pozostawia natomiast rozdział drugi opisujący główne techniki i metody detekcji phishingu. Obejmuje on tylko pewną część obszernej literatury dziedzinowej i nasuwa pytanie dotycząca stopnia kompletności analiz w tym aspekcie. Stopień ten zostałby zwiększony przy zastosowania systematycznego podejścia analizy literatury.

Ogólnie Doktorant wykazał dobrą znajomość współczesnej literatury z dyscypliny naukowej, której dotyczy rozprawa.

6. Podsumowanie

W podsumowaniu stwierdzam, że mimo pewnych uchybień przedstawionych powyżej, przedłożona do recenzji rozprawa doktorska wykonana przez Pana kpt. mgr. inż. Mariusza Cezarego Szarka przedstawia oryginalny i znaczący dorobek Doktoranta oraz potwierdza jego szeroką wiedzę i dobrą znajomość literatury dziedzinowej.

W mojej opinii spełnia zatem ustawowe wymagania stawiane rozprawom doktorskim.

Na tej podstawie wnioskuję o jej dopuszczenie do następnego etapu postępowania w sprawie nadania stopnia doktora.

Rafał Leszczyna