Instytut Informatyki Teoretycznej i Stosowanej
Polskiej Akademii Nauk

Rozszerzone Streszczenie Dysertacji Doktorskiej

Samo-Nadzorujące się uczenie w czasie rzeczywistym dla wykrywania włamań w bezpiecznym Internecie Rzeczy

MERT NAKIP

PROMOTOR: PROF. DR. EROL GELENBE

Gliwice, Polska
2023

Streszczenie

Systemy Internetu Rzeczy są budowane ze stosunkowo prostych urządzeń o niskiej mocy obliczeniowej. Dlatego trudno implementować w nich złożone algorytmy metod bezpieczeństwa. Preferowane są mniej pracochłonne zabezpieczenia, zwłaszcza wykorzystujące mechanizmy uczenia maszynowego (ML). Proces uczenia wykonywany jest często offline z wykorzystaniem dużych zbiorów danych zebranych w czasie symulacji, co może być czasochłonne i dostarczać niepoprawne, mylące wyniki. Niniejsza praca bada otwarte kwestie w tej dziedzinie i proponuje sposoby umożliwiające proste i w pełni online i uczenie się wykrywania intruzów oparte na ML, torując drogę ku bezpiecznemu IoT.

Najpierw rozwijany jest System Wykrywania Intruzji (IDS) uczący się normalnych wzorców ruchu sieci IoT i wykrywający zarówno złośliwe pakiety ruchu sieciowego, jak i skompromitowane urządzenia. System opiera się na modelu Deep Random Neural Network (DRNN) w połączeniu z oryginalnie zaproponowanymi metrykami ruchu oraz klasyfikatorze opartym na statystyce Whiskera (SWBC). W każdym przypadku wykrycia złośliwego ruchu i identyfikacji skompromitowanego urządzenia proponujemy zestaw oryginalnych metryk ruchu sieciowego umożliwiających dokładne rozpoznawanie wzorców ruchu Botnet oraz śladów atakującego. Opracowano także nowy algorytm SWBC klasyfikujący pakiety jako łagodne i złośliwe, ucząc się kryterium klasyfikacji na podstawie wyników DRNN na danych treningowych. Przedstawiono także offline i quasi-online (inkrementalne i sekwencyjne) algorytmy uczenia dla naszego IDS.

Następnie oceniamy wydajność naszego IDS zarówno z algorytmami uczenia offline, jak i quasi-online dla ataków Botnet DDoS, DoS oraz ataków „dnia zerowego" na trzech dostępnych publicznie zbiorach danych. Wyniki pokazują bardzo dobrą wydajność i niski czas obliczeń naszego IDS w porównaniu z dobrze znanymi modelami ML. Wyniki ujawniają także potencjał uczenia online w wykrywaniu intruzów.

Następnie, aby umożliwić w pełni online uczenie się ML oparte na IDS bez interwencji ludzi, proponujemy nowe podejście: samonadzorujące się wykrywanie intruzów (SSID). W celu nauczenia wykorzystywanego IDS, ramka SSID zbiera i etykietuje pakiety ruchu opierając się wyłącznie na decyzjach IDS i ich statystycznie mierzonej wiarygodności. Ramka SSID umożliwia IDS szybkie dostosowywanie się do zmieniających się cech ruchu sieciowego, eliminuje potrzebę zbierania danych offline, zapobiega ludzkim błędom w etykietowaniu danych i unika kosztów pracy związanych ze szkoleniem modelu i doświadczalnym zbieraniem danych. Dlatego - jak sugerują też wyniki eksperymentalne na publicznych zbiorach danych dotyczących złośliwego ruchu i wykrywania skompromitowanych urządzeń przy użyciu dobrze znanych modeli ML, SSID jest bardzo użyteczny dla rozwijania dla systemów IoT uczenia się online ML opartego na IDS.

Publikacje Autora

Poniżej wymieniono publikacje autora, które zostały uwzględnione w tej pracy lub powstały na jej podstawie. Rozszerzona lista publikacji autora, w tym tych opublikowanych podczas studiów doktoranckich autora, ale niezwiązanych bezpośrednio z treścią tej pracy, można znaleźć w pracy.

Artykuły w czasopismach:

- M. Nakıp and E. Gelenbe, "Fully Online Self-Supervised Learning Framework for Machine Learning based Intrusion Detection," in *arXiv*, 2023, *Preprint*.
- E. Gelenbe and M. Nakıp, "Traffic Based Sequential Learning During Botnet Attacks to Identify Compromised IoT Devices," in *IEEE Access*, vol. 10, pp. 126536-126549, 2022.

Artykuły konferencyjne:

- E. Gelenbe and M. Nakıp, "Real-Time Cyberattack Detection with Offline and Online Learning," *2023 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN)*, London, United Kingdom, 2023, pp. 01-06.
- E. Gelenbe and M. Nakıp, "G-Networks Can Detect Different Types of Cyberattacks," *2022 30th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, Nice, France, 2022, pp. 9-16.
- M. Nakıp and E. Gelenbe, "Botnet Attack Detection with Incremental Online Learning," *2021 Security in Computer and Information Sciences (EuroCybersec)*, Nice, France, 2022, pp. 51-59.
- M. Nakıp and E. Gelenbe, "MIRAI Botnet Attack Detection with Auto-Associative Dense Random Neural Network," *2021 IEEE Global Communications Conference (GLOBECOM)*, Madrid, Spain, 2021, pp. 01-06.

