Data Days 2022

Enhancing Anomaly Detection Of DDoS Attacks In Streaming Cloud Server Data Using Active Learning And Human Feedback

Abtin Mahyar
Saeid Cheshmi
Ghazal Rafiei

July 20, 2023

Abstract

We propose an anomaly detection system for detecting DDoS attacks in a cloud service using active learning and human-in-the-loop feedback. The system combines LSTM networks with an asymmetric auto-encoder to capture temporal dependencies in time-series data. Unsupervised training is employed, and instances with high uncertainty are selected through active learning for domain expert labeling. The feedback from experts refines the model's capabilities, adapting to evolving network conditions and data distributions. The proposed architecture offers a flexible, robust, and adaptable solution for real-time anomaly detection in cloud server network traffic, effectively mitigating the impact of concept drift and data drift. By integrating active learning and human expertise, the system continuously improves its performance, ensuring accurate detection of DDoS attacks and other anomalies in cloud services.

Table of Contents

List of Figures

Introduction

In recent years, cloud-based services have experienced exponential growth, becoming integral components of various industries and organizations. However, this rapid expansion has also attracted malicious actors seeking to disrupt these services through Distributed Denial of Service (DDoS) attacks. DDoS attacks pose a severe threat to cloud servers, overwhelming them with a massive volume of traffic, resulting in service unavailability and potential financial losses. Detecting and mitigating these attacks in real-time is crucial to ensuring the reliability and continuity of cloud services. Traditional anomaly detection systems often struggle to keep pace with the dynamic and evolving nature of DDoS attacks, leading to false positives or delayed responses. Moreover, cloud server networks' continuous evolution and traffic patterns dynamic nature necessitate a flexible and adaptive detection mechanism.

In this context, we present a novel anomaly detection system for identifying DDoS attacks in cloud services. Our approach leverages the integration of active learning and human-in-the-loop feedback, aiming to improve the efficiency and accuracy of anomaly detection in time-series data, particularly in the context of cloud server network traffic. By incorporating domain experts into the detection process, our method enables the model to learn from human analysts' collective intelligence and expertise, adapting to emerging threats and novel attack patterns in real-world scenarios.

The integration of active learning and human-in-the-loop feedback in our anomaly detection system empowers the model to learn from human experts and iteratively improve its performance. By intelligently selecting instances for labeling through active learning, we reduce the labeling burden on experts, enabling them to focus on the most challenging and critical data points. The feedback from domain experts provides the model with accurate and reliable labels, enhancing its understanding of intricate temporal patterns and anomalies. As a result, our method effectively combines the strengths of machine learning algorithms and human expertise, creating a synergy that outperforms traditional unsupervised approaches. The continuous adaptation to evolving network conditions and emerging threats ensures that our system remains at the forefront of anomaly detection in cloud services, safeguarding against DDoS attacks and promoting the seamless operation of cloud-based infrastructures.

Background

Streaming anomaly detection. It is the process of continuously monitoring and analyzing real-time data streams to identify unusual or abnormal patterns or events. In this method, data is processed as it arrives, rather than in batches, allowing for rapid detection of anomalies as they occur. Algorithms, such as statistical methods or machine learning models, are applied to the streaming data to detect deviations from normal behavior. The goal is to promptly identify and flag anomalies, enabling timely responses and actions to prevent potential issues or threats. Streaming anomaly detection is crucial for various applications, including monitoring critical systems, detecting fraudulent activities, and maintaining the optimal performance of dynamic environments where data is constantly evolving.

Human-in-the-loop (HITL). It is a concept that involves integrating human intelligence and decision-making into automated or artificial intelligence systems. In HITL, humans are actively engaged in various stages of a process or system, collaborating with and providing feedback to the machine learning or automation algorithms. This collaboration enables humans to validate, refine, or correct the outcomes of automated processes, particularly in scenarios where algorithms may encounter uncertainty, ambiguity, or complex situations that require human judgment. By incorporating human expertise, HITL aims to enhance the overall performance, reliability, and ethical considerations of AI systems, bridging the gap between machine capabilities and human understanding to achieve more accurate and contextually appropriate results.[1]

Data Drift. It refers to the situation where the statistical properties of the incoming data change over time. These changes can be gradual or abrupt and can result from various factors such as changes in user behavior, hardware upgrades, or software updates. Data drift affects the features and characteristics of the data while keeping the underlying concept (or the target variable) unchanged. In the context of anomaly detection, data drift can lead to shifts in the distribution of normal data, causing the model to produce more false positives or false negatives.

Concept Drift. It, on the other hand, occurs when the relationship between the input features and the target variable (i.e., the concept being learned) changes over time. This means that the underlying concept that the model was originally trained to learn is no longer valid, leading to a degradation in the model's performance. In the context of anomaly detection, concept drift can arise from changes in the network behavior or the emergence of new types of anomalies that the model has not encountered before.

