

INFORMATION ON DOCTORAL DISSERTATION

Title of Thesis:

Network Anomaly Detection Models based on Deep learning and Data Fusion

Major: **Information Systems**

Code: **9.48.01.04**

Name of PhD candidate: **Bui Cong Thanh**

Committees:

- 1. Associate Professor.Doctor. Hoang Minh**
- 2. Associate Professor.Doctor. Nguyen Quang Uy**

Academic Institution: Posts and Telecommunications Institute of Technology

NEW RESULTS OF THE DISSERTATION:

The rapid development of the computer network in all aspects of its infrastructure has put it under the pressure of modern cyber-attacks. Seeking solutions to identify and prevent cyber attacks is a crucial task that has attracted the research community, mainly investigating to enhance network anomaly detection (NAD) method. The thesis aimed to improve network anomaly detection algorithms; the result is presented in the theoretical algorithms, which are later supported by empirical results. The contributions of the thesis can be summarized as follows.

(1) The thesis proposed novel models of Network Anomaly Detection based on deep learning; they are named Clustering-Shrink AutoEncoder (KSAE) and Double-Shrink AutoEncoder (DSAE). These solutions are to solve the challenges of the SAE model, which is a typical deep learning-based NAD model, achieves convincing results on a wide range of network security datasets. Amongst them, DSAE is constructed in a very different way as other methods based on AutoEncoder. DSAE uses both reconstruction errors (RE) and the latent vector to build the anomaly score. The experiment shows that DSAE provides a highly reliable in detecting an intrusion that mimics the normal data, which usually leads SAE challenging to address.

(2) The thesis proposes a novel fusion-based framework for network anomaly detection called OFuseAD (One-Class Fusion-based Anomaly Detection). Since the stand-alone NAD method faces an issue, it can efficiently perform a problem while it can not perform well on another problem. OFuseAD provides a general architecture to build a fusion-based model in semi-supervised learners (unavailable of positive data and its corresponding label to evaluate the algorithm). This model takes advantage of the individual stand-alone NAD, and it can also eliminate the decision threshold to meet the

requirement of deploying in the real-world network.

(3) The thesis introduces the method to apply Dempster - Shafer (D-S) theory for network anomaly detection. It provides a novel function called One-Class Basic probability assignment (OBPA), which are often too complicated to apply D-S theory. It also presents a solution to solve the limitation of DS-theory's rule, which is named DRC_AD, in unequal performance amongst individual local detectors.

APPLICATION AND USED IN THE REAL WORLD OR FUTURE WORKS:

(1) The OFuseAD can be utilized on another than the network security domain of anomaly detection problems. In general, the model detection based on OFuseAD can be considered in circumstances of the OCC problem.

(2) Dempster - Shafer (D-S) theory with OBPA function and DRC_AD rule can be considered for developing the application.

(3) The thesis introduces a novel metric for measuring the generalization ability of OCCs. This metric can be considered to estimate the weight of OCC methods.

(4) Besides that, this thesis still has some limitations: the OFuseAD model for network anomaly detection assumes that the original information source for the local detectors to observe is the same, leading the model less flexible; the computation complexity of OFuseAD is based on the OCC detector, with the distance-based and density-based method, which usually get the high cost.

(5) In the future, several developments from OFuseAD are considerable such as using more detectors, with the heterogeneity of sensors including a decision from the security expertise, other platform security application, or physical detectors such as voltage metrics of security devices.

Research supervisors

PhD candidate

Associate Professor. Doctor. Hoàng Minh

Bui Cong Thanh