Rozszerzone Streszczenie Rozprawy Doktorskiej

Motywacja

Większość (około 52%) urządzeń Internetu Rzeczy (IoT) to urządzenia niskokosztowe i niewymagające częstej konserwacji, dysponujące ograniczonymi zasobami. Często posiadają one luki w oprogramowaniu, albo słabe zabezpieczenia przy logowaniu, co czyni je łatwiejszym celem dla atakujących niż urządzenia obsługiwane przez użytkowników. Dlatego uważa się, że 70% wszystkich urządzeń IoT jest podatnych na ataki [1].

Liczba urządzeń IoT gwałtownie rośnie wraz z rozszerzaniem się obszarów zastosowania, prowadząc do większej liczby naruszeń bezpieczeństwa i wynikających z tego ataków cybernetycznych. Na przykład szacuje się, że przeciętny inteligentny dom jest celem dla $12,000$ ataków w ciągu jednego tygodnia. W rezultacie cyberbezpieczeństwo sieci IoT stało się jednym z głównych wyzwań, a ataki cybernetyczne na urządzenia IoT są uważane za główne zagrożenie przez 33% firm zajmujących się cyberbezpieczeństwem [2].

Zwiększenie cyberbezpieczeństwa sieci IoT jest kluczowe dla ich pracy, a System Wykrywania Intruzów (IDS) jest bardzo ważnym elementem zwiększenia cyberbezpieczeństwa systemu sieciowego, ponieważ umożliwia systemom zarządzania siecią podejmowanie wczesnych działań i reagowanie na ataki przed wystąpieniem szkód. Z drugiej strony urządzenia IoT często są wdrażane w masowych sieciach IoT [3] i działają z minimalną (jeśli nie zerową) ingerencją człowieka. Chociaż głównie sugerowane były systemowe podejścia do poprawy bezpieczeństwa systemów cyberfizycznych [4,5], jest trudno (jeśli nie niemożliwe) obciążyć proste urządzenia IoT złożonymi funkcjami bezpieczeństwa.

Można powiedzieć, że opracowanie zaawansowanych metod bezpieczeństwa, takich jak IDS oparte na uczeniu maszynowym (ML), dla urządzeń i sieci IoT stanowi ogromne wyzwanie z następujących powodów: 1) Urządzenia IoT często mają niewystarczająco małe zasoby obliczeniowe do wdrażania złożonych algorytmów. 2) Algorytmy oparte na danych (np. ML) wymagają dużych ilości danych, które są trudne do zebrania, ponieważ wymagają dużo pracy, wysokich kosztów rozwoju i długiego czasu wdrażania. 3) Zaawansowane metody bezpieczeństwa są głównie dostosowywane do indywidualnego systemu lub sieci, do którego są stosowane, ponieważ ich parametry są bezpośrednio optymalizowane dla tego systemu. W rezultacie, gdy te algorytmy muszą być wdrażane dla nowego systemu, znaczna część pracy musi być powtarzana ręcznie.

Wkład Rozprawy Doktorskiej

Głównym celem niniejszej pracy jest badanie uczenia się online dla IDS opartego na ML w kierunku tworzenia bezpiecznych systemów IoT. Nasze badania mają na celu rozwiązanie powyższych problemów i

dostarczenie lekkiego, łatwego do wdrożenia algorytmu do wykrywania intruzów, który jest jedną z kluczowych metod zapewnienia bezpieczeństwa.

1. Proponujemy nowatorskie zasady uczenia się nadzorowanego dla IDS opartego na ML, nazywane Frameworkiem do Wykrywania Intruzów z Samouczącym się Nadzorem (SSID), który umożliwia całkowicie online uczenie się parametrów IDS bez potrzeby ingerencji człowieka. W ramach frameworku SSID statystycznie mierzymy wiarygodność IDS opartego na ML, biorąc pod uwagę jego zdolność do uogólniania oraz pakiety ruchu, które IDS nauczyło się rozpoznawać. W tym celu prezentujemy miary do oszacowania zdolności uogólniania i reprezentatywności nauczonego ruchu.

   Główne zalety frameworku SSID to:

   - umożliwia IDS łatwe dostosowywanie się do zmiennych w czasie charakterystyk ruchu w sieci,
   - eliminuje potrzebę zbierania danych w trybie offline,
   - zapobiega ludzkim błędom w etykietowaniu danych, oraz
   - unika kosztów pracy związanych z szkoleniem modelu i zbieraniem danych poprzez eksperymenty.

2. Rozwijamy lekki IDS oparty na ML przy użyciu Głębokiej Losowej Sieci Neuronowej (DRNN). Poprzez oryginalnie zdefiniowane pomiary ruchu, IDS uczy się normalnych (nieszkodliwych) wzorców ruchu, a następnie identyfikuje nieprawidłowe zmiany ruchu, które mogą wskazywać na możliwy atak. W tym celu:

   - Określamy oryginalne metryki, które łatwo obliczyć, używając jedynie informacji nagłówkowych z pakietów ruchu, oraz są wysoce skuteczne w analizie wpływów ataków Botnet na ruch sieciowy oraz w przechwytywaniu sygnatur atakującego.
   - Opracowujemy trzy procedury uczenia dla opracowanego IDS dla uczenia offline, sekwencyjnego i przyrostowego.
   - Opracowujemy algorytm klasyfikacji, nazywany Łagodnym Klasyfikatorem opartym na Statystycznym Wąsiku (SWBC), który identyfikuje złośliwy ruch, porównując rzeczywisty ruch z oczekiwanym ruchem szacowanym przez pamięć autoasocjacyjną.

