Dr hab. inż. Wojciech Mazurczyk, profesor uczelni
Politechnika Warszawska
Wydział Elektroniki i Technik Informacyjnych
Instytut Informatyki
Ul. Nowowiejska 15/19
00-665 Warszawa

27.04.2020

**RECENZJA ROZPRAWY DOKTORSKIEJ
DLA RADY WYDZIAŁU TELEKOMUNIKACJI, INFORMATYKI
I ELEKTROTECHNIKI
UNIWERSYTETU TECHNOLOGICZNO-PRZYRODNICZEGO
W BYDGOSZCZY**

Tytuł rozprawy: „Zastosowanie metod uczenia maszynowego do wykrywania ataków sieciowych"

Autor rozprawy: mgr inż. Marek Pawlicki

1. Jakie zagadnienie naukowe jest rozpatrzone w pracy (teza rozprawy) i czy zostało ono dostatecznie jasno sformułowane przez autora?

Tematyka recenzowanej rozprawy doktorskiej koncentruje się na wybranych aspektach zapewniania bezpieczeństwa sieciowego. W szczególności jej problematyka naukowa poświęcona jest zagadnieniom wykrywania ataków w sieciach teleinformatycznych wykorzystując do tego celu techniki uczenia maszynowego oraz zabezpieczaniu takich rozwiązań przed manipulacją ze strony atakującego. Tematyka podjęta w rozprawie jest ważna i aktualna, gdyż na przestrzeni ostatnich lat można zaobserwować znaczący wzrost różnego rodzaju ataków sieciowych, co w rezultacie przekłada się na straty finansowe i wizerunkowe firm, instytucji, organizacji oraz zwykłych użytkowników sieci.

W związku z powyższym niezbędne jest prowadzenie badań naukowych w sposób dwutorowy. Z jednej strony konieczne jest wskazywanie nowych, potencjalnych sposobów detekcji zagrożeń sieciowych z wykorzystaniem technik uczenia maszynowego i ulepszenie istniejących rozwiązań. Natomiast z drugiej strony istotne jest także opracowywanie nowatorskich sposobów ograniczających lub wręcz uniemożliwiających atakującemu oddziaływanie na systemy bezpieczeństwa, gdyż brak takich zabezpieczeń może prowadzić do ich niskiej skuteczności lub braku możliwości wykrycia aktywnych zagrożeń.

Z tej perspektywy rozprawa doktorska mgr inż. Marka Pawlickiego wpisuje się w bieżący nurt prac badawczych bezpieczeństwa sieciowego, gdzie obecnie jednym z głównych kierunków badawczych cyberbezpieczeństwa na świecie jest ocena możliwości wykorzystania technik uczenia maszynowego do wykrywania cyberataków.

We wstępie rozprawy zdefiniowano następującą tezę (podrozdział 1.2):

„Możliwe jest opracowanie złożonych algorytmów uczenia maszynowego dla skutecznej detekcji ataków na podstawie przepływów sieciowych oraz opracowanie technik poprawy ich wiarygodności i bezpieczeństwa (wykrywania ataków typu adversarial learning)."

W mojej ocenie teza rozprawy została postawiona prawidłowo i jest jasno sformułowana. W celu jej udowodnienia Doktorant zaproponował i przebadał szereg nowych podejść oraz usprawnień pozwalających zrozumieć zależności pomiędzy poszczególnymi komponentami tego typu rozwiązań, a w konsekwencji zwiększyć ich efektywność. Z drugiej strony przeanalizowane zostały także szczegółowo ataki na metody detekcji wykorzystujące techniki uczenia maszynowego i zaproponowano rozwiązania im przeciwdziałające.

2. Jaka jest przydatność rozprawy z punktu widzenia nauk technicznych, czy założenia przyjęte przez autora są uzasadnione?

