Warszawa, 18 marca 2024 r.

Prof. dr hab. inż. Ewa Niewiadomska-Szynkiewicz
Wydział Elektroniki i Technik Informacyjnych
Politechnika Warszawska
ewa.szynkiewicz@pw.edu.pl

RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY DYSCYPLINY NAUKOWEJ INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA WOJSKOWEJ AKADEMII TECHNICZNEJ

Tytuł rozprawy: Budowanie mechanizmu obrony przed dedykowanymi kampaniami phishingowymi

Autor rozprawy: mgr inż. Mariusz Szarek

1. Ogólna charakterystyka rozprawy. Cel badań.

Zespoły ds. cyberbezpieczeństwa z różnych krajów potwierdzają, że z każdym rokiem sukcesywnie obserwuje się wzrost liczby incydentów i ataków w cyberprzestrzeni. Są one coraz bardziej wyrafinowane i trudniejsze do wykrycia. Celem jest zazwyczaj wykonywanie niepożądanych operacji na komputerze ofiary skutkujących m.in. dystrybucją złośliwego oprogramowania w sieci, przechwytywaniem wrażliwych informacji oraz zakłócaniem pracy. Od kilku lat obserwuje się, że najczęstszym typem ataku są oszustwa komputerowe, które stanowią 80-90% wszystkich obsługiwanych incydentów bezpieczeństwa.

Przedmiotem badań udokumentowanych w rozprawie są metody i techniki wykrywania kampanii phishingowych, czyli oszustw, w których wykorzystuje się inżynierię społeczną do wyłudzenia poufnych informacji, zainfekowania komputera szkodliwym oprogramowaniem, czy też nakłonienia ofiary do wykonania określonych działań. To obecnie jeden z najpopularniejszych typów oszustw komputerowych.

W opiniowanej rozprawie Doktorant bardzo szczegółowo omówił różne modele cyberataków, koncentrując się na technikach realizacji ataku typu phishing. Szczególną uwagę zwrócił na atrybuty wiadomości, które mogą być wykorzystane do wykrywania oszustw komputerowych. Przedstawił stosowane powszechnie podejścia do detekcji phishingu, wskazując na ich niedoskonałości i słabe strony. Jako główny cel badań przyjął opracowanie autorskiej metody zakładającej szczegółową analizę przesyłanych w sieci wiadomości i wykorzystanie algorytmów uczenia maszynowego do ich klasyfikacji, a w efekcie do wykrywania phishingu. W sformułowanej w pracy tezie badawczej stwierdził,

że proponowane w pracy rozwiązanie umożliwi detekcję nowych ataków phishingowych realizowanych zgodnie z niestosowanymi dotychczas schematami, zawierającymi nieznane wzorce. W tym kontekście uważam tematykę pracy za istotną i aktualną, o rezultatach, które mogą być zastosowane w praktyce do podniesienia poziomu bezpieczeństwa cyberprzestrzeni. Fakt ten przesądza o pozytywnej ocenie wybranego tematu jako przedmiotu opiniowanej rozprawy doktorskiej.

Podsumowując, rozprawa ma głównie charakter konstrukcyjny i doświadczalny. Rozważania teoretyczne koncentrują się na analizie przesyłanych w sieci teleinformatycznej komunikatów i wskazaniu atrybutów komunikacji typowych dla tego typu ataków oraz technik związanych z ich propagowaniem. Na tej podstawie autor zaproponował techniczne i nietechniczne wskaźniki phishingu. Część konstrukcyjna pracy obejmuje opracowanie i realizację narzędzia do wykrywania kampanii phishingowych. Część doświadczalna zawiera wyniki eksperymentów wykonanych dla danych testowych przygotowanych przez Doktoranta m.in. zasobów sieci Internet. Zamieszczone w pracy wyniki badań pozwalają uznać, że Doktorant osiągnął założony cel oraz potwierdzają słuszność podjętych przez niego działań.