3. W końcu opracowujemy nowy system do identyfikacji skompromitowanych urządzeń IoT (botów) oparty jedynie na ruchu w sieci, bez potrzeby dostępu do stanu urządzenia czy treści wiadomości. Wraz z złośliwym ruchem, kluczowe jest zidentyfikowanie skompromitowanych urządzeń, aby utorować drogę do skutecznego zapobiegania rozprzestrzenianiu się ataku lub złagodzenia jego wpływów na sieć.

Wyniki Badań

System Wykrywania Intruzów z Uczeniem Offline i Kwazi-Online:

W niniejszej pracy dyplomowej najpierw opracowujemy IDS oparty na wykrywaniu anomalii z uczeniem offline i kwazi-online w celu wykrywania zarówno złośliwego ruchu, jak i skompromitowanych

urządzeń IoT podczas ataków Botnet lub ataków typu zero-day. Ten IDS, przedstawiony na Rysunku 1, składa się z trzech głównych funkcji: ekstrakcji metryk ruchu sieciowego, szacowania oczekiwanych wartości metryk dla normalnego ruchu "nieszkodliwego", oraz podejmowania ostatecznej decyzji o tym, czy analizowane metryki wskazują na włamanie.

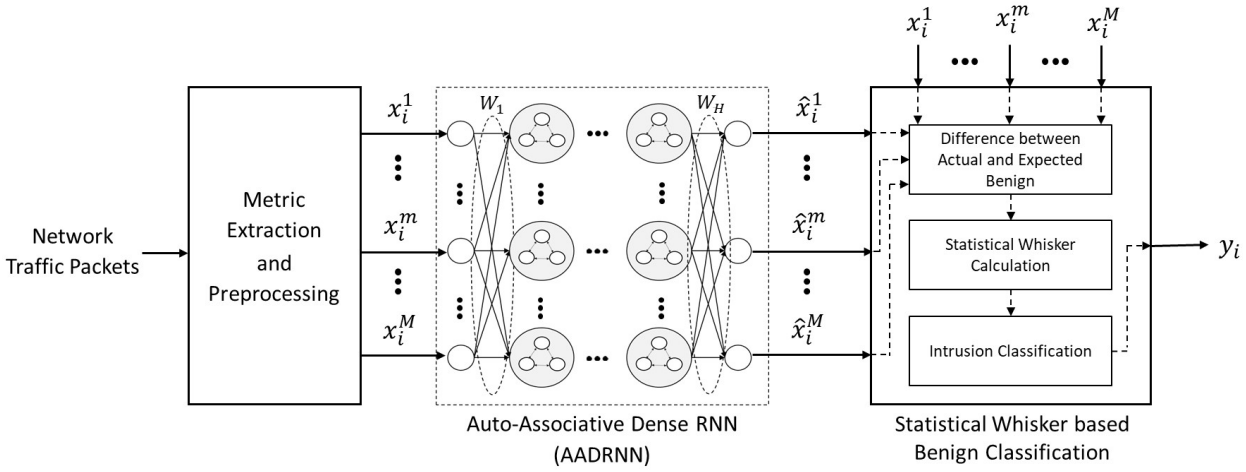

Figure 1: Architektura detektora ataków opartego na DRNN z jego trzema modułami: Ekstrakcja Metryk i Przetwarzanie Wstępne, Auto-Asocjacyjna Głęboka Losowa Sieć Neuronowa (AADRNN) oraz Klasyfikacja Benign oparta na Statystycznym Wąsiku

Aby zaobserwować wpływ intruzji na ruch w sieci i przechwycić sygnatury atakującego, proponujemy oryginalne metryki ruchu sieciowego specjalnie dla każdego z zadań: wykrywania złośliwego ruchu oraz identyfikacji skompromitowanego urządzenia. W szczególności, do wykrywania złośliwego ruchu, przedstawiamy trzy metryki mające na celu pomiar gęstości całkowitego ruchu sieciowego, natomiast do identyfikacji skompromitowanego urządzenia przedstawiamy sześć metryk służących do pomiaru gęstości otrzymanego i przesłanego ruchu przez indywidualne urządzenie.

Tworzymy pamięć autoasocjacyjną przy użyciu modelu DRNN, nazywaną AADRNN, aby szacować wartości metryk, które spodziewane są podczas normalnej pracy rozważanej sieci. Aby uzyskać rzeczywiste wartości metryk z ich zaszumionych wersji, model DRNN jest szkolony wyłącznie przy użyciu normalnych pakietów ruchu. W tym celu opracowujemy algorytmy uczenia autoasocjacyjnego w trybie offline i online:

- **Algorytm Uczenia Offline** opiera się na algorytmie półnadzorowanym przedstawionym w [6]. Podczas procesu uczenia macierz wag na każdej warstwie DRNN jest określana poprzez minimalizację funkcji kosztu z regularyzacją L1 za pomocą Szybkiego Iteracyjnego Algorytmu Skurczu i Progu (FISTA) [7]. Funkcja kosztu głównie mierzy kwadratową odległość euklidesową między wektorem wejściowym a wyjściowym rozważanej warstwy.