Praktyczna przydatność rozprawy dla nauk technicznych jest potencjalnie duża. Autor proponuje szereg nowatorskich rozwiązań zarówno w kontekście wykorzystania technik uczenia maszynowego do detekcji ataków sieciowych jak i sposobów wykrywania prób manipulacji tych rozwiązań. Doktorant przedstawia także wyniki badań eksperymentalnych dla zróżnicowanych zbiorów danych oraz systematycznie weryfikuje szereg zagadnień związanych z technikami uczenia maszynowego związanymi np. z balansowaniem zbiorów, czy wpływem doboru hiperparametrów na wyniki detekcji ataków sieciowych. Należy także stwierdzić, że założenia przyjęte przez mgr Pawlickiego w rozprawie należy uznać za racjonalne oraz uzasadnione, gdyż pozwalają one na dokonanie oceny proponowanych rozwiązań w warunkach odpowiadających tym, które występują obecnie w sieciach teleinformatycznych.

3. Czy autor rozwiązał postawione zagadnienie i czy użył właściwej do tego metody?

W celu udowodnienia postawionej w pracy doktorskiej tezy Autor, zaproponował nowe sposoby detekcji ataków sieciowych wykorzystujące metody uczenia maszynowego oraz rozwiązanie pozwalające na wykrycie ataków typu *adversarial learning* na tego typu detektory.

Dla każdego z wymienionych elementów w sposób szczegółowy mgr Pawlicki scharakteryzował przyjęte założenia, sposób realizacji technik uczenia maszynowego oraz zastosowane środowisko eksperymentalne jak i wykorzystaną metodologię badawczą. Następnie Doktorant przeprowadził badania eksperymentalne, analogiczne dla każdego z zaproponowanych elementów oraz przedstawił uzyskane wyniki zarówno w postaci wykresów jak i formie tabelarycznej. Uzyskane wyniki eksperymentalne zostały przedstawione i omówione są w sposób wystarczająco jasny i zrozumiały, a wyciągnięte na ich podstawie wnioski są poprawne.

Podsumowując, należy uznać, że rozwiązanie postawionego zadania badawczego zostało przeprowadzone w sposób zasadniczo prawidłowy. Pewne kwestie dyskusyjne zostały natomiast zawarte w punkcie 6 niniejszej recenzji.

4. Jaki charakter ma rozprawa (teoretyczny, doświadczalny, konstrukcyjny), jaka jest jej pozycja w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Rozprawa została przygotowana w języku polskim i składa się z 8 rozdziałów, tabel wyników badań, listy tabel i rysunków, wykazu skrótów oraz bibliografii. Całość pracy obejmuje 120 stron. Wyniki własne doktoranta przedstawione są w rozdziałach 3-4 oraz 6-7. Stan

dotychczasowych badań w głównym zakresie tematycznym rozprawy przedstawiono w rozdziale 2 dla wykorzystania technik uczenia maszynowego do detekcji ataków sieciowych oraz w rozdziale 5, gdzie dokonano przeglądu metod ataku typu *adversarial learning*.

Rozprawa doktorska mgr inż. Marka Pawlickiego ma charakter konstrukcyjno-eksperymentalny, co w szczególności potwierdza zawartość rozdziałów 3-4 oraz 6-7. Doktorant opracowuje metodyki badawcze i przeprowadza szereg eksperymentów różnych konfiguracji technik uczenia maszynowego w celu opracowania dogodnej metody detekcji ataków sieciowych wykorzystując do tego celu zbiory danych zawierające starsze (NSL-KDD) i nowe (CICIDS2017) zagrożenia spotykane w sieciach teleinformatycznych. Ponadto, Autor proponuje także nowatorski detektor pozwalający wykrywać próby atakującego mające na celu wprowadzenie w błąd rozwiązania detekcji ataków sieciowych poprzez celową manipulację danych, które ono wykorzystuje (ataki typu *adversarial learning*) oraz przeprowadza systematyczną analizę skuteczności zaproponowanego podejścia.

