dr hab. Jarosław Bylina
Instytut Informatyki
Uniwersytet Marii Curie-Skłodowskiej
Pl. M. Curie-Skłodowskiej 5
20-031 Lublin
email: jaroslaw.bylina@umcs.pl

Lublin, 10 grudnia 2023

Recenzja rozprawy doktorskiej

Tytuł rozprawy:	**Samonadzorujące się uczenie w czasie rzeczywistym dla wykrywania włamań w bezpiecznym Internecie Rzeczy**
Autor rozprawy:	**mgr Mert Nakıp**
Promotor rozprawy:	**prof. dr Erol Gelenbe**
Dziedzina:	**nauki inżynieryjno-techniczne**
Dyscyplina:	**informatyka techniczna i telekomunikacja**

1 Tematyka rozprawy i jej aktualność

Pan magister Mert Nakıp zajął się w swojej pracy problematyką bezpieczeństwa specyficznego, ale bardzo ważnego i prężnie rozwijającego się obszaru Internetu, mianowicie Internetu Rzeczy (IoT).

Ze względu na tę właśnie specyfikę działania i zastosowania, urządzenia IoT są małe, tanie, prawie bezobsługowe. Za tym idzie ich względnie niskie bezpieczeństwo — w porównaniu do bardziej tradycyjnych obszarów Internetu. Co więcej, opracowanie skutecznych metod ochrony przed włamaniami do systemów IoT napotyka wiele problemów — takich jak:

- stosunkowo niskie moce obliczeniowe i pojemności pamięci pojedynczych urządzeń, co utrudnia stosowanie skomplikowanych algorytmów;
- różnorodność rozwiązań i konkretnych zastosowań, a więc potrzeba dostosowania zabezpieczeń do konkretnych sytuacji.

Tematyka rozprawy jest więc bardzo aktualna. Rozwój IoT, a także coraz szersze jego stosowanie w każdym niemal aspekcie życia wymusza zadbanie o bezpieczeństwo tego rodzaju systemów.

2 Problem naukowy sformułowany w rozprawie

Głównym celem pracy było zbadanie możliwości zastosowania uczenia się w czasie rzeczywistym do systemów wykrywania włamań (IDS) i użycia go w urządzeniach IoT.

Autor przy tej okazji chciał także stworzyć lekki, prosty do wdrożenia algorytm wykrywania włamań oparty na samonadzorującym się uczeniu maszynowym działającym w pełni w czasie rzeczywistym (online), ale także z możliwością uczenia na danych zebranych uprzednio (offline) oraz całkowicie bez jakiejkolwiek interwencji ludzkiej.

Problem podejmowany przez Autora jest sformułowany bardzo dobrze, przejrzyście i bardzo trafnie w świetle bieżących zainteresowań informatyki technicznej i telekomunikacji. Rzeczony problem rozwiązany został poprawnie, przy użyciu właściwych do tego celu metod.

3 Zawartość i charakter rozprawy

Rozprawa składa się z 6 rozdziałów. Pierwszy z nich jest krótkim wstępem, w którym zawarte jest wprowadzenie do pracy, a także jej cele, wraz z uzasadnieniem oraz listą własnych prac opublikowanych w czasopismach i na konferencjach. Rozdział 2 poświęcony jest przeglądowi zagadnień związanych ogólnie z cyberbezpieczeństwem Rozdział 3 to przegląd sposobów wykrywania włamań/ataków w aspekcie Internetu Rzeczy. Rozdział 4 stanowi pierwszą część autorskiego wkładu Doktoranta, mianowicie omówienie systemu działającego na danych zebranych (offline, quasi-online). Kolejny rozdział to natomiast

opis systemu działającego w czasie rzeczywistym (online), który jest częścią drugą wkładu Autora. Ostatni rozdział jest podsumowaniem pracy.

Rozprawa jest więc samodzielnym dziełem, w którym Doktorant opisuje autorskie rozwiązanie konkretnego problemu powołując się także na swoje publikacje w międzynarodowych czasopismach naukowych.

4 Oryginalny wkład Autora w dyscyplinę

Najważniejsze osiągnięcia autorskie Doktoranta przedstawione w rozprawie to:

- stworzenie pamięci autoasocjacyjnej oraz procedur uczenia maszynowego dla niej na danych zebranych (offline, quasi-online), a także klasyfikacji szkodliwego ruchu z użyciem tejże pamięci;
- przedstawieniu miar statystycznych szacujących zdolności powstałego systemu;
- stworzenie środowiska opartego na samonadzorującym uczeniu maszynowym w czasie rzeczywistym dostosowanego do wykrywania ataków w systemach urządzeń IoT poprzez eliminację gromadzenia danych, wyłączenie interwencji człowieka, przystosowanie do warunków ruchu w sieciach urządzeń IoT;
- użycie powstałego narzędzia do detekcji szkodliwego ruchu oraz do identyfikacji przejętego przez włamywacza urządzenia.