- **Algorytm Uczenia Przyrostowego - Kwazi-Online** umożliwia korzystanie z detektora ataków bez konieczności offline'owego zbierania nieszkodliwego ruchu. W tym celu łączy algorytm uczenia półnadzorowanego opracowany w [6] z algorytmem uczenia sekwencyjnego opracowanego w [8]. Algorytm ten składa się z etapów inicjalizacji i przyrostowego uczenia kwazi-online. Etap inicjalizacji trwa przez transmisję I pakietów i jest traktowany jako zimny start proponowanego IDS. W związku z tym przesyłanie pierwszych I pakietów jest uznawane za pakiety nieszkodliwe, a wagi połączeń

AADRNN są początkowo uczone przy użyciu tych pakietów. Etap przyrostowego uczenia kwazi-online działa w oknach czasowych. Na końcu każdego okna tylko wagi połączeń warstwy wyjściowej AADRNN są aktualizowane, aby nauczyć się nieszkodliwego ruchu zgromadzonego w tym oknie.

Ostateczna decyzja jest podejmowana poprzez porównanie oczekiwanych wartości metryk (czyli wyjścia z AADRNN) z rzeczywistymi wartościami metryk. W tym celu opracowujemy nowatorski algorytm Klasyfikatora Benign opartego na Statystycznym Wąsiku, który wykrywa intruzję, jeśli rzeczywiste metryki znacząco różnią się od oczekiwanych oszacowanych metryk. Istotność różnicy, jak również wszystkie parametry algorytmu, są określane wyłącznie na podstawie próbek pakietów używanych do szkolenia, które są znane jako nieszkodliwy ruch.

Wydajność proponowanego IDS jest oceniana pod względem wykrywania złośliwego ruchu oraz identyfikacji skompromitowanych urządzeń podczas ataków Botnet, jak również podczas wykrywania ataków typu zero-day (nieznane). Podczas oceny wydajności korzystamy z publicznie dostępnych zestawów danych - mianowicie Kitsune [9], KDD Cup'99 [10] i BotIoT [11] - i porównujemy wydajność proponowanego IDS z dobrze znanymi modelami ML: Regresją Liniową (LR), Metodą Najmniejszych Absolutnych Skurczów i Selektorów Operatorów (Lasso), Regresją Najbliższych Sąsiadów (KNN), Maszyną Wektorów Nośnych (SVM), Perceptronem Wielowarstwowym (MLP) oraz Pamięcią Długookresową Krótkookresową (LSTM).

Wyniki w Tabeli 1 pokazują, że proponowany IDS (AADRNN) znacząco przewyższa inne metody, osiągając 99.82% procent prawdziwie pozytywnych i 99.98% procent prawdziwie negatywnych. Ponadto Rysunek 2 ujawnia, że czas treningu AADRNN wynosi mniej niż 0.1 sekundy, podczas gdy podejmuje decyzje online w około 0.5 μ sekundy.

Table 1: Porównanie metod wykrywania ataków pod względem dokładności, jak również odsetków prawdziwie pozytywnych, fałszywie negatywnych, prawdziwie negatywnych i fałszywie pozytywnych

Wykrywanie Ataków Metody	Dokładność	Prawdziwie Pozytywne	Fałszywie Negatywne	Prawdziwie Negatywne	Fałszywie Pozytywne
AADRNN	99.84	99.82	0.18	99.98	0.02
KNN	99.79	99.79	0.21	99.75	0.25
Lasso	99.78	99.75	0.25	99.95	0.05
Proste Progowanie	93.18	93.09	6.94	93.63	6.37

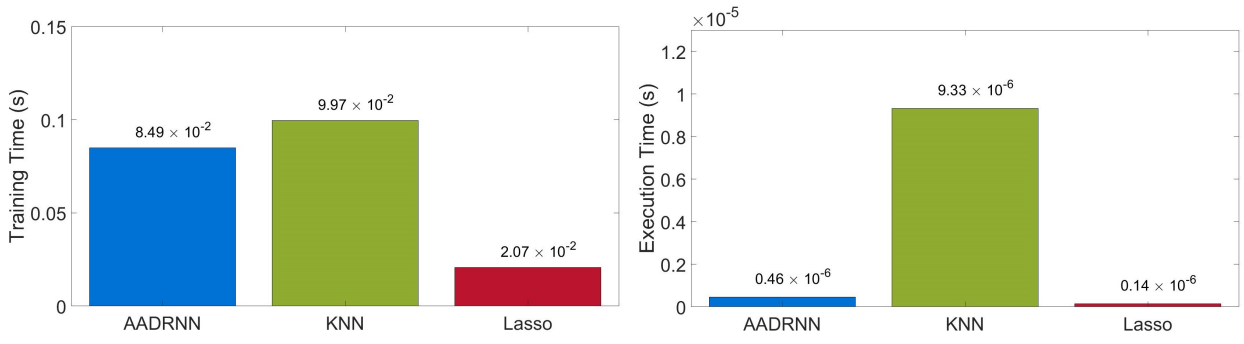

Figure 2: (lewo) Czasy szkolenia i (prawo) czasy wykonania różnych metod wykrywania ataków

Gdy proponowany IDS jest testowany pod kątem wykrywania różnych rodzajów ataków cybernetycznych jednocześnie, wyniki przedstawione na Rysunku 3 ujawniają, że 1) dokładność prognozowania wynosi ponad 98% dla 21 spośród 37 typów ataków, oraz 2) proponowany IDS znacząco przewyższa najnowszy Klasyfikator Jednoklasowy oparty na SVM (SVM-OCC).

