Assignment 4

Name:
Amjad Alqahtani
wkh221

Problem Statement

The paper mentioned that there is an increment of sophistication in cyberattacks and we need to have more effective and optimized Network Intrusion Detection Systems (NIDS) to encounter the attacks. The authors propose a novel framework for encountering anomalies. The framework is based on leveraging both image processing and deep learning for enhanced NIDS performance.

A traditional way of machine learning-based NIDS struggles to maintain the trade-off between efficiency and effectiveness. There are attempts at optimization through feature selection, data augmentation, and hybrid algorithms. However, the evolution of cyberattacks are high and challenge these approaches.

This paper is based on a multi-tier solution. First, it reduces the feature set for efficiency. Second, it transforms non-image data that comes from logs or requests into image format. The transformed images are enhanced using a Gabor filter. Third, the images are processed with a CNN to improve anomaly detection.

This framework uses three NIDS datasets called : CSE-CIC-IDS 2018, CIC-IDS 2017, and ISCX-IDS 2012. From the experiment evaluations, the framework has shown high detection accuracy across these datasets and outperforms previous methods approximately 92% detection rate.

To sum up, this paper presents a promising approach to handling complex and varied network traffic patterns using a combination of feature reduction and CNN-based image classification.

Problem with Existing Solution

Let me start by saying that finding an effective ML-based solution for NIDS is a challenging task and has some limitations.

First, as network workloads change, ML models must be regularly updated to remain accurate. However, this process is both time consuming and resource intensive.

Second, existing solutions often employ hybrid approaches for data preprocessing, feature selection, and predictive algorithms. Recently, DL methods, particularly CNNs, have gained popularity in NIDS because of their ability to reconstruct features and learn intricate patterns from images.

Third, to apply CNNs to NIDS, non-image network data must be converted into an image format, commonly done through Fourier transformations or multidimensional matrix representation. The multidimensional matrix approach is efficient, but compromises feature correlations, thus impacting detection accuracy. Fourier-based transformations can be computationally complex and struggle with scalability on large datasets.

Fourth, hybrid approaches, which combine various ML techniques, often lack precision when detecting minor attack types or rare threats. These approaches may focus on more common attack patterns, leading to biased models that overlook critical but infrequent anomalies.

Fifth, CNNs depend heavily on the transformation of non-image network data into image formats suitable for CNN-based models. Hybrid approaches that combine feature selection and DL techniques have been developed to improve accuracy, but these models often struggle to detect minor attack types effectively.

Finally, high accuracy in DL-based NIDS models often requires high-resolution image transformations and extensive computational resources. This creates a trade-off, as reducing

resource usage by lowering image resolution or simplifying transformations can decrease detection precision, potentially missing subtle patterns indicative of an attack.

Proposed Solution

The proposed solution combines feature reduction and image transformation for improved detection using CNNs. This framework consists of a few primary steps:

The framework begins with data preprocessing to clean and prepare the network traffic data. Any missing values, duplicates, or other inconsistent data items are addressed to ensure data quality and consistency. This step is critical for efficient processing and accurate feature selection.

To optimize computational efficiency, an augmented feature selection process is applied, filtering out irrelevant features. This reduction not only decreases data complexity but also enhances detection performance by focusing on relevant patterns. To do this, the framework has to imply some normalization techniques. This step ensures that the data retains patterns crucial for anomaly detection.

After selecting essential features, the framework uses a DeepInsight-based approach to convert the non-image data into a 2D image format. This method maps features in a way that maintains feature relationships, which is important for accuracy. The transformed images use a Gabor filter for enhancement. Gabor filter is a technique that emphasizes spatial patterns and textures. This enhancement makes patterns within the data clearer for the CNN, which helps in accurately identifying potential anomalies.

The final step of the framework is classification using a CNN model. The CNN is designed with multiple convolutional and dense layers that extract and analyze intricate patterns within the image data to increase detection accuracy by capturing anomalies that may go

undetected by simpler models. The CNN model processes the enhanced images, learning to distinguish between normal and abnormal patterns in network traffic.

To sum up, by reducing the feature set, it decreases computational overhead, making it more efficient than methods that process all available features. Unlike Fourier or spectrogram-based approaches, which may compromise feature correlations, the DeepInsight-based method used in this framework preserves these relationships.

The framework improves the accuracy of CNN-based detection. Moreover, applying a Gabor filter to the transformed images led to enhancing CNN's ability to detect anomalies.