2. Syntetyczna analiza treści rozprawy. Charakter rozprawy

Zasadnicza część rozprawy składa się z pięciu rozdziałów. Rozdział pierwszy zawiera wprowadzenie w tematykę pracy. Przedstawiony jest kontekst rozważanego zagadnienia oraz zaprezentowane uzasadnienie podjętego problemu badawczego. Doktorant wskazuje na bardzo silny trend wzrostowy w zakresie zagrożeń w sieci. Omawia dwa powszechnie uznawane modele ataków cybernetycznych, tj. Cyber Kill Chain oraz MITREW ATT&CK. Szczególną uwagę zwraca na zjawisko phishingu. Podaje podstawowe definicje, kryteria, opisuje proces, którego efektem jest atak. Prezentuje liczne scenariusze takiego ataku oraz różnego rodzaju techniki realizacji kampanii phishingowych i odpowiadające im rodzaje phishingu. Rozróżnia ataki typu phishing i spam. Wymienia czynniki, które istotnie przyczyniły się do zainteresowania tego typu atakami przez cyberprzestępców. Zwraca uwagę na zagrożenia związane z rozwojem metod sztucznej inteligencji i ich coraz powszechniejszym zastosowaniem do generacji nowych ataków, w tym obejmujących oszustwa komputerowe. Jako przykład wskazuje na możliwości szkodliwego wykorzystania ChatGPT. Analizuje stosowane przez atakujących metody inżynierii społecznej oraz podatności użytkowników na zagrożenia w sieci. Materiał zawarty w tym rozdziale jest bardzo obszerny, zawiera duży zasób wiedzy na temat zjawiska phishingu, stanowi punkt wyjścia dla treści prezentowanych w dalszej części pracy.

Rozdział drugi jest poświęcony przeglądowi różnych podejść do wykrywania kampanii phishingowych. Prezentowane są proste techniki zakładające tworzenie list zawierających niedozwolone lub podejrzane adresy URL oraz IP, które są potencjalnymi kandydatami generującymi niebezpieczne komunikaty. Omawiane są reguły detekcji stosowane do wykrywania ataków czy incydentów bezpieczeństwa. Zwrócona jest uwaga, że w przypadku phishingu badane jest głównie podobieństwo treści. Następnie uwaga koncentruje się na zastosowaniu technik optymalizacji, w tym metod czerpiących inspiracje z biologii. Przedstawione są prace, w których do detekcji zostały wykorzystane algorytmy genetyczne optymalizacji. Ostatnią grupą analizowanych technik są podejścia wykorzystujące eksplorację danych i algorytmy uczenia maszynowego. Autor prezentuje kilka rozwiązań opisanych w literaturze polegających na zastosowaniu klasycznych klasyfikatorów, takich jak: naiwny klasyfikator Bayes'a, drzewa decyzyjne, klasyfikator asocjacyjny, maszyna wektorów nośnych. Omawiane są różne koncepcje, m.in. polegające na wyszukiwaniu ukrytych korelacji między danymi i tworzeniu odpowiednich klas, analizowaniu formatowania HTML

wiadomości email, klasyfikowaniu wiadomości na podstawie zestawu zdefiniowanych cech itd. Doktorant dość dokładnie analizuje opisywane podejścia. Szczególną uwagę zwraca na ich ograniczenia, które dokładnie omawia. Wskazuje m.in. na brak odporności prezentowanych rozwiązań na zmieniające się techniki atakujących. Ponadto, ze względu na fakt, iż większość opisywanych metod wykorzystuje wzorce lub schematy generowania zidentyfikowanych wcześniej ataków, występują poważne problemy z wykryciem nieznanych ataków (ataki 0-day).

Rozdziały trzy, cztery i pięć są kluczowe, gdyż zawierają wyniki prac badawczych Doktoranta. W rozdziale trzecim są bardzo szczegółowo analizowane techniki propagowania oraz własności złośliwych wiadomości. Efektem tych analiz jest autorska lista atrybutów, która zdaniem Doktoranta może być wykorzystana do klasyfikacji wiadomości i wykrycia oszustwa. Ostatecznie zaproponowanych zostało 26 atrybutów, którym w procesie analizy i oceny wiadomości przypisywane są odpowiednio wartości 0, 1 i „null". Atrybuty z przypisanymi wartościami stanowią listę wskaźników, która jest wykorzystywana w opracowanych przez Doktoranta algorytmach detekcji kampanii phishingowych. Rozdział jest bardzo obszerny, wskazane atrybuty wiadomości są szczegółowo analizowane na podstawie historycznych ataków. Omawiane są typowe wartości atrybutów oraz te, które wskazują na możliwy atak, różne scenariusze ataków oraz sytuacje, które mogą wpływać na błędne wyniki klasyfikacji. Autor porusza wiele wątków, co z jednej strony świadczy o wysokich kompetencjach w zakresie analizowania phishingu, z drugiej ta wielowątkowość i zbyt częste odbieganie od głównego tematu nieco utrudnia studiowanie rozprawy.

