1. **Tytuł projektu:**

Wielowymiarowe modele uczenia maszynowego do analizy ruchu sieciowego ze szczególnym uwzględnieniem sieci przemysłu 4.0

2. **Słowa kluczowe**
   Wykrywanie anomalii w sieciach teleinformatycznych, klasyfikacja ruchu w sieciach teleinformatycznych, metody uczenia maszynowego, sieci splotowe
3. **Instytucja finansująca (nr umowy)**

Inicjatywa Doskonałości – Uczelnia Badawcza (1820/5/Z01/POB3/2021)

4. **Okres realizacji**

01.01.2021-31.12.2022

5. **Dofinansowanie (w tym w 2021)**
   279 400,00

6. **Partnerzy**
   -

7. **Kierownik projektu**
   Dr inż. Waldemar Graniszewski

8. **Zespół projektowy**

Dr hab. inż. Marcin Iwanowski, prof. PW

Mgr inż. Jacek Krupski

Inż. Damian Rybicki

Inż. Krystian Szefler

9. **Cel projektu (max. 1000 znaków)**

Głównym celem badawczym projektu jest przeanalizowanie i eksperymentalne przebadanie metod uczenia maszynowego, w szczególności sieci splotowych (konwolucyjnych) oraz porównanie ich z innymi technikami klasyfikacji ruchu sieciowego, w tym sieciach Przemysłu 4.0.

10. **Streszczenie (max. 1 strona)**

Tematem proponowanych prac badawczych jest automatyczna analiza rodzaju ruchu sieciowego na podstawie analizy strumienia ruchu w sieciach teleinformatycznych. W efekcie takiej analizy możliwe będzie selekcjonowanie poszczególnych rodzajów ruchu, czy też wykrywanie anomalii, stanowiących zagrożenie dla systemów informatycznych i nadzorowanych przez nie obiektów. W szczególności proponowane prace będą zogniskowane w dwóch obszarach: metodologicznym i aplikacyjnym.

Odnośnie pierwszego z nich - metodologicznego, badania będą dotyczyły zastosowania metod uczenia maszynowego, w tym uczenia głębokiego operujących na danych wielowymiarowych wygenerowanych na podstawie ruchu sieciowego oryginalnie reprezentowanego w formie szeregu czasowego. Dzięki takiemu przekształceniu możliwe będzie zastosowanie metod analitycznych znanych z wielowymiarowych modeli danych w uczeniu maszynowym, w tym także intensywnie rozwijanych obecnie splotowych (konwolucyjnych) sieci neuronowych. Analiza dostępnej literatury wykazała, że warto rozwijać prace w tym zakresie.

W odniesieniu do drugiego aspektu - aplikacyjnego, jako główne pole badawcze, będą wybrane systemy przemysłowe - ICS (ang. Industrial Control Systems). Cztery lata temu, w Unii Europejskiej została wprowadzona dyrektywa 2016/1148 Parlamentu Europejskiego i Rady Europy w sprawie środków na rzecz wysokiego wspólnego poziomu bezpieczeństwa sieci i systemów informatycznych na terytorium Unii. Cyberbezpieczeństwo ICS jest, zgodnie z tą dyrektywą, procesem holistycznym. Jednym z istotnych środków bezpieczeństwa jest segmentacja krytycznych systemów od reszty infrastruktury i odpowiednie monitorowanie ruchu sieciowego. Jednym z kluczowych elementów ochrony wielowarstwowej (ang. defence in depth) jest, m.in., zastosowanie systemów zapobiegania włamaniom – IPS (ang. Intrusion Prevention Systems). Te systemy oprócz standardowych reguł bezpieczeństwa stosują również metody uczenia maszynowego i sztucznej inteligencji. W ramach projektu weryfikowana będzie możliwość wykorzystanie metod modeli wielowymiarowych. Projekt będzie realizowany przez zespół złożony z przedstawicieli dwóch dyscyplin ITT oraz AEE. W ramach tych prac badawczych zostanie zrealizowana istotna część pracy doktorskiej.

