ARIA
APPLIED RESEARCH IN ACTION

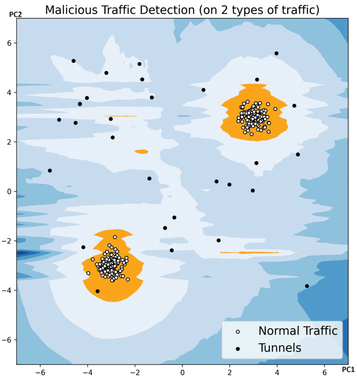
Detecting Malicious DNS Traffic Using Neural Networks

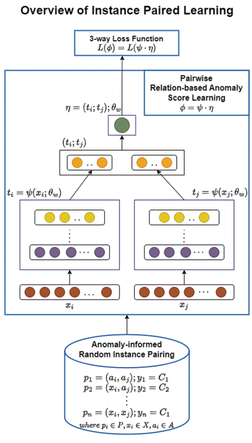
Enhancing Network Security Using Weakly-Supervised Learning and Pairwise Relation Prediction Network

Rajesh Marudhachalam and Ruixuan Zhang

Murat Erdogdu
ACADEMIC SUPERVISOR

Jeff Ip
INDUSTRY SUPERVISOR





PROJECT SUMMARY

Detecting DNS (Domain Name System) tunnels, covert channels exploited for malicious activities, pose a substantial threat to internet stability and integrity. We studied this problem and propose an efficient approach to detect tunnels by distinguishing benign and malicious DNS traffic that involves comprehensive data exploration, including Perturbation and Exploratory Data Analysis, to inform feature selection for model development [2,3]. Our approach employs a weakly-supervised learning framework that uses 3-way pairwise relation modelling [1] to detect tunnels as anomalies, that leverage's learning pairwise relation features and anomaly scores through predicting the relation of any two randomly sampled training instances. This pairwise approach seamlessly augments the training dataset. The results demonstrate a robust system, excelling at preserving benign traffic integrity while effectively identifying tunnels. Our research introduces a novel neural-network based architecture to cybersecurity deployed in a real-time environment, fortified by monitoring and a self-auditing function. Furthermore, it not only bolsters the defences against tunnels but also sets the stage for the continued development of advanced detection techniques and the adaptation of cybersecurity strategies to combat the ever-evolving landscape of malicious DNS tunnelling.

REFERENCES

Pang, G., Shen, C., Jin, H., & Hengel, A. van den. (2019). Deep Weakly-supervised Anomaly Detection. https://doi.org/10.48550/arxiv.1910.13601

Wang, Y., Zhou, A., Liao, S., Zheng, R., Hu, R., & Zhang, L. (2021). A comprehensive survey on DNS tunnel detection. Computer Networks (Amsterdam, Netherlands : 1999), 197, 108322–. https://doi.org/10.1016/j.comnet.2021.108322

Bekker, J., & Davis, J. (2020). Learning from positive and unlabeled data: a survey. Machine Learning, 109(4), 719–760. https://doi.org/10.1007/s10994-020-05877-5

BLUECAT™

Computer Science
UNIVERSITY OF TORONTO
Master of Science in Applied Computing