W rozdziale czwartym opisana jest opracowana przez Doktoranta metoda automatycznego wykrywania wiadomości phishingowych. Autor odwołuje się do wskazanych w rozdziale trzecim atrybutów wiadomości, które można wykorzystać do ich klasyfikacji, i ostatecznie wybiera 19 atrybutów z zaproponowanej listy, które posłużą do detekcji ataków. Przedstawiony na str. 177 rysunek 47 pokazuje schemat działania algorytmu realizującego koncepcję Autora. Proces wykrywania ataków jest realizowany w dwóch etapach. W pierwszym dokonuje się wstępnego przetworzenia wiadomości pod kątem wybranych atrybutów i ustawienia wartości wskaźników, w drugim uruchamiane są algorytmy uczenia maszynowego w celu przypisania analizowanej wiadomości do jednej z trzech klas: normalna, phishing, spam. W rozdziale opisanych jest pięć klasycznych algorytmów uczenia maszynowego: maszyna wektorów nośnych (SVM), naiwny klasyfikator Bayes'a, drzewo decyzyjne, las losowy, regresja logistyczna.

Rozdział piąty zawiera wyniki eksperymentów, których celem była weryfikacja poprawności działania i ocena skuteczności kilku wariantów opracowanego i zrealizowanego algorytmu klasyfikacji wiadomości. Zdefiniowane są metryki oceniające jakość rozwiązania – typowe stosowane w ocenie klasyfikatorów. Prezentowane rezultaty dotyczą dwóch serii eksperymentów. Celem pierwszej była ocena poprawności wydobywania i analizy rozważanych atrybutów z wiadomości i ustawiania odpowiednich wartości wskaźników. Druga seria eksperymentów była wykorzystana do oceny jakości klasyfikacji wiadomości na: normalne, spam i phishing. Rozdział prezentuje liczne wyniki testów przeprowadzonych dla różnych zbiorów uczących. Badany jest wpływ zbalansowania liczności podzbiorów zawierających dane normalne, phishing i spam na jakość klasyfikacji. Opisane są różne podejścia do tworzenia zbalansowanych zbiorów.

Pracę kończy podsumowanie zawierające syntetyczne omówienie przeprowadzonych analiz oraz uzyskanych wyników. Doktorant wskazuje na zalety opracowanej metody wykrywania wiadomości phishingowych, wspomina też o jej ograniczeniach i pokazuje możliwe kierunki rozwoju.

3. **Ocena analizy źródeł i sposobu sformułowania wniosków wynikających z analizy źródeł.**

Ogółem Autor w rozprawie odwołuje się do 100 pozycji bibliograficznych związanych z tematyką pracy. Poza publikacjami naukowymi są tu odwołania do dokumentów, standardów, raportów zespołów badawczych oraz firm zajmujących się cyberbezpieczeństwem. W ostatnich latach pojawiło się wiele propozycji związanych m.in. z wykorzystaniem metod uczenia maszynowego do detekcji malware, złośliwych kampanii, kampanii phishingowych i spamowych. Uwzględniając to, zamieszczony w rozdziale drugim przegląd prac naukowych mógłby być nieco obszerniejszy i głębszy. Niezbyt liczną grupę stanowią ponadto publikacje z renomowanych czasopism.