Należy także podkreślić, że badania prowadzone przez Autora rozprawy wpisują się w bieżący i ważny nurt badań naukowych, którego celem jest uzyskanie odpowiedzi na pytanie na ile skuteczne i bezpieczne będzie zastosowanie technik uczenia maszynowego w cyberbezpieczeństwie. Z tej perspektywy należy uznać, że rozprawa Doktoranta wpisuje się bardzo dobrze w stan wiedzy oraz poziom techniki reprezentowany obecnie przez literaturę światową.

5. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora?

Oceniając dorobek rozprawy stwierdzam, że mgr Pawlicki wniósł oryginalny wkład w dziedzinę bezpieczeństwa sieciowego. W ocenie recenzenta główne cele pracy sformułowane w jej wstępie zostały osiągnięte. Autor na podstawie aktualnego stanu wiedzy zawartego w literaturze oraz własnych doświadczeń i przemyśleń, w swojej rozprawie doktorskiej dokonał systematycznej analizy wykorzystania technik uczenia maszynowego do detekcji ataków sieciowych (wraz z przebadaniem wielu „zewnętrznych" i „wewnętrznych" elementów mających wpływ na efektywność tych metod) oraz możliwych zagrożeń dla tego typu rozwiązań. Następnie, na tej podstawie opracował koncepcje nowych i ulepszonych metod detekcji oraz przeprowadził badania eksperymentalne.

Do najistotniejszych osiągnięć ocenianej rozprawy zaliczyć należy:

1) wykorzystanie sieci typu Gated Recurrent Unit (GRU) jako ekstraktora cech oraz ich przebadanie w wielu wariantach,
2) systematyczne przebadanie wpływu doboru hiperparametrów sieci na uzyskane wyniki detekcji,
3) zbadanie wpływu różnych metod balansowania danych na wyniki osiągane przez techniki uczenia maszynowego w tym zbiorze,
4) połączenie wszystkich wcześniej przebadanych komponentów (tj. balansowanie, optymalizacja hiperparametrów oraz GRU) w jedno rozwiązanie detekcji ataków sieciowych,
5) propozycja autorskiego rozwiązania wykrywania ataków typu *adversarial learning* z grupy *evasion* oraz systematyczna analiza jego skuteczności.

Ponadto, co warte podkreślenia, rezultaty zawarte w rozprawie zostały przedstawione i opublikowane przez Doktoranta w postaci współautorskich artykułów naukowych. Na dorobek ten składa się siedem publikacji o zasięgu międzynarodowym (w tym w trzech

występuje on jako pierwszy autor, a w czterech jako drugi), w tym dwie prace opublikowano w czasopismach z tzw. listy filadelfijskiej a jedną w materiałach konferencji z listy CORE B. Są to:

1. Rafał Kozik, Marek Pawlicki, Michał Choraś, Witold Pedrycz: Practical Employment of Granular Computing to Complex Application Layer Cyberattack Detection. Complexity 2019: 5826737:1-5826737:9 (2019) – (czasopismo o IF: 2.591)
2. Rafał Kozik, Marek Pawlicki, Michał Choraś: Cost-Sensitive Distributed Machine Learning for NetFlow-Based Botnet Activity Detection. Security and Communication Networks 2018: 8753870:1-8753870:8 (2018) – (czasopismo o IF: 1.376)
3. Marek Pawlicki, Michał Choraś, Rafał Kozik: Recent Granular Computing Implementations and its Feasibility in Cybersecurity Domain. ARES 2018: 61:1-61:6 (konferencja CORE B)
4. Michał Choraś, Marek Pawlicki, Rafał Kozik: The Feasibility of Deep Learning Use for Adversarial Model Extraction in the Cybersecurity Domain. IDEAL (2) 2019: 353-360
5. Marek Pawlicki, Rafal Kozik, Michal Choraś: Artificial Neural Network Hyperparameter Optimisation for Network Intrusion Detection. ICIC (1) 2019: 749-760
6. Marek Pawlicki, Adam Marchewka, Michał Choraś, Rafał Kozik: Gated Recurrent Units for Intrusion Detection. IP&C 2019: 142-148
7. Rafał Kozik, Marek Pawlicki, Michał Choraś: Sparse Autoencoders for Unsupervised Netflow Data Classification. IP&C 2018: 192-199

