Projekt zespołowy badawczy

Ćwiczenie numer: 2

Temat: **State of the art – analiza literatury**

Politechnika Białostocka

Wykonujący ćwiczenie:

- Jakub Kowalewski

Studia dzienne

Kierunek: Informatyka, II stopnia

Semestr: II

Grupa zajęciowa: PS 3

Prowadzący ćwiczenie: dr inż. Tomasz Grześ

Data wykonania ćwiczenia:
26.01.2025 r.

Min. 4 publikacje szczegółowo przeanalizowane ze zbioru min. 12 publikacji dotyczących tematu wybranych na podstawie streszczenia.

1. S. Chidambaranathan, M. Santhanaraj, E. Ajitha, S. F. A. S, S. B and T. Kathirvel, "Automated Detection of Encryption Algorithms Using AI Techniques," 2024 International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS), Bengaluru, India, 2024, pp. 1270-1274, doi: 10.1109/ICICNIS64247.2024.10823274.

   W dzisiejszej erze cyfrowej algorytmy kryptograficzne algorytmy kryptograficzne odgrywają istotną rolę w ochronie wrażliwych informacji, utrzymywaniu poufności i obronie przed cyberatakami. Wraz z rosnącą ilością metod szyfrowania, identyfikacja konkretnego algorytmu użytego do wygenerowania szyfrogramu staje się coraz większym wyzwaniem. Aby temu zaradzić, w ramach tego artykułu zaproponowano uczenie maszynowe, które automatycznie identyfikuje algorytmy kryptograficzne przy użyciu klasyfikatora Random Forest. System przetwarza ogromny zbiór danych o nazwie „encrypted-data", który zawiera 39 różnych algorytmów i trybów, wyodrębniając istotne cechy, takie jak częstotliwość bajtów i entropia z szyfrogramów. Te tworzą 257-wymiarowy wektor cech używany do klasyfikacji. Modele Random Forest są trenowane z tych danych. Podczas testowania modele są ładowane i używane do przewidywania algorytmu szyfrowania dla danego szyfrogramu. Metoda zaproponowana w artykule oferuje skalowalne i wydajne podejście do identyfikacji kryptograficznej, przewyższając tradycyjne techniki.

2. O, Sharif, & Mansoor, S.P.. (2011). Performance Evaluation of Classifiers used for Identification of Encryption Algorithms. 10.13140/2.1.4498.1768.

   W artykule poruszono kwestię znaczenia wydajności algorytmu klasyfikującego w rozpoznawaniu wzorców i uczenia maszynowego. W artykule opisano użycie technik rozpoznawania wzorców do identyfikacji algorytmów szyfrowania. Pod uwagę wzięto cztery różne algorytmy szyfrów blokowych: DES, IDEA, AES i RC2. Do klasyfikacji wykorzystano algorytmy takie jak: Naive Bayes, SVM, Neural Network, uczenie oparte na instancjach, Bagging, AdaBoostM1, Rotation Forest i Decision Tree. Wyniki badań przeprowadzonych przez autorów wskazują, że rozpoznawanie wzorców jest przydatną techniką przy identyfikacji algorytmów szyfrowania. Dodatkowo użycie jednego klucza szyfrowania zapewnia lepszą klasyfikację niż w przypadku użycia różnych kluczy. Dodatkowym wnioskiem jest stwierdzenie, że im więcej plików zostanie wykorzystanych przy trenowaniu modeli, tym dokładność klasyfikacji będzie wyższa.

3. Gancarczyk, G., Dąbrowska-Boruch, A., Wiatr, K. Efektywność parametrów statystycznych w detekcji informacji szyfrowanej. *Pomiary Automatyka Kontrola*. 2010, R. 56, nr 10, (10), s. 1137–1143.