5 Analiza źródeł i zastany stan wiedzy

Bibliografia recenzowanej rozprawy doktorskiej obejmuje 220 pozycji — zacytowanych w odpowiednim kontekście. Źródła te dobrze przedstawiają bieżący stan wiedzy na tematy poruszane w pracy. W większości są to źródła nowe (ostatnie pochodzą z 2023 roku), co całkowicie naturalne (i pożądane) w tego rodzaju tematyce, ale Autor nie stroni też od cytowania kilku starszych (bardziej klasycznych) publikacji (z roku 1981 czy też 2001), co pozwala osadzić pracę w całości dorobku naukowego dyscypliny oraz dyscyplin pokrewnych i pokazuje, że zainteresowanie bezpieczeństwem w Internecie (szczególnie Internecie Rzeczy) nie jest nowe, ale i nie słabnie.

Wyczerpującym przedstawieniem stanu wiedzy zastanej są rozdziały 2 oraz 3, w którym znacząca większość z pozycji bibliografii całej pracy jest wymieniona i opisana pod kątem tematyki rozprawy. Co ważne, w rozdziale 1 Autor prezentuje swoje prace (związane z rozprawą, a także inne).

6 Znaczenie wkładu Autora

Kandydat w swej rozprawie zajmuje się opracowaniem — teoretycznym i praktycznym — oraz budową i przetestowaniem zaplanowanego autorskiego środowiska SSID (*Self-Supervised Intrusion Detection*).

w rozdziale 4 Autor opisał opracowany IDS (ang. *intrusion detection system*, system wykrywania włamań) oparty na anomaliach z uczeniem się w trybie offline i quasi-online, który potrafi wykrywać zarówno podejrzany ruch, jak i zagrożone urządzenia IoT podczas różnego rodzaju ataków. Składa się on z trzech głównych funkcji służących do wyodrębniania metryk ruchu sieciowego, szacowania oczekiwanych wartości metryk dla normalnego ruchu i podejmowania ostatecznej decyzji, czy analizowane metryki wskazują na naruszenie bezpieczeństwa.

Ten IDS obejmuje następujące nowatorskie rozwiązania:

- Aby zaobserwować zmiany w ruchu sieciowym i przechwycić sygnatury atakującego, Autor zaproponował oryginalne metryki ruchu sieciowego — specjalnie dla wykrywania złośliwego ruchu i identyfikacja zaatakowanych urządzeń. W szczególności, w przypadku wykrywania szkodliwego ruchu przedstawił trzy wskaźniki mierzące gęstość całkowitego ruchu sieciowego, natomiast w przypadku identyfikacji zaatakowanych urządzeń — sześć wskaźników mierzących gęstość ruchu odbieranego i przesyłanego przez pojedyncze urządzenie.

- Autor stworzył AADRNN — pamięć autoasocjacyjną, korzystającą z modelu DRNN — w celu oszacowania wartości metryki, jakich można spodziewać się podczas normalnej pracy rozważanej sieci. W tym celu model DRNN jest szkolony przy użyciu wyłącznie normalnych pakietów ruchu.

- W końcu, w tej części Autor pokazuje własny algorytm klasyfikatora wykrywającego zagrożenie.

Wydajność proponowanego IDS Autor ocenia pod kątem wykrywania złośliwego ruchu i identyfikacji zagrożonych urządzeń — także pod kątem wykrywania ataków nieznanych. Autor wykorzystuje publicznie dostępne zbiory danych i porównuje wydajność wydajność proponowanego IDS z sześcioma znanymi modelami ML. Według zamieszczonych wyników, proponowany IDS znacznie przewyższył istniejące metody w obu badanych sytuacjach — zarówno pod względem wykrywania szkodliwego ruchu, jak i identyfikowania zaatakowanych urządzeń.

Ponadto ta część wykazała możliwości w zakresie dalszego rozwoju — w celu stworzenia systemów IDS działających w czasie rzeczywistym — co Autor opisuje w rozdziale 5. Tutaj właśnie zaproponował on środowisko SSID, które działa w pełni w czasie rzeczywistym, bez interwencji człowieka (co bardzo ważne w środowisku urządzeń IoT).

SSID realizuje działania:

- ocenia wiarygodność decyzji o stwierdzeniu włamania;

- identyfikuje normalny i złośliwy ruch;

- wybiera i etykietuje pakiety ruchu sieciowego w sposób samonadzorowany, w oparciu wyłącznie o decyzje IDS i zaufanie SSID do tych decyzji;

- biorąc pod uwagę wiarygodność IDS, wybrane pakiety szkoleniowe i najnowszy stan bezpieczeństwa sieci, SSID określa, kiedy należy zaktualizować parametry IDS

— w ten sposób SSID eliminuje potrzebę gromadzenia danych, eliminuje potrzebę etykietowania danych (i koszty oraz błędy ludzkie związane z nim) oraz dostosowuje się do zmieniających się parametrów ruchu.

