Chapter 12
Scenario-Aware Anomaly Detection in Insider Activities Using CT-GAN Data Augmentation

K. Kamatchi
https://orcid.org/0000-0003-1243-5071
Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, University in Morai, India

E. Uma
Anna University, India

ABSTRACT

Insider threats are cyberattacks perpetrated by individuals with legitimate organizational access. Detecting insider attacks is challenging due to their subtle nature and potential for causing significant harm. One popular technique for spotting dangerous insiders is anomaly detection. However, internal analysis of danger remains a largely untapped study area due to the lack of actual events data and the unequal distribution of classes in datasets. In order to enhance underutilised minority class instances and enrich the data with diverse and significant examples drawn from pragmatically malicious events, we develop a Conditional Tabular Generative Adversarial Network (CT-GAN) in this research. Comprehensive experiments utilizing the CMUCERT Version 4.2 dataset showcase the impact of CT-GAN-augmented data in enhancing the detection of anomalies within insider activity analysis. Additionally, the approach is assessed against several existing methods using various criteria and performance metrics.

DOI: 10.4018/979-8-3373-2115-8.ch012

INTRODUCTION

Organizations are continually exposed to various cyber threats, but insider attacks carried out by individuals with direct or indirect access to internal systems are especially troubling. These attacks are particularly dangerous because the perpetrators are already integrated into the organization, making them familiar with its systems and processes, whether physically or virtually. Recent insights indicate that insider threats will become more critical and top-of-mind for security professionals in 2023. According to Cybersecurity Insiders' 2023 Insider Threat Report, 74% of organizations consider themselves moderately to severely vulnerable to insider attacks, with 60% of respondents having encountered one in 2022, (Cybersecurity Insiders, 2023). This represents a marked increase in frequency, underscoring the growing difficulty in detecting and preventing such threats, which often involve legitimate accounts and credentials, rendering them harder to distinguish from normal user activity. Insider threats unfold gradually, as attackers carefully plan and execute their actions over time. They often combine various strategies, relying on past experiences and established attack patterns. Studies on insider threats show that these incidents involve multiple fraudulent actions targeting trusted resources within an organization Lippi et al., (2025), Almelehy et al., (2025), Alboalebrah and Al-augby, (2025), Almedires et al., (2025), Ang et al., (2025), Alshuaibi et al., (2025), Alotaibi et al., (2025), Aljumaiah et al., (2025).

Researchers have categorized these activities into distinct scenarios to better understand and combat them. Key examples include Data Exfiltration, System Sabotage, and Intellectual Property Theft, which provide insight into the diverse tactics used by insiders, (Greitzer, 2019; Jones, 2024). Although the effect and rising incidence of insider threats are well acknowledged, the dearth of actual-world information due to privacy issues and inequitable data distributions have rendered insider threat research a less studied subject in cybersecurity. Identifying insider threats is usually framed as an anomaly detection problem, with malicious actions being infrequent and hard to detect. In many real-world applications (Le & Zincir-Heywood, 2021; Yuan & Wu, 2021), Anomaly Detection (AD) is crucial, particularly in cybersecurity. The majority of the time, anomaly detection techniques which usually use binary or one-class classification models are used to detect insider threats. Supervised learning models often overfit when trained on limited class data, focusing on dominant classes, and reducing training effectiveness. This leads to poor performance when applied to new data. While several techniques can reduce overfitting, gathering more data remains the most effective approach. Ensuring that every class is represented with meaningful samples significantly boosts the accuracy and efficiency of anomaly detection systems. In order to improve multi-class AD and address the problem of scarce information for minority classes in insider threat situations, this study uses a