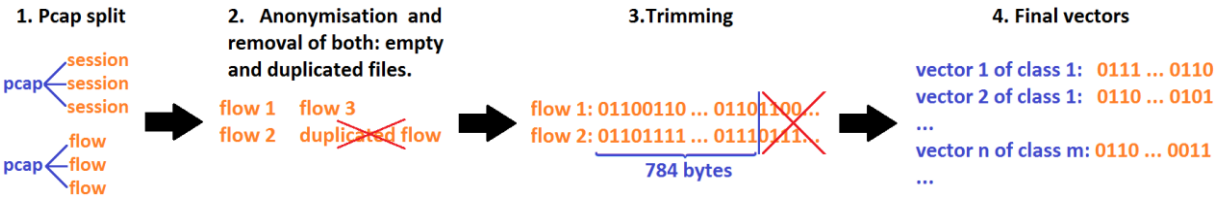
11. Dotychczasowe osiągniecia (max 2000 wyrazów)

Projekt jest w fazie intensywnych studiów literaturowych i analizowania istniejących metod uczenia głębokiego, w szczególności metod splotowych (konwolucyjnych) wykorzystywanych do klasyfikacji ruchu w sieciach teleinformatycznych.

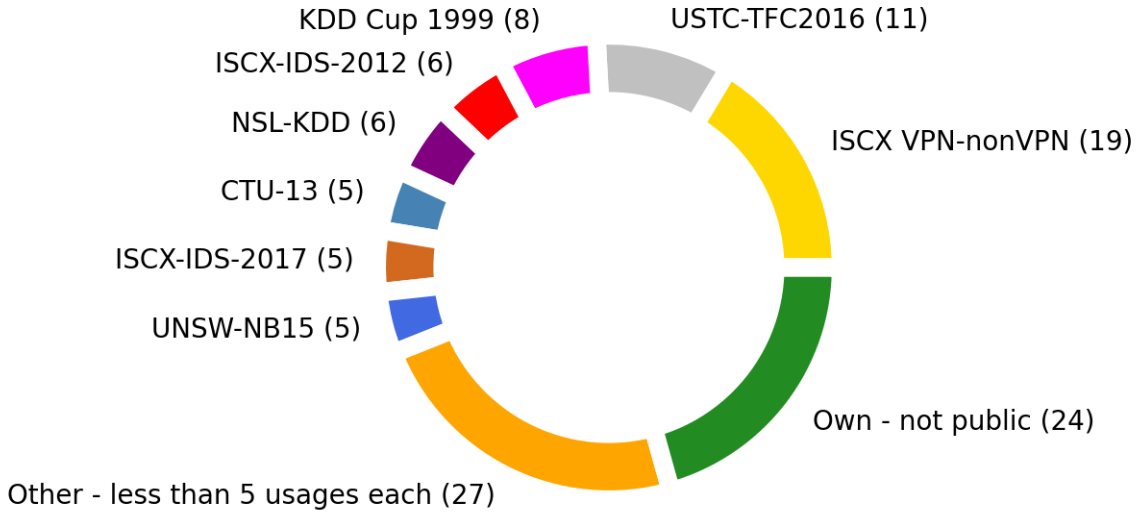
12. Publikacje

J. Krupski, D. Rybicki, W. Graniszewski, M. Iwanowski, Deep learning vs. traditional approaches to malware traffic classification - a comparative study, The 12th International Conference on Image Processing and Communications (IP&C'21), Bydgoszcz 2021 (accepted)

13. Materiały graficzne



Rys. 1 Sposób wstępnego przetwarzania danych wejściowych do systemu klasyfikacji ruchu sieciowego.



Rys. 2 Wykorzystanie zbiorów danych w poszczególnych publikacjach wykorzystujących sieci splotowe do klasyfikacji ruchu sieciowego.