Dorobek ten jak na stopień rozwoju kariery mgr Pawlickiego uważam za ważny i znaczący. Dodatkowo, co wartym podkreślenia recenzowana rozprawa powstała w trakcie realizacji trwających, międzynarodowych projektów badawczych w ramach Programu Horyzont 2020: *Infrastress* i *Sparta*, w których Doktorant aktywnie uczestniczy.

6. Jakie są słabe strony rozprawy i jej główne wady?

Rozprawa jako całość nie ma istotnych wad, choć od strony edytorskiej mogła by być napisana lepiej, a jej struktura mogłaby być bardziej przejrzysta. Niemniej jednak w trakcie jej czytania nasuwają się pewne uwagi o charakterze dyskusyjnym:

- Tematyka rozprawy dotyczy technik uczenia maszynowego i tematów pokrewnych, ale w pracy brak jest ogólnego wstępu związanego z tą tematyką zawierającego informacje wprowadzające np. o ewolucji rozwiązań uczenia maszynowego, ich klasyfikację, przedstawienie podstaw metod użytych dalej w rozprawie oraz terminów (np. *dropout*, *overfitting*, itp.).
- Struktura pracy – Doktorant zdecydował się wydzielić dwie główne części rozprawy – jedna jest związana z propozycją detekcji ataków sieciowych wykorzystujących techniki uczenia maszynowego, a w drugiej omawiane są zagadnienia związane z atakami wpływającymi na efektywność takich rozwiązań. Takie podejście należy uznać za prawidłowe. Niemniej jednak dalszy podział na rozdziały i podrozdziały mógłby być lepszy. Przykładowo, rozdział dotyczący propozycji nowego detektora ataków typu *evasion* zawiera 10 stron i opisano w nim m.in. metodykę przeprowadzonych badań, a na jego końcu Autor odnosi się do wyników badań oraz umieszcza odwołania do tabeli, które znajdują się już w kolejnym bardzo krótkim (3 stronicowym) rozdziale zawierającym w zasadzie same tabele. W wyniku takiego podejścia rozprawa traci na spójności i czytelności. Innym przykładem jest rozdział 3, a w szczególności podrozdział *3.1 Proponowana metoda wykorzystująca sztuczne sieci neuronowe*, gdzie w zasadzie umieszczono głównie informacje wprowadzające dotyczące sztucznych sieci neuronowych, a sama autorska metoda opisana jest dopiero w kolejnych podrozdziałach. Ponadto, podrozdział *3.3 Metody balansowania zbiorów*

danych powinien być podrozdziałem 3.2.1, gdyż jest tematycznie powiązany z rozdziałem 3.2 (taka sama sytuacja jest w przypadku podrozdziałów 4.5 i 4.6). Dodatkowo, ciągle w podrozdziale 3.3 następują odwołania do tabel z wynikami, które zostały umieszczone dopiero w rozdziale 4, co jak wspomniano powoduje brak ciągłości myśli i spójności tej części rozprawy.

