Explainable Malware Detection Using Predefined Network Flow

Boryau Hsupeng*, Kun-Wei Lee*, Te-En Wei*, Shih-Hao Wang*

**Cybersecurity Technology Institute, Institute for Information Industry, Taiwan, R.O.C*

bojsahsupeng@iii.org.tw, kwlee@iii.org.tw, dwyanewei @iii.org.tw, shwang@iii.org.tw

***Abstract*—As the internet has become an indispensable part of modern life, defences against cybersecurity attacks have become an important topic and a considerable number of studies have been made to provide reliable tactics to defend against cyberattacks. Flow export protocols and technologies provide several advantages in network monitoring. By using flow data aggregated from packets, the amount of data to be analysed has been significantly reduced and it is often said to be more scalable than packet-based traffic analysis. With the help of modern Artificial Intelligent algorithms, AI can be trained with flow data to predict hackers' attacks and types of malware. In this paper, we will present CSTITool, a CICFlowMeter-based flow extraction tool, to feature extraction with an aim of improving the model performance. The flow features will be used to train a machine learning-based model for hackers' attacks and malware classification. To provide interpretability, an explainable AI will be introduced to help understand the relation between the prediction and the features.**

***Keyword*—Explainable AI, Cybersecurity, XGBoost, Malware Classification, SHAP**

Boryau Hsupeng received his M.S. degree in the Department of Space Science and Engineering from National Central University, Taoyuan, Taiwan, in 2015. He is currently an engineer in Cybersecurity Technology Institute, Institute for Information Industry, Taiwan. His current and previous research interests include magnetospheric physics, space plasma physics, numerical simulation, and machine learning.

Kun-Wei Lee received his M.S. degree in statistics from National Chengchi University, Taipei, Taiwan, in 2017. He is currently an engineer in Cybersecurity Technology Institute, Institute for Information Industry, Taiwan. His research interests include data mining, machine learning, deep learning, and generative adversarial network.

Te-En Wei received his Ph.D. degree in the Department of Computer Science and Information Engineering at National Taiwan University of Science and Technology, Taipei, Taiwan. He is currently an acting director in Cybersecurity Technology Institute, Institute for Information Industry, Taiwan. His research includes Network Security, Application Security, and Financial Forecasting.

Shih-Hao Wang received his post-graduate degree of Management Information Systems from National Pingtung University of Science and Technology. He is a section manager at Institute for Information Industry. He has conducted research and design in cyber security, network communication, and fleet management.