W rozdziale II, który jest głównie poświęcony omówieniu i analizie źródeł, opisane są różne techniki wykrywania phishingu, od metod wykorzystujących listy skompromitowanych i godnych zaufania adresów URL i IP, przez metody regułowe, heurystyki (algorytmy genetyczne) po algorytmy uczenia maszynowego. W pracy brakuje omówienia narzędzi, w których stosowane są sztuczne sieci neuronowe, w tym sieci głębokie. W ostatnich latach podejmowane były liczne próby, w których badano skuteczność tego typu modeli w zadaniach wykrywania cyberataków, w tym wiadomości phishingowych. Ze względu na fakt, że opracowany w ramach pracy doktorskiej algorytm zakłada analizę treści wiadomości wskazane byłoby również krótkie odniesienie się do metod i narzędzi dedykowanych przetwarzaniu języka naturalnego i uzasadnienie, dlaczego nie były one wykorzystywane.

W sekcji V.2.2 opisywane są, zastosowane przez Doktoranta, metody uczenia maszynowego (SVM, naiwny klasyfikator Bayes'a, drzewo decyzyjne, las losowy, regresja logistyczna). W przypadku wszystkich tych opisów brakuje odwołań do publikacji twórców tych rozwiązań. Publikacje te nie zostały również zamieszczone w bibliografii.

Podsumowując, uważam, że mimo wymienionych braków przedstawiona w rozprawie analiza źródeł oraz postawione na podstawie przeglądu wnioski świadczą o wiedzy autora w przedmiocie rozprawy. Doktorant analizuje prezentowane w literaturze podejścia i rozwiązania do wykrywania wiadomości phishingowych, formułuje zagadnienia badawcze i proponuje autorską metodę.

4. **Analiza poprawności rozwiązania przedstawionego zadania, poprawność przyjętych założeń i wybranych metod.**

Doktorant rozwiązał postawione w pracy zagadnienie. Przeprowadził bardzo głęboką analizę wiadomości phishingowych, ich charakterystycznych cech, technik propagowania, wykorzystywanych podatności itd. Analizując opisane w literaturze tematu rozwiązania, w szczególności ich ograniczenia oraz wykorzystując obserwacje innych badaczy opracował autorską metodę i zrealizował kilka wariantów algorytmów wykorzystujących różne metody uczenia maszynowego do wykrywania kampanii phishingowych. Przygotował odpowiednie oprogramowanie oraz zbudował zbiory rzeczywistych danych do trenowania i testowania zrealizowanych algorytmów. Rozważył różne metody rozszerzania zbiorów danych trenujących w przypadku ich niezbalansowania. Wykonał, wnikliwie przeanalizował i udokumentował liczne eksperymenty badawcze.

Zdaniem recenzenta przyjęta metoda badawcza obejmująca dokładne omówienie proponowanego rozwiązania, sformułowanie odpowiednich algorytmów, ich wstępną weryfikację, a następnie wykonanie licznych testów dla różnych zestawów danych rzeczywistych jest właściwa dla badanego zagadnienia. Uzyskane wyniki potwierdzają poprawność rozwiązania. Autor w rozprawie doktorskiej uzasadnia, że zaproponowana

technika może wspierać detekcję zagrożeń w systemach teleinformatycznych, a uzyskane rezultaty pokazują jej skuteczność. To czego brakuje w rozprawie, to studium porównawczego z wybranym/wybranymi rozwiązaniami proponowanymi w literaturze. Brakuje również eksperymentów potwierdzających wprost to co zostało zawarte w tezie rozprawy – skuteczność opracowanej metody w wykrywaniu nowych ataków, dla których nie posiadamy wzorca.

5. **Oryginalność rozprawy, samodzielny dorobek autora, pozycja rozprawy w stosunku do stanu wiedzy prezentowanego w literaturze światowej.**

Rozprawa zawiera nowe oryginalne rezultaty. Do szczególnie ważnych należy zaliczyć:

- Szczegółowa analiza phishingu, wyodrębnienie atrybutów wiadomości świadczących o potencjalnym oszustwie i opracowanie listy wskaźników, które mogą być wykorzystane do automatycznego wykrywania wiadomości phishingowych przez odpowiednie narzędzia programistyczne.
- Opracowanie listy wskaźników, tzw. nietechnicznych, które mogą wspierać eksperta w procesie wykrywania kampanii phishingowych.
- Opracowanie metody detekcji wiadomości phishingowych i realizacja kilku wariantów algorytmu różniących się wykorzystanymi technikami uczenia maszynowego.
- Wykazanie poprawności i skuteczności zaproponowanych rozwiązań za pomocą eksperymentów wykonanych na rzeczywistych danych.