- Zaburzona kolejność definiowania pojęć i wprowadzania informacji w rozprawie. Przykładowo, na str. 37 Autor wprowadza zjawisko „znikającego gradientu" wskazując go jako wadę funkcji *Sigmoid*, jednak sama definicja i istota tego problemu zostały opisane dopiero na str. 49. W innym miejscu, w rozdziale 3.3 przedstawione są wyniki związane z balansowaniem zbiorów danych dla CICIDS2017, jednak informacje wprowadzające dotyczące tego zbioru zostały umieszczone dopiero w rozdziale 4. Także w podrozdziale 6.3.2 (str. 102), Doktorant najpierw odwołuje się pod rysunkiem 21 do czterech metod ataków bez nazwania, o które metody chodzi a wymienia je z nazwy dopiero poniżej na str. 102-103.
- W podrozdziale *3.1.2 Hiperparamtery i poprawa wyników wybranych algorytmów poprzez ich optymalizację* Doktorant pisze, że spośród szeregu dostępnych funkcji aktywacji wybiera do dalszych badań jedynie cztery (funkcję sigmoidalną i jej twardą wersję, tangens hiberboliczny oraz obciętą funkcję liniową). Nie jest jednak jasne czemu akurat te cztery funkcje aktywacji zostały wybrane – brak jest jakiejkolwiek dyskusji na ten temat oraz uzasadnienia wyboru popartego analizą literatury.
- W podrozdziale 4.5 Autor pisze: „Po szeregu eksperymentów okazało się, że żadna z bardziej złożonych metod z kategorii "undersampling" nie osiąga wyników lepszych od tych pochodzących z wykorzystania techniki "random subsampling", w wypadku zbioru CICIDS2017", jednak analiza wyników umieszczonych w tabelach 5-9 pokazuje, że nie jest to stwierdzenie prawdziwe. Przykładowo, metoda *Tomek-links* dla *Random Forest* i *Naive Bayes* osiąga wyższe wartości skuteczności (accuracy).
- Brak bądź przedstawienie w ograniczonym zakresie uzasadnienia przyjętych parametrów konfiguracyjnych:
    - W podrozdziale 4.9 dotyczącym wykorzystania architektury GRU określono parametry konfiguracyjne eksperymentów następująco: „W teście wykorzystano architekturę GRU o dwóch warstwach po 40 neuronów każda, z warstwami *dropout* ustawionymi na 0.2 zaraz po nich dla zbioru NSL-KDD, oraz podobna architekturę tylko z 78 neuronami na pierwszej ukrytej warstwie i 38 na drugiej dla CICIDS2017", jednak nie przedstawiono żadnego uzasadnienia dla wyboru tych konkretnych wartości.
    - W rozdziale 6 Doktorant pisze: „Wykorzystana sieć złożona była z dwóch warstw ukrytych po 25 neuronów każda, ReLU w roli funkcji aktywacji i ADAM jako optymizator." jednak znów nie przedstawiono argumentów za wykorzystaniem takiej właśnie konfiguracji. Taka sama sytuacja ma miejsce w rozdziale 6.3.3 w przypadku opisu konfiguracji detektora.
    - W podrozdziale 4.11 stwierdzono, że „W celu przeprowadzenia eksperymentu zbiór CICIDS2017 podzielono na 3 części", ale nie podano charakteru ani szczegółów tego podziału.
    - W podrozdziale 6.3.2 wskazano, że „Ze zbioru B losowo wybrano 1397 próbek klasy "ATTACK"(...)" – również w tym miejscu nie jest jasne czemu taka liczba próbek została wybrana.

- W rozdziale 5.4.1 podczas omawiania ataku typu Carlini and Wagner Doktorant napisał, że spośród ataków zdefiniowanych w pozycji [19] wybrano L_2 jednak nie przedstawił żadnego uzasadnienia tej decyzji.