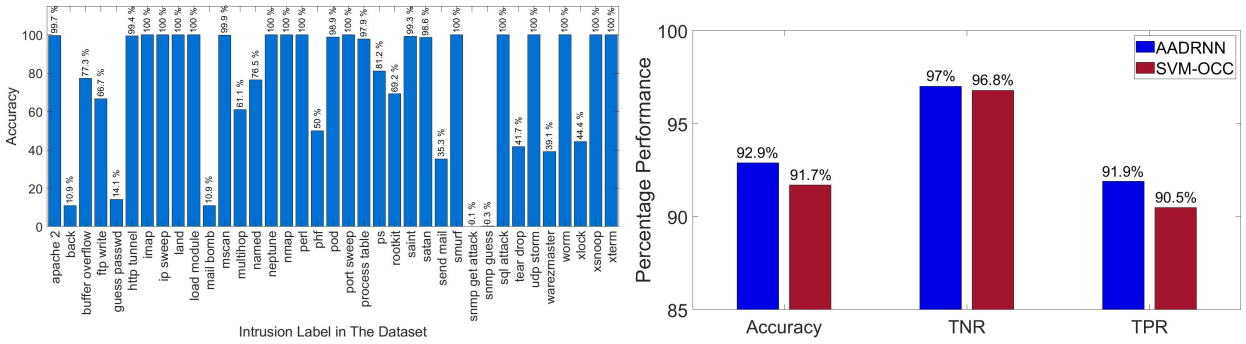

Figure 3: (lewo) Wydajność proponowanego IDS w wykrywaniu każdego typu ataku w zestawie danych KDD z uczeniem offline oraz (prawo) jego porównanie z Klasyfikatorem Jednoklasowym opartym na SVM (SVM-OCC)

Następnie porównanie między IDS z przyrostowym uczeniem kwazi-online a IDS z uczeniem offline, przedstawione na Rysunku 4, pokazuje, że IDS z przyrostowym uczeniem kwazi-online osiąga wydajność wykrywania złośliwego ruchu zbliżoną do wydajności IDS z uczeniem offline, przy znacząco mniejszej liczbie pakietów używanych do treningu. Można by powiedzieć, że uczenie kwazi-online to bardzo obiecujące podejście do wykrywania włamań w sieciach IoT.

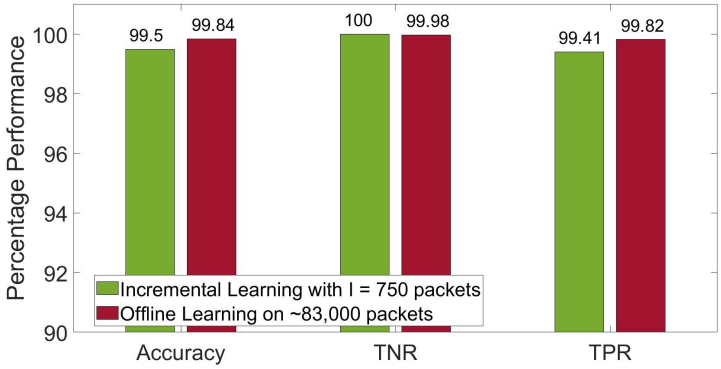

Figure 4: Porównanie wydajności IDS opartego na AADRNN w trybie Uczenia Przyrostowego z 750 pakietami a tym w trybie Uczenia Offline z około $83,000$ pakietami

Ponadto proponowany IDS do identyfikacji skompromitowanych urządzeń, nazywany CDIS, jest oceniany pod kątem różnych typów ataków DoS i DDoS dostępnych online. Wyniki wydajności na Rysunku 5 pokazują, że CDIS może skutecznie wykrywać potencjalnie złośliwe urządzenia podczas różnych typów ataków DDoS, w których złośliwe oprogramowanie rozprzestrzenia się na urządzeniach IoT.

Podsumowując, wyniki pokazują, że proponowany IDS znacząco przewyższa istniejące metody zarówno w wykrywaniu złośliwego ruchu, jak i identyfikacji skompromitowanych (zarażonych) urządzeń. Ponadto