Przedstawione wyniki mają istotne znaczenie w zakresie rozwoju metod detekcji kampanii phishingowych i podniesienia poziomu cyberbezpieczeństwa.

6. **Analiza poprawności prezentacji wyników pracy**

Recenzowana rozprawa doktorska liczy 273 strony, zawiera 76 rysunków, 58 tabel oraz bibliografię obejmującą 100 publikacji. Zasadnicza treść obejmuje 243 strony, pozostałe zawierają spis treści, wykaz rysunków, tabel, skrótów, trzy dodatki oraz bibliografię. Pewnym utrudnieniem dla systematycznego czytelnika jest brak w opracowaniu spisu oznaczeń używanych w formułach matematycznych.

Układ rozdziałów jest generalnie prawidłowy, ale mogłyby być one lepiej napisane. Obecnie opisy są dość chaotyczne i często niejasne. Autor porusza wiele wątków, co z jednej strony świadczy o wysokich kompetencjach w zakresie analizowania phishingu, z drugiej ta wielowątkowość i zbyt częste odbieganie od głównego tematu rozdziału nieco utrudnia studiowanie rozprawy. Praca znacznie by zyskała, gdyby rozdziały były krótsze, opisy konkretne, bez nadmiernych szczegółów, a przede wszystkim powtórzeń. O tym jakie własności wiadomości sugerują, że jest ona oszustwem dowiadujemy się praktycznie w każdym rozdziale. Zastrzeżenia dotyczą przede wszystkim struktury i treści rozdziałów 3, 4 i 5. Czytelność pracy znacznie by się poprawiła, gdyby wszystkie rozważania dotyczące własności wiadomości wskazujących na jej złośliwość, wybrane przez Autora wskaźniki do automatycznej detekcji phishingu itd., zostały zamieszczone w rozdziale trzecim. Opisy dotyczące tych zagadnień zawarte w rozdziale czwartym burzą jego strukturę, np. sekcja IV.2 „Problem poprawności cech". Opisy punktów w sekcji IV.1 są nadmiarowe – wystarczyło wymienić wybrane atrybuty, zdanie na str. 161 (sekcja IV.1.1) „Konstruując metodę wykrywania wiadomości phishingowych, konieczne jest opracowanie modelu...." nic nowego nie wnosi – było to poruszane we wcześniejszych rozdziałach. Podobnie, wyniki analiz wiadomości, np. zliczanie liczby kropek w adresach URL itd. powinny znaleźć się

w rozdziale trzecim. Dlaczego w tym rozdziale mamy sekcje poświęcone przygotowaniu wektora cech (sekcja IV.3.3), a tym bardziej opis możliwego rozszerzenia baz gromadzących historyczne ataki (sekcja IV.3.3.1)? Rozdział czwarty powinien w całości być poświęcony opracowanej przez Autora metodzie detekcji i realizujących ją algorytmów - opisowi autorskich dokonań. Powinien się koncentrować i rozpocząć od celu metody, tj. klasyfikacji wiadomości na normalne, spam i phishing oraz kluczowego dla rozprawy, ogólnego schematu algorytmu (rys. 43, str. 177). Schemat ten powinien być znacznie dokładniejszy, dokładniej omówiony, powinien stanowić bazę dla dalszych rozważań. W następnej kolejności wskazane byłoby zwarte omówienie kolejnych kroków algorytmu, łącznie z informacją z jakich narzędzi się w nich korzysta. Z opisu dowiadujemy się, że do „kodowania cech" zastosowano maszynę wektorów nośnych (SVM), która jest omówiona kilka stron dalej. Podobnie jest z rozdziałem piątym. Opisy są bardzo chaotyczne. Na przykład we wstępie do rozdziału (str. 196) mówi się o tym, że weryfikacja algorytmu detekcji jest wykonywana w dwóch etapach. Na stronie 217 mamy już sześć etapów. W sekcji V.2.3 podanych jest mnóstwo szczegółowych, często bardzo technicznych informacji (np. o dodaniu obsługi wyjątków), przez co umyka to co jest najważniejsze. Ponadto w tym miejscu pojawia się wątpliwość, czy opracowane narzędzie jest uniwersalne, skoro dla konkretnego eksperymentu trzeba było wprowadzać szereg modyfikacji? Z opisu wynika, że rozdział ten jest poświęcony weryfikacji poprawności implementacji i wprowadzaniu modyfikacji w kodzie. Jeśli tak, to nie ma uzasadnienia na umieszczanie go w rozdziale z wynikami eksperymentów. Nie jest do końca jasne na czym polegała różnica w przypadku równoważenia zbiorów – dwa podejścia opisane na str. 223, tj: „równoważenie zbioru po jego wczytaniu" i „utworzenie zrównoważonego zbioru na bazie zbioru oryginalnego".