- Powierzchowna dyskusja uzyskanych rezultatów eksperymentów:
  - W podrozdziale 4.5 dyskusja wyników eksperymentalnych przedstawiona w tabelach 4-12 zajmuje jedynie 1,5 strony. W ocenie recenzenta omówienie poszczególnych tabel powinno być bardziej dogłębne wraz z odniesieniem w tekście rozprawy do poszczególnych grup wyników (w tym wyników liczbowych) wraz z dokładną analizą porównawczą i dociekaniem przyczyn takiej postaci rezultatów. Taki sam mankament występuje także w przypadku innych podrozdziałów rozdziału 4.
  - Natomiast w rozdziale 7 przy omawianiu uzyskanych wyników pojawiają się wartości liczbowe, jednak całość dyskusji rezultatów z tabel 32-36 to jedynie około pół strony tekstu.
- W podrozdziale 5.2, na stronie 82 wyróżniono cztery możliwości znajomości sposobu działania atakowanego algorytmu uczenia maszynowego przez atakującego. Jeden z tych sposobów to „wiedza niepełna", ale w pierwszych dwóch grupach (wiedza o zbiorze danych oraz wiedza reprezentacji cech) dopuszcza się wiedzę częściową (więc także niepełną). Jaka jest zatem różnica między tymi sposobami?
- W rozdziale 6, gdzie przedstawiono procedurę ekstrakcji modelu do przeprowadzenia badań eksperymentalnych użyto jedynie zbiór danych NSL-KDD, który Doktorant we wcześniejszej części rozprawy (podrozdział 6.1.1) sam ocenił jako przestarzały. Skąd taki wybór?
- W rozdziale 8 przedstawiono wnioski z rozprawy, jednak w ocenie recenzenta brakuje tam wskazania dalszych możliwych kierunków rozwoju prac badawczych w tej dziedzinie, które Doktorat mógłby/planuje podjąć.

W rozprawie można odnaleźć także niedociągnięcia edytorskie bądź językowe. Przykładowo:

- Nieprecyzyjne tłumaczenie terminów bądź ich błędne użycie: przykładowo, skrót IDS (Intrusion Detection System) w literaturze polskojęzycznej tłumaczy się zazwyczaj jako System Wykrywania Włamań lub System Wykrywania Intruzów, a nie jako System Wykrywania Ataków jak to użyto w pracy. Ponadto, ataki skierowane przeciwko konkretnej osobie (str. 10) określa się terminem „spear phishing", a nie „spear fishing".
- Używanie wyrażeń żargonowych lub tłumaczenia „na siłę" z języka angielskiego, przykładowo: „Silnik korelacji to narzędzie analityczne które jako wkład pobiera..." a powinno być raczej: „Silnik korelacji to narzędzie analityczne które na wejściu pobiera...". Inne przykłady: „brana pod uwagę jest tylko garstka prostych ataków...", „Restartowanie analizy od zera...", „Ewaluacja skuteczności", „lepsze osiągi...", „w celu mitygacji...", „na powierzchnię wypłynął", „relewantne", itp.
- Występujące błędy językowe – przykładowo: jest „Przewaga podejścia opartego o sygnatury" zamiast „Przewaga podejścia opartego na sygnaturach", „wyrażana jak ukazano..." zamiast „wyrażana jak zawarto/przedstawiono...", „stale zyskuje na wadze", zamiast „stale zyskuje na znaczeniu", „lub lepsza co inne" zamiast „lub lepsza niż inne".
- Literówki, przykładowo, w tezie pracy jest „wiarydgodnosi", a powinno być „wiarygodności", na stronie 32 „zdecydowania większość badań" zamiast

„zdecydowana większość badań", natomiast na stronie 85 umieszczono „Zadanie jest uważane z wykonane", a powinno być „Zadanie jest uważane za wykonane".