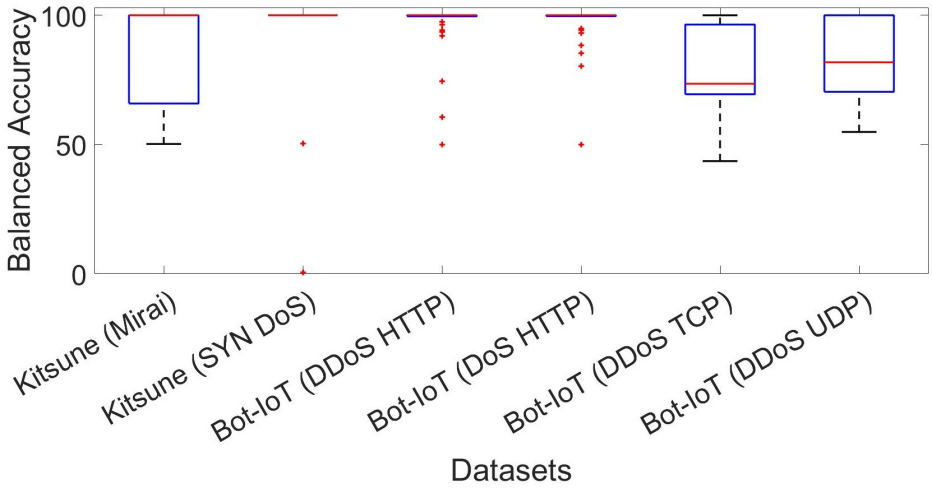

Figure 5: Wydajność proponowanego IDS w identyfikacji skompromitowanych urządzeń IoT podczas różnych typów ataków z uczeniem sekwencyjnym

algorytmy uczenia kwazi-online sekwencyjnego i przyrostowego wykazały duży potencjał w zakresie rozwoju wysoce wydajnego IDS z uczeniem online, który wymaga niskiego czasu obliczeniowego i małych danych. Z drugiej strony, IDS z uczeniem online wciąż wymaga udoskonalenia – jak to ma dostarczyć ramy Wykrywania Intruzji z Samouczącym się Nadzorem (SSID) – dla dokładniejszego wykrywania intruzji.

Samo-Nadzorujące się Uczenie w Czasie Rzeczywistym:

Jako jedno z głównych wkładów tej pracy proponujemy nowatorską ramę SSID, która jest zaprojektowana do pełnego online'owego szkolenia dowolnego danego IDS – którego parametry są obliczane przy użyciu ruchu sieciowego. Rama SSID automatycznie wybiera pakiety ruchu normalnego do uczenia się i decyduje, kiedy aktualizować parametry wykorzystywanego algorytmu. W ten sposób całkowicie eliminuje interwencję człowieka, potrzebę offline'owego zbierania danych (oznaczonych lub nieoznaczonych) oraz offline'owego treningu. W związku z tym proponowana rama różni się znacząco od istniejących prac [12–18].

Rama SSID, jak pokazano na Rysunku 6, składa się z dwóch kolejnych etapów uczenia: początkowego uczenia i uczenia online. Początkowe uczenie ma na celu szybkie dostosowanie parametrów IDS do sieci, w której IDS jest nowo wdrażany, podczas gdy uczenie online ma na celu aktualizację parametrów za każdym razem, gdy aktualizacja jest wymagana, aby zapewnić wysoką dokładność wykrywania przez IDS.

Jak można zobaczyć na Rysunku 7, podczas rzeczywistego działania IDS, równolegle do wykrywania, rama SSID wykonuje następujące główne zadania:

- Nieustannie szacuje wiarygodność decyzji dotyczących intruzji, aby identyfikować normalny i złośliwy ruch, mierząc zdolność IDS do uczenia się i generalizowania na podstawie danych dostarczonych przez SSID oraz stopień, w jakim te dane mogą reprezentować bieżące wzorce ruchu w sieci.
- Aby dostarczyć dane treningowe dla IDS, rama SSID wybiera i oznacza pakiety ruchu sieciowego w sposób samouczący się, opierając się wyłącznie na decyzjach IDS oraz zaufaniu SSID do tych decyzji.
- Biorąc pod uwagę wiarygodność IDS, wybrane pakiety treningowe oraz najnowszy stan bezpieczeństwa sieci, rama SSID określa, kiedy zaktualizować parametry IDS.

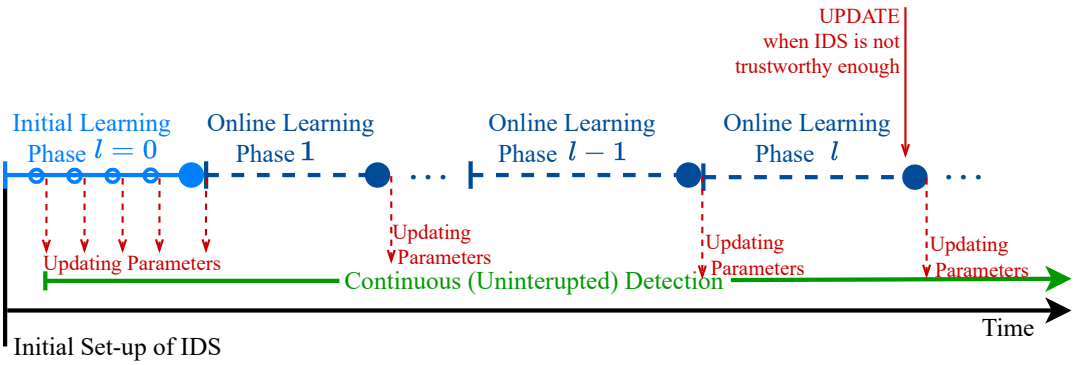

Figure 6: Procesy wykrywania i uczenia się w IDS w ramach Ramy Wykrywania Intruzji z Pełnym Online Samouczącym się Nadzorem (SSID)