Strukturę rozdziałów czwartego i piątego burzą opisy wybranych metod uczenia maszynowego oraz technik generowania dodatkowych danych w celu równoważenia zbiorów uczących. Powinny być one wprowadzone znacznie wcześniej, np. w rozdziale drugim z przeglądem literatury lub odrębnym rozdziale, dedykowanym metodom uczenia maszynowego i tworzenia zbiorów danych do ich trenowania. Zamieszczony opis SVM dotyczy tylko klasyfikatora liniowego. Autor nie wspomina nic o wersjach z nieliniowym jadrem, a dalej, na str. 189 Autor mówi o „wyborze jądra". Opisy kończą dane techniczne – podawane są nazwy klas, metod itd. Nie ma zwyczaju przedstawiania w pracach doktorskich szczegółów dotyczących konkretnych implementacji, a jeśli już, to tego typu informacje powinny być zamieszczone w dodatku.

Redakcja pracy budzi pewne zastrzeżenia, nie jest do końca profesjonalna. Przyjęty styl definiowania rozdziałów i poszczególnych sekcji nie jest wygodny dla czytelnika. Główne zastrzeżenia dotyczą prezentacji formuł matematycznych. Wzory są pisane za pomocą różnego typu czcionek o różnej wielkości. W profesjonalnych tekstach naukowych formuły są pisane takim rozmiarem czcionki jak pozostały tekst. Warto było wykorzystać profesjonalne środowiska do wprowadzania formuł matematycznych. Objaśnienia niektórych zmiennych powinny być bardziej szczegółowe. Formuła (2.4), dla P(v)=0 lub 1 otrzymamy dzielenie przez zero, formuła (2.6), chyba powinno być „m".

W pracy występują dość liczne błędy językowe, gramatyczne i interpunkcyjne. Razi konsekwentne stosowanie słowa „ilość" do rzeczowników policzalnych (powinno być „liczba"). Nie jest też jasne, dlaczego Doktorant używa czasu przyszłego do opisów prac, które wykonał w przeszłości. Zaproponowana notacja jest w pewnych miejscach nieoczywista i często przypadkowa. Mieszane są pojęcia „cecha" i „wskaźnik". Często trudno jest się zorientować co autor ma na myśli, np. „prawidłowe działanie funkcji algorytmu" (str. 198), „mechanizm bazujący na heurystyce odnośnika URL" (str. 185),

„zastosowanie lepszych warunków logicznych" (str. 230), „przekrzywione dane" (str. 191), „optymalny parametr uczenia" (str. 218), „według trójpodziału" (str. 96), itd.

W treści rozdziałów nie ma żadnych odwołań do zamieszczonych rysunków i tabel. Utrudnia to czytanie pracy. Na wykresach prezentujących wyniki eksperymentów brakuje opisów osi. Rysunek 62 – co Autor rozumiał przez „progres uczenia"? Czym różni się w tym przypadku „równoważenie" od „balansowania"? Oba słowa w rozważanym kontekście mają takie samo znaczenie.