- Str. 26 – w trakcie omawiania artykułu [79] Doktorant wprowadza oznaczenia CNN_SEQ, CNN_IMG, czy CNN_ASM bez ich wcześniejszego wyjaśnienia.
- Odwołania do nieistniejących rysunków np. na str. 40 i 97 znajduje się odwołanie do rys. 6.1.3.
- Na str. 40 dwukrotnie w tym samym miejscu cytowana jest para cytowań: [79][82].
- W całej rozprawie Doktorant zamiast terminu „Tabela" używa słowa „Tablica".
- Nie jest jasne czemu podpis rysunku 7 odnosi się jedynie do klasy "BENIGN", przecież inne klasy też uległy zmianie po przeprowadzeniu procedury „Random Subsampling".
- Czemu kolorystyka użyta do oznaczenia klas ruchu sieciowego na rysunku 8 jest zupełnie inna niż na rysunkach 5-7 zawierających te same klasy? Analiza wyników byłaby łatwiejsza, gdyby na rysunkach 5-10 użyto tych samych kolorów.
- Użycie różnych separatorów dziesiętnych w tekście pracy – często wykorzystywany jest przecinek dziesiętny, ale także w wielu miejscach pojawia się używana przede wszystkim w krajach anglosaskich kropka.
- Na str. 57 podano, że liczba klas w zbiorze danych CICIDS2017 wynosi 13, ale z analizy rysunków 5-10 wynika, że jest ich jednak 14. Skąd wynika ta różnica?
- Na str. 58 napisano „Z wyników "recall" niezbalansowanego zbioru (tab. 3)", jednak wydaje się, że chodzi tu w rzeczywistości o odwołanie do tabeli 4.
- Na str. 67 umieszczono informację o podzbiorze „Wtorek" jednak wcześniej w pracy nie opisano, co się na niego składa.
- Nie jest jasne czemu w tabeli 20 Doktorant zdecydował się przyjąć inną kolejność kolumn niż w tabeli 15 – utrudnia porównanie znajdujących się tam danych.
- Cytowania do zbiorów MNIST oraz CIFAR-10 pojawiają się w rozprawie wielokrotnie (a powinny pojawić się jedynie podczas ich pierwszego wprowadzenia).
- Na str. 82 Autor pisze: „Obecne rozwiązania zaproponowane przez badania", a powinno być „Obecne rozwiązania zaproponowane w dotychczas przeprowadzonych pracach naukowych".
- Na str. 87 stwierdzono „Ataki typu *exploratory* można podzielić na następujące podkategorie:" a powinno być „Ataki typu *exploratory black-box* można podzielić na następujące podkategorie:".
- W rozdziale 5.3 wprowadza się podgrupy ataków typu *extraction*, jednak całkowicie pominięto ich istotę i opis.
- Pozycja literatury [30] odwołująca się do zbioru danych KDD jest nieprawidłowa.

7. Czy rozprawa świadczy o dostatecznej wiedzy autora i znajomości współczesnej literatury z zakresu dyscypliny naukowej, jakiej rozprawa dotyczy?

Tak. Jak już wspomniano, szczegółowa analiza współczesnej literatury w zakresie wykorzystania metod uczenia maszynowego do detekcji ataków sieciowych została zawarta w rozdziale 2, natomiast zagadnienia wykrywania ataków typu *adversarial learning* na tego typu rozwiązania umieszczono w rozdziale 5. Doktorant wykazał się wystarczającą

znajomością stanu wiedzy w dziedzinie będącej przedmiotem pracy oraz umiejętnością analizy literatury i poprawnego formułowania wniosków na jej podstawie. Dowodzi to, że ma duże rozeznanie w dotychczas przeprowadzonych badaniach dotyczących zarówno detekcji ataków sieciowych z wykorzystaniem technik uczenia maszynowego jak i ataków na tego typu rozwiązania oraz sposobów im przeciwdziałania.

Rozprawa zawiera 137 pozycji bibliografii. Warto jednak zauważyć, że 9 pozycji na liście cytowanej literatury znajduje się tam dwukrotnie (np. [21] i [22], [40] i [41], [53] i [54], itp.).

Pewnym mankamentem tej części rozprawy jest jednak brak jednoznacznego scharakteryzowania unikalnych cech proponowanych przez Doktoranta rozwiązań na tle dotychczas przeprowadzonych badań znanych z literatury. Przykładowo, w podrozdziale 2.3 omówionych zostało wiele prac naukowych z zakresu wykrywania ataków sieciowych z wykorzystaniem technik uczenia maszynowego, które podsumowano w podrozdziale 2.4 i wskazano ich ograniczenia, natomiast w żaden sposób nie wskazano, że to proponowane przez Doktoranta w rozprawie rozwiązanie pozwala te niedostatki zniwelować.

