Klasifikácia a podobnosť rodín malvéru

Peter Chomič

Vedúci práce: JUDr. RNDr. Pavol Sokol, PhD.

Konzultant: Mgr. Ladislav Bačo

Motivácia

- 2017 – 12M vzoriek nového malwaru – potreba automatickej klasifikácie
- Machine learning a data mining
- Pomoc pri analýze malwaru
- Využitie aj pri automatickej tvorbe nových tried a detekcii malwaru

Zdroj: Jung, Byungho, Taeguen Kim, and Eul Gyu Im. "Malware classification using byte sequence

Ciele

1) Vytvorenie dátovej sady pre klasifikáciu vzoriek malvéru
2) Porovnanie prístupov ku klasifikácii a určovaniu podobnosti malvéru
3) Vytvorenie modelu, implementácia a vyhodnotenie pre klasifikáciu vzoriek malvéru do jednotlivých rodín malvéru

Riešenia

- Množstvo druhov dát, dynamické aj statické (inštrukcie, volania knižníc...)
- Viacero interpretácií (obrázky, n-gramy, sekvencie, grafy, DNA)
- Klastrovanie: K-means, hierarchické, euclid distance, mean shift, affinity
- Klasifikácia: neurónové siete, SVM, K-NN, rozhodovacie stromy, regresia
- Ensemble learning

Zdroj: Jung, Byungho, Taeguen Kim, and Eul Gyu Im. "Malware classification using byte sequence

Literatúra

1) Saxe, J., Sanders, H.: Malware data science - attack detection and attribution, San Francisco, No starch press. 2018.

2) Monnappa, K.A.: Learning malware analysis, Packt. 2018.

3) Gibert, Daniel, Carles Mateu, and Jordi Planes. "An End-to-End Deep Learning Architecture for Classification of Malware's Binary Content." *International Conference on Artificial Neural Networks*. Springer, Cham, 2018.

4) Yan, Jinpei, Yong Qi, and Qifan Rao. "Detecting malware with an ensemble method based on deep neural network." *Security and Communication Networks* 2018 (2018).

5) Yakura, Hiromu, et al. "Malware Analysis of Imaged Binary Samples by Convolutional Neural Network with Attention Mechanism." *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy*. ACM, 2018.

6) Jung, Byungho, Taeguen Kim, and Eul Gyu Im. "Malware classification using byte sequence information." *Proceedings of the 2018 Conference on Research in Adaptive and Convergent Systems*. ACM, 2018.

Ďakujem za pozornosť

Peter Chomič
peter.chomic@student.upjs.sk