Autor ocenia także wydajność SSID w przypadku wykrywania złośliwego ruchu i identyfikacji zainfekowanych urządzeń. Wyniki pokazuję, że modele te osiągają wysoką wydajność w porównaniu z tymi samymi modelami z uczeniem się na danych zebranych i przyrostowych.

7 Redakcja rozprawy i prezentacja wyników

Struktura pracy jest uporządkowana i przejrzysta. Podział na poszczególne rozdziały jest logiczny. Bardzo wyraźnie zaznaczone są dokonania autorskie Doktoranta (wraz z wymienionymi publikacjami).

Pracę dobrze się czyta — napisana jest starannie zarówno pod względem językowym, jak i typograficznym. Szczególnie podkreślić należy dobór odpowiedniego narzędzia do składu (LaTeX) — bez niego napisanie pracy wypełnionej skomplikowanymi wzorami byłoby trudne i nie dałoby pożądanego wyniku. Bez większych zarzutów należy też odnieść się do wyglądu, opisu i czytelności ilustracji, wykresów oraz tabel, których duża liczba ułatwia czytanie i zrozumienie rozprawy. Także interpretacja wykresów zawarta w pracy jest poprawna.

Warto też pozytywnie wyróżnić znajdujące się na początku tabele skrótów i oznaczeń, które bardzo pomagają w sprawnej analizie rozdziałów dotyczących autorskich rozwiązań Kandydata.

8 Słabe strony i uwagi krytyczne

Praca ma bardzo niewiele elementów, do których można mieć wyraźne uwagi krytyczne.

- Przede wszystkim, do pełnego zrozumienia i docenienia pracy Doktoranta, brakuje w tekście rozprawy wskazania repozytorium z pełnym kodem źródłowym, co jest już pewnym standardem w nauce (ze względu na konieczność powtarzalności doświadczeń). Bardzo niewiele dowiadujemy się z samej pracy o realizacji algorytmów w języku programowania (poza zdawkową uwagą o implementacji w Pythonie), o użytych bibliotekach itd. Ten brak pozostawia duży niedosyt w czytelniku, bowiem możliwość analizy kodu (nie mówiąc o jego uruchomieniu) rzuca zwykle bardzo dużo światła na pracę.

- O ile Autor bardzo dobrze wprowadza czytelnika w zagadnienia cyberbezpieczeństwa i Internetu Rzeczy właściwie od podstaw, o tyle traktuje zagadnienia związane z podstawami uczenia maszynowego jako rzecz oczywistą. Niewielki dodatkowy rozdział (np. między rozdziałem 3 a 4) na ten temat byłby bardzo dobrym uzupełnieniem całości.

- Brakuje w rozprawie bardziej szczegółowych rozważań dotyczących sprzętu IoT, na jaki jest przeznaczone stworzone środowisko. Ze względu na specyfikę urządzeń IoT, wydaje się istotnym dokładniejsze ustosunkowanie się do konkretnego sprzętu i ocena, jak w warunkach polowych wygląda wydajność i skuteczność opracowanych algorytmów.

- Bardzo drobnymi problemami są nieliczne błędy interpunkcyjne (braki kropek po niektórych wzorach kończących zdanie) czy typograficzne (jak niewłaściwie dobrane rozmiary nawiasów, zapis kursywą niektórych nazw funkcji matematycznych we wzorach, czy też użycie symbolu '*sign*' zamiast ogólnie przyjętego 'sgn' [bez 'i' oraz bez kursywy]).

Powyższe uwagi krytyczne nie wpływają na merytoryczną wartość pracy w żaden sposób, a są jedynie uchybieniami zrozumiałymi przy tego rodzaju i wielkości pracy, a także subiektywnym zdaniem recenzenta.

9 Podsumowanie i wniosek końcowy

Po analizie rozprawy Doktoranta, mogę stwierdzić, że jest ona przygotowana rzetelnie i wnosi znaczący wkład w dyscyplinę *informatyka techniczna i telekomunikacja*. Potwierdza ona też zdolność Kandydata do prowadzenia dalszej pracy naukowej samodzielnie. Świadczy ona o ugruntowanej ogólnej wiedzy Autora w zakresie nauk inżynieryjno-technicznych i szczegółowej wiedzy odpowiadającej zakresowi badań.

Stwierdzam, że recenzowana rozprawa pt. „Samonadzorujące się uczenie w czasie rzeczywistym dla wykrywania włamań w bezpiecznym Internecie Rzeczy" spełnia warunki określone w Ustawie *Prawo o szkolnictwie wyższym i nauce*. W związku z tym, wnioskuję o dopuszczenie rozprawy doktorskiej Pana mgr. Merta Nakipa do publicznej obrony i dalszych etapów postępowania doktorskiego.

Jarosław Bylina

Jarosław Bylina