ABSTRACT

Insider threat is a potential threat to an organization that results in financial and reputation losses while exposing sensitive information. Past research extensively focused on external threats, and overlooked on both intentional and unintentional insider threats. Several researchers majorly focused on detecting such insider activities but fail to mitigate both intentional and unintentional insider threats. Few challenges such as mishandling imbalanced dataset and fail to incorporate feature engineering techniques, limited mitigation strategies are encountered. This research employs a hybrid machine learning approach to identify insider threats and incorporated behavioural biometrics with user profiling to mitigate both intentional and unintentional insiders effectively.

A methodology comprising of three phases is proposed. It consist of Preprocessing and Insider Detection (P&ID) in Phase I, Unintentional Insider Mitigation (UIM) in Phase II, and Intentional Insider Mitigation (IIM) in Phase III. P&ID consist of two layers - Preprocessing, and Insider Detection. In Layer 1, log data is preprocessed using data integration, encoding and tuned the nearmiss-2 sampling technique to obtain a balanced data to diminish the class imbalance problem. In Layer 2, a hybrid B-SVM combining Support Vector Machines (SVM) and Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH) is applied. It classifies users into genuine, intentional insiders, and unintentional insiders. The proposed method achieved a 99.15% detection accuracy, with a low misclassification rate of 0.85% for detecting both intentional and unintentional insider threats.

Once unintentional insiders are detected, the unintentional insiders are mitigated in UIM phase. UIM phase consist of two layers – Feature engineering, and Core behavior identification. In Layer 1, Clonal Kernel Principal Component Analysis (CKPCA) is proposed for feature engineering. CKPCA integrates population subset selection, kernel mean embedding, and dimensionality reduction to improve feature representation. These features are further analyzed using Deep Belief Networks (DBN) in Layer 2 that achieved 99.84% authentication accuracy and a 0.15% Equal Error Rate (EER) of 0.15%. This phase significantly minimizes false alarms and ensures a reliable mitigation process for unintentional insiders.

In IIM phase, the detected intentional insiders are mitigated using user profiling mechanism based on their authentication outcome. IIM phase consist of three layers – Data pre-processing, Model training and evaluation, and User profiling. In Layer 1, data pre-processing is done using label encoding and train-test split. In layer 2, Decision tree is modeled to categorize users low-risk and high-risk. In Layer 3, Low-risk users with legitimate activities are profiled into the Allowlist, while users displaying malicious intent with high-risk are placed on the Denylist. This adaptive profiling ensures that intentional threats are neutralized without affecting genuine users.

The methodology was validated using two datasets namely the CERT Insider Threat Dataset and the CIC Darknet Dataset. P&ID detected 8 intentional and one unintentional insider among 250,078 daily logs using CERT Dataset. P&ID is validated with darknet dataset, detected 4,783 intentional-Darknet users and 68 unintentional-Darknet users where (VPN: 42) (Tor: 21) (NonVPN: 5) among 134,305 daily activities. UIM mitigated one unintentional insider as an intentional insider using CERT log activities. UIM mitigated 68 unintentional-Darknet users as 64 Intentional-Darknet and 4 benign users using darknet dataset. IIM profiled 57 genuine users in Allowlist and 8 intentional insiders in Denylist using CERT dataset. Using CIC Darknet dataset, the IIM profiled 5063 benign users in Allowlist and 4847 Intentional-Darknet users in Denylist.

This study offers a practical and highly effective solution for insider threats in environments where user log data is analyzed. By combining hybrid machine learning models with behavioral biometrics and user profiling, the approach ensures accurate detection and mitigation of both intentional and unintentional threats. This approach can be applied in any environment where user log is prevalent.