ABSTRACT

The increasing complexity of cyberattacks in the Internet of Things (IoT) architecture requires an Intrusion Detection System (IDS) that is not only accurate in detecting threats, but also capable of maintaining user data privacy. Conventional IDS approaches based on a centralized architecture are considered ineffective because they are vulnerable to single points of failure and pose a risk of data leaks. To address this, Federated Learning (FL) emerges as an innovative solution that enables collaborative model training among devices without need to send raw data to a central server. However, the effectiveness of FL is significantly influenced by the choice of aggregation method, which impacts model convergence, detection accuracy, and system resilience, especially when handling data with non-independent and identically distributed (non-IID) characteristics. Several researchers have discussed aggregation methods in FL related to non-IID data, but without explicitly evaluating the effect of regularization strength on FedProx under non-IID data distribution conditions. This thesis proposes the development of a Collaborative Intrusion Detection System (CIDS) based on FL using two aggregation methods, FedAvg and FedProx to evaluate attack detection performance on the CICIoT2023 dataset. Experiments were conducted on IID and non-IID data scenarios using a Deep Neural Network (DNN) model and feature selection techniques based on the Feature Importance Gain values from XGBoost. The results show that on IID data, the combination of FedAvg and feature selection provides the highest accuracy of 97.76%. Meanwhile on non-IID data, the combination of FedProx (μ = 1.0) and feature selection achieves the best accuracy of 96.92%.

Keywords: Internet of Things, CIDS, Federated Learning, Deep Learning, CICIoT2023