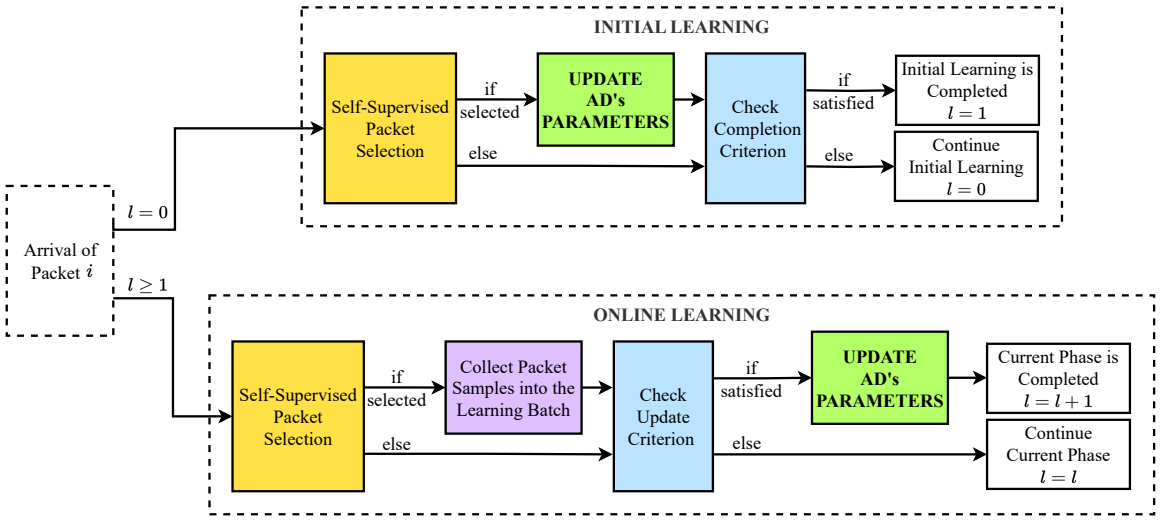

Figure 7: Schemat blokowy procesu uczenia w ramie SSID dla online'owego samouczącego się uczenia parametrów IDS

Następnie oceniamy wydajność ramy SSID dla dwóch zadań: wykrywania złośliwego ruchu oraz identyfikacji skompromitowanych urządzeń, mając na celu poprawę bezpieczeństwa sieci IoT:

Dla wykrywania złośliwego ruchu, dwa różne modele ML, DRNN i MLP, są wdrażane z ramą SSID i testowane na zestawie danych Kitsune [9]. Wyniki na Rysunku 8 ujawniają, że modele ML szkolone w ramach ramy SSID, nie wymagające offline'owego zestawu danych, osiągają znacząco wysoką wydajność w porównaniu z tymi samymi modelami z uczeniem offline.

Dla identyfikacji skompromitowanych urządzeń, wydajność CDIS jest testowana w uczeniu sekwencyjnym oraz w ramie SSID na danych 6 różnych ataków cybernetycznych dostarczanych przez dwa publiczne zestawy danych Kitsune [9] oraz Bot-IoT [11]. Wyniki na Rysunku 9 pokazały, że użycie SSID znacząco poprawia wydajność CDIS w większości przypadków.

Podsumowując, proponowana rama SSID eliminuje potrzebę offline'owego zbierania danych, zapobiega ludzkim błędom w etykietowaniu danych i unika kosztów pracy związanych z treningiem modelu oraz zbieraniem danych przez eksperymenty. Ponadto, rama SSID umożliwia IDS łatwe dostosowywanie się do zmiennych w czasie charakterystyk ruchu sieciowego, znacząco poprawiając jego wydajność w wykrywaniu intruzji.

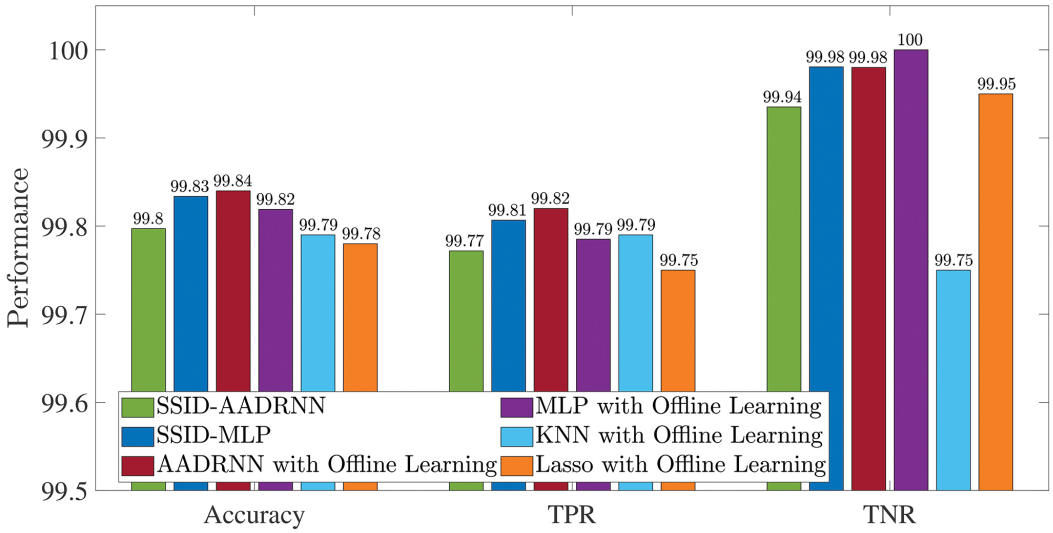

Figure 8: Porównanie wydajności modeli ML w ramie SSID z modelami z uczeniem offline