   Autorzy artykułu zwracają uwagę na fakt, że informacja szyfrowana tak samo jak wszystkie inne typy danych mogą zostać poddane analizie statystycznej. Wyznaczanie parametrów takich jak wartość średnia, wariancja, czy też entropia nie jest zbyt skomplikowana. Autorzy artykułu używają w tym celu nowoczesnych narzędzi numerycznych takich jak MATLAB, Mathcad czy też Microsoft Excel. Artykuł ma na celu odpowiedź na pytanie, czy obliczenie tych parametrów dla plików zaszyfrowanych niesie za sobą wiedzę, którą można wykorzystać w pozytywny sposób. Autorzy podkreślają, że informacje te można wykorzystać w celu określenia, czy informacja jest zaszyfrowana, czy też nie. Artykuł zostanie wykorzystany jako wzór, na którym analogicznie będzie wykonywana entropia danych przygotowanych na potrzeby mojego eksperymentu jednak z tą różnicą, że zostanie przeze mnie technologia i język Python wraz z jego bibliotekami.

4. S. O. Sharif, L. I. Kuncheva and S. P. Mansoor, "Classifying encryption algorithms using pattern recognition techniques," 2010 IEEE International Conference on Information Theory and Information Security, Beijing, China, 2010, pp. 1168-1172, doi: 10.1109/ICITIS.2010.5689769.

   Artykuł skupia się na próbie identyfikacji słabych punktów w algorytmach używanych do szyfrowania kodu lub metodach używanych do generowania kluczy. W przeprowadzanym badaniu wykorzystano techniki rozpoznawania wzorców do identyfikacji algorytmów szyfrowania dla szyfrów blokowych. Badaniu poddano algorytmy DES, IDEA, AES i RC działające w trybie ECB. Zbadano również takie techniki klasyfikacji jak Naive Bayes, SVM, Neural Network, AdaBoostM1, Random Forest i Decision Tree. Celem badania było znalezienie najlepszego algorytmu klasyfikacji do identyfikacji metody szyfrowania. W przypadku tego badania najlepszym okazał się klasyfikator oparty o algorytm Random Forest.

5. De Gaspari, F., Hitaj, D., Pagnotta, G. *et al.* Reliable detection of compressed and encrypted data. *Neural Comput & Applic* **34**, 20379–20393 (2022).

6. Olber, P. (2023). Sztuczna inteligencja i przestępczość przyszłości w kontekście kryminalistycznych badań informatycznych. Przegląd Policyjny, 149(1), 138-155.

7. A. D. Dileep and C. C. Sekhar, "Identification of Block Ciphers using Support Vector Machines," The 2006 IEEE International Joint Conference on Neural Network Proceedings, Vancouver, BC, Canada, 2006, pp. 2696-2701, doi: 10.1109/IJCNN.2006.247172.

8. A. Cufoglu, M. Lohi and K. Madani, "Classification accuracy performance of Naïve Bayesian (NB), Bayesian Networks (BN), Lazy Learning of Bayesian Rules (LBR) and Instance-Based Learner (IB1) - comparative study," 2008 International Conference on Computer Engineering & Systems, Cairo, Egypt, 2008, pp. 210-215, doi: 10.1109/ICCES.2008.4772998.

9. Lin, X., Zhang, C., Dule, T. (2011). On Achieving Encrypted File Recovery. In: Lai, X., Gu, D., Jin, B., Wang, Y., Li, H. (eds) Forensics in Telecommunications, Information, and Multimedia. e-Forensics 2010. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 56. Springer, Berlin, Heidelberg.

10. Sorokin, I. Comparing files using structural entropy. *J Comput Virol* **7**, 259–265 (2011).

11. Saraf, Kundankumar Rameshwar, Vishal Prakash and Amit Mishra. “Text and Image Encryption Decryption Using Advanced Encryption Standard.” (2014).

12. Atikah, Nur & Rizky Ashila, Mutia & Setiadi, De Rosal Ignatius Moses & Rachmawanto, Eko & Sari, Atika. (2019). AES-RC4 Encryption Technique to Improve File Security. 1-5. 10.1109/ICIC47613.2019.8985825.