Dobrym uzupełnieniem przedstawionego podsumowania obu przeglądów literatury (z części 1 i 2) byłyby także tabele, w których pokazano by w przejrzysty sposób poszczególne osiągnięcia rozwiązań zaproponowanych dotychczas i na tym tle podkreślono nowatorskie aspekty wprowadzonego przez autora rozprawy rozwiązania.

8. Czy autor wykazał umiejętność poprawnego i przekonywującego przedstawienia uzyskanych przez siebie wyników (zwięzłość, jasność, poprawność redakcyjna rozprawy)?

Tak. Doktorant w rozprawie wykazał się umiejętnością przedstawienia uzyskanych wyników badawczych oraz poprawnego wyciągania wniosków z nich płynących. Jak jednak wspomniano już w sekcji 6 niniejszej recenzji zasadniczo dyskusja uzyskanych rezultatów mogłaby być bardziej pogłębiona, a parametry konfiguracyjne wykorzystane w prowadzonych eksperymentach lepiej umotywowane.

9. Czy i jaka jest przydatność rozprawy dla gospodarki narodowej?

Praktyczna przydatność rozprawy dla nauk technicznych jak także dla gospodarki narodowej jest potencjalnie duża. Detekcja ataków sieciowych ze względu na dynamiczną ewolucję zagrożeń nie jest zagadnieniem trywialnym i tym bardziej wartościowe są rozwiązania, w tym te zaproponowane w rozprawie, wykorzystujące techniki uczenia maszynowego umożliwiające ich wykrywanie.

Z tej perspektywy proponowanie i analizowanie własności nowych sposobów wykrywania zagrożeń sieciowych jest kluczowe dla zapewnienia odpowiedniego poziomu bezpieczeństwa w sieciach firmowych, wojskowych, operatorów telekomunikacyjnych, instytucji/organizacji oraz zwykłych użytkowników. Warto także podkreślić, że to właśnie rozwiązania oparte na technikach uczenia maszynowego uznaje się obecnie za rozwiązania przyszłościowe i z największym potencjałem dla zapewnienia cyberbezpieczeństwa w nowoczesnych sieciach teleinformatycznych. Z drugiej strony ważnym jest także zapewnienie, żeby tego typu mechanizmy nie były podatne na manipulację ze strony atakującego. W obu wymienionych powyżej zagadnienia rozprawa mgr Pawlickiego proponuje nowatorskie podejścia oraz przedstawia wyniki badań potwierdzające ich efektywność, zatem w tym aspekcie przedłożoną rozprawę należy ocenić wysoko.

10. Czy rozprawa spełnia wymagania stawiane rozprawom doktorskim przez obowiązujące przepisy?

Powyżej przedstawione uwagi merytoryczne oraz redakcyjne nie mają istotnego wpływu na jakość oraz wagę przedstawionych rozwiązań i nie obniżają znacząco wartości pracy. Przedstawione przez Doktoranta zagadnienia badawcze zostały ujęte wystarczająco szczegółowo, a uzyskane wyniki są znaczące oraz potwierdzają osiągnięcie z powodzeniem założonych w rozprawie celów.

Biorąc pod uwagę zaprezentowany dorobek naukowy Doktoranta, a w szczególności jego dorobek publikacyjny, na którym bazuje rozprawa (7 publikacji o zasięgu międzynarodowym, w tym dwie publikacje w czasopiśmie z tzw. listy filadelfijskiej i jedna w materiałach konferencji z listy CORE B) uważam, że recenzowana praca spełnia w stopniu bardzo dobrym wymagania stawiane rozprawom doktorskim przez obowiązującą ustawę o stopniach i tytule naukowym.

dr hab. inż. Wojciech Mazurczyk, profesor uczelni