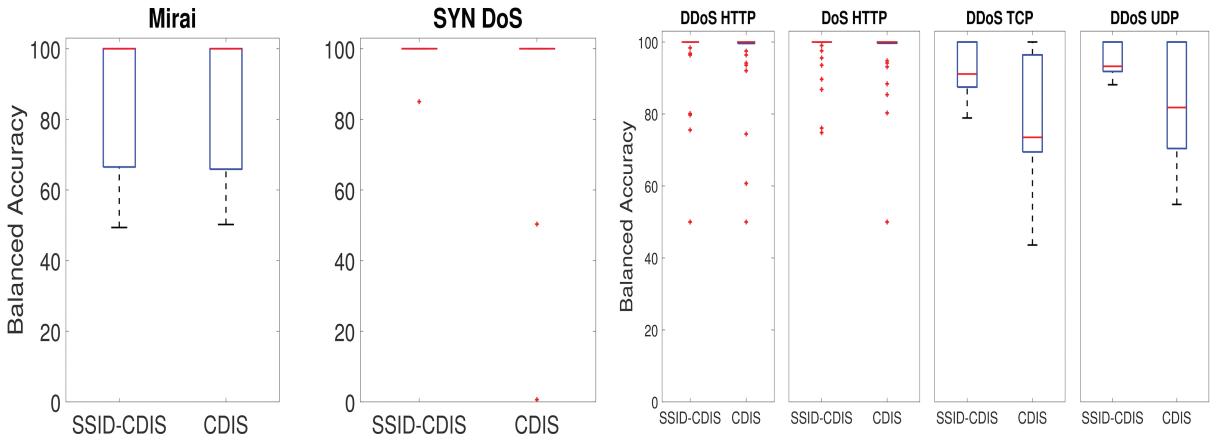

Figure 9: Porównanie wydajności CDIS szkolonego w ramie SSID z tym szkolonym w sekwencyjnym uczeniu kwazi-online na zestawach danych (lewo) Kitsune i (prawo) Bot-IoT

Bibliography

[1] "Hp study reveals 70 percent of Internet of Things devices vulnerable to attack," 2014, accessed: 2023-03-22. [Online]. Available: https://www.hp.com/us-en/hp-news/press-release.html?id=1744676

[2] Intersog, "IoT Security Statistics: 6 Facts [Updated]," Dec 2021, accessed: 2023-03-03. [Online]. Available: https://intersog.com/blog/iot-security-statistics/

[3] Cisco, *Cisco Annual Internet Report (2018–2023)*, Mar. 2020. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

[4] G. Matta, S. Chlup, A. M. Shaaban, C. Schmittner, A. Pinzenöhler, E. Szalai, and M. Tauber, "Risk management and standard compliance for cyber-physical systems of systems," *Infocommunications Journal*, vol. 13, no. 2, pp. 32–39, June 2021.

[5] S. Maksuti, M. Zsilak, M. Tauber, and J. Delsing, "Security and autonomic management in system of systems," *Infocommunications Journal*, vol. 13, no. 3, pp. 66–75, September 2021.

[6] E. Gelenbe and Y. Yin, "Deep learning with dense random neural networks," in *International Conference on Man–Machine Interactions*. Springer, 2017, pp. 3–18.

[7] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[8] N.-y. Liang, G.-b. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1411–1423, 2006.

[9] "Kitsune Network Attack Dataset," August 2020. [Online]. Available: https://www.kaggle.com/ymirsky/network-attack-dataset-kitsune

[10] "KDD Cup 1999 Data." [Online]. Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[11] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset," *Future Generation Computer Systems*, vol. 100, pp. 779–796, 2019.

[12] H. M. Song and H. K. Kim, "Self-supervised anomaly detection for in-vehicle network using noised pseudo normal data," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1098–1108, 2021.

[13] Z. Wang, Z. Li, J. Wang, and D. Li, "Network intrusion detection model based on improved byol self-supervised learning," *Security and Communication Networks*, vol. 2021, pp. 1–23, 2021.

[14] X. Zhang, J. Mu, X. Zhang, H. Liu, L. Zong, and Y. Li, "Deep anomaly detection with self-supervised learning and adversarial training," *Pattern Recognition*, vol. 121, p. 108234, 2022.

[15] H. Kye, M. Kim, and M. Kwon, "Hierarchical detection of network anomalies: A self-supervised learning approach," *IEEE Signal Processing Letters*, vol. 29, pp. 1908–1912, 2022.

[16] E. Caville, W. W. Lo, S. Layeghy, and M. Portmann, "Anomal-e: A self-supervised network intrusion detection system based on graph neural networks," *Knowledge-Based Systems*, vol. 258, p. 110030, 2022.

[17] M. Abououf, R. Mizouni, S. Singh, H. Otrok, and E. Damiani, "Self-supervised online and lightweight anomaly and event detection for IoT devices," *IEEE Internet of Things Journal*, vol. 9, no. 24, pp. 25 285–25 299, 2022.

[18] W. Wang, S. Jian, Y. Tan, Q. Wu, and C. Huang, "Robust unsupervised network intrusion detection with self-supervised masked context reconstruction," *Computers & Security*, vol. 128, p. 103131, 2023.