7. Słabe strony rozprawy i jej główne wady

Słabe strony rozprawy, to nawiązując do poprzedniego punktu, mało klarowna i zbyt chaotyczna prezentacja wyników prac badawczych, w tym opracowanej metody wykrywania wiadomości phishingowych oraz wyników przeprowadzonych eksperymentów. Uwagi dotyczą zarówno braku precyzyjnego, zwartego opisu celu pracy oraz opracowanej metody wykrywania kampanii phishingowych i wykonanych wariantów algorytmu, jak i przyjętej formy prezentacji metody. Obniża to istotnie czytelność pracy i zrozumienie koncepcji proponowanego rozwiązania. Spore zastrzeżenia budzi opis eksperymentów, a szczególnie brak jawnego podania liczności zbiorów uczących po ich zbalansowaniu, przy wykorzystaniu opisanych w pracy technik. Na stronach 272-274 Autor podał macierze pomyłek dla zbiorów testowych uzyskanych za pomocą sześciu metod, z których można wnioskować jakie były rozmiary zbiorów, ale niestety nie dołączył żadnego komentarza i objaśnień. Tego typu informacje powinny być zamieszczone w rozdziałach z wynikami testów, a nie tylko w dodatku. Wpływają na ocenę wiarygodności wyników.

Główna uwaga krytyczna dotyczy eksperymentów, które zostały przeprowadzone przy bardzo dużym niezbalansowaniu danych trenujących i testowych: próbki normalne=7, phishingowe=47 (rozdz. V.1), próbki normalne = 80, spam=378, phishing=1168 (rozdz. V.2). Przy takim rozkładzie danych trudno jest cokolwiek wnioskować, szczególnie że oceniano klasyfikatory tylko na podstawie dokładności klasyfikacji (ACC - *accuracy*). W rozdziale V.1.2 zostały opisane podstawowe miary oceny klasyfikatorów. Brakuje pola pod krzywą ROC. Nie jest oczywiste, dlaczego do oceny wykonanego detektora została wybrana tylko jedna z nich. Autor nie uzasadnił wyboru. Miara ACC nie zawsze daje pełny obraz, zazwyczaj podaje się dodatkowo wartości miar precyzja i czułość lub pole pod krzywą ROC.

W rozdziale I, na str. 82 sformułowana jest teza rozprawy:

> *„...połączenie powyżej opisanych sposobów detekcji pozwoli na wykrywanie wcześniej niestosowanych schematów ataku, używanie nieznanych do tej pory wzorców – co eliminuje wady klasycznych metod wykrywania phishingu."*

Autor do tezy odnosi się wyraźnie dopiero w podsumowaniu pracy. Stwierdza, że niektóre wskazane przez niego własności wiadomości phishingowej, dotychczas nie uwzględniane w procesie detekcji, pozwolą na wykrywanie nowych ataków. Nie jest jasne czy były wykonane testy porównawcze pokazujące jakość detekcji nowych (np. sztucznie wygenerowanych) wiadomości phishingowych w przypadku, gdy detektory były wyuczone na danych zawierających wspomniane wskaźniki oraz danych, w których nie były one uwzględnione.

W pracy brakuje porównania autorskiego rozwiązania z wybranymi metodami opracowanymi przez innych badaczy. Takie studium porównawcze potwierdziłoby, że metoda zaproponowana przez Doktoranta jest dobrą alternatywą.

8. Przydatność rozprawy dla nauk technicznych

Mimo wymienionych powyżej słabych stron pracy uważam, że przedstawione w rozprawie wyniki wnoszą istotny wkład w dyscyplinę informatyka techniczna i telekomunikacja, a konkretnie w rozwój badań w zakresie cyberbezpieczeństwa. Opracowana metoda detekcji wiadomości phishingowych może przyczynić się do zwiększenia skuteczności ochrony przed tego typu atakami, również tymi, które jeszcze nie były identyfikowane. Metoda ma duży potencjał rozwojowy. Wykonane eksperymenty badawcze potwierdzają jej efektywność. Należy podkreślić, że tematyka rozprawy jest bardzo aktualna.

9. Podsumowanie i wniosek końcowy

Uważam, że rozprawa mgr inż. Mariusza Szarka Ostapa spełnia wymagania stawiane rozprawom doktorskim przez przepisy Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. 2023, poz. 742) w odniesieniu do rozpraw doktorskich. W związku z tym wnoszę o przyjęcie rozprawy i dopuszczenie Autora do dalszych, przewidzianych przepisami, etapów przewodu doktorskiego.

podpis