Active learning. It is a machine learning approach that involves iteratively selecting and labeling the most informative instances from an unlabeled dataset to improve the performance of a model. Instead of relying solely on a large set of pre-labeled data, active learning actively queries an oracle, which can be a human expert or a pre-trained model, to obtain labels for selected data points. The key idea is to strategically choose samples that are challenging or uncertain for the current model, aiming to reduce the labeling effort while achieving high accuracy. By iteratively updating the model with new labeled data, active learning effectively enhances the model's ability to generalize and make accurate predictions, making it particularly valuable in scenarios where acquiring labeled data is costly or time-consuming [2].

DDoS attacks. A Distributed Denial of Service (DDoS) attack on cloud servers is a malicious attempt to overwhelm and disrupt the normal operation of cloud-based services by flooding the servers with an overwhelming amount of traffic or requests. In a DDoS attack, a large number of compromised devices, known as botnets, are coordinated by the attacker to send an excessive volume of data, requests, or connection attempts to the cloud servers simultaneously. This massive influx of traffic exhausts the server's resources, such as CPU, memory, and network bandwidth, rendering the service unavailable to legitimate users. DDoS attacks can have severe consequences, leading to downtime, loss of revenue, and damage to the reputation of the targeted cloud service provider. Implementing robust DDoS mitigation measures, such as traffic filtering, rate limiting, and load balancing, is essential for safeguarding cloud servers against such malicious attacks.

Methodology

3.1 Proposed Model

The proposed anomaly detection system leverages a novel architecture that combines Long Short-Term Memory (LSTM) networks with an asymmetric auto-encoder for efficient anomaly detection in time-series data, with a specific focus on cloud server network traffic. The asymmetric auto-encoder, an essential component of this architecture, is specially designed to identify anomalies based on reconstruction errors. Unlike traditional auto-encoders, the asymmetric auto-encoder incorporates encoder and decoder structures intentionally imbalanced in capacity. The encoder is designed to possess higher capability, enabling it to effectively capture the underlying structure of normal instances while facing challenges in reconstructing anomalous patterns.

During the detection process, the LSTM encoder generates a compact latent representation of the input time series. This latent representation is subsequently used by the decoder of the asymmetric auto-encoder to reconstruct the original time series. The discrepancy between the reconstructed time series and the original input serves as the anomaly score for each instance. Instances exhibiting more intricate temporal patterns or deviating significantly from the normal data distribution yield higher reconstruction errors, thus being identified as potential anomalies.

To achieve unsupervised training, the proposed system utilizes the entire dataset, comprising both normal and anomalous instances. The LSTM-based architecture learns to accurately reconstruct the time series data, with a particular focus on minimizing the reconstruction errors for normal instances. By introducing a threshold on the anomaly scores, the system classifies instances with reconstruction errors surpassing the threshold as anomalies. This unsupervised approach enables the model to generalize effectively to new and unseen data while remaining robust to varying degrees of anomaly complexity inherent in cloud server network traffic.

In summary, the proposed anomaly detection system effectively captures temporal dependencies and complex patterns in time-series data through the combination of LSTM networks and an asymmetric auto-encoder. By employing reconstruction errors as anomaly scores and implementing an unsupervised approach with adaptive thresholding, the system exhibits enhanced performance and adaptability in real-world scenarios.

An important consideration in deploying anomaly detection systems in resource-constrained environments, such as cloud servers, lies in the need for efficient models with faster inference times and reduced memory requirements. While LSTM networks and asymmetric auto-encoders are effective for anomaly detection, they may lead to models with large parameter sizes and longer inference times. However, knowledge distillation presents a promising solution to address these challenges.

Knowledge distillation [3] empowers the proposed anomaly detection system to develop smaller, faster, and more generalized models suitable for real-world applications. By leveraging knowledge from a larger teacher model, the student model can accurately detect anomalies in cloud server network traffic while maintaining low computational costs. With advancements in knowledge distillation techniques, this approach offers a cost-effective means to keep the system up-to-date and robust in detecting anomalies, ensuring the security and reliability of cloud server infrastructures with minimal resource overhead. Through the integration of knowledge distillation, the proposed system not only enhances its performance but also extends its applicability to resource-constrained environments, making it a viable solution for real-time and efficient anomaly detection in cloud server networks.

3.2 Data

3.2.1 Features

Network Traffic Data. The most fundamental data for DDoS attack detection is network traffic data. This includes information such as the number of incoming requests, the source IP addresses, packet sizes, request types, and traffic volume over time. Analyzing this data helps the anomaly detection system to establish a baseline of what normal traffic looks like for the cloud service. Unusual spikes or patterns in network traffic can indicate a potential DDoS attack.

Request Rate. Monitoring the rate at which requests are coming into the cloud service is essential. DDoS attacks often involve a massive increase in the number of requests, overwhelming the service. By tracking the request rate, the anomaly detection system can flag unusual surges that may signify an attack.

Geolocation Data. Understanding the geographical origin of incoming traffic can be valuable in detecting DDoS attacks. If a disproportionate amount of traffic is coming from a specific region or from a large number of distinct locations simultaneously, it could be indicative of an orchestrated attack.

Protocol Distribution. Analyzing the distribution of protocols in network traffic can help distinguish legitimate traffic from malicious requests. DDoS attacks may exploit specific vulnerabilities in particular protocols, leading to an abnormal protocol distribution.

Resource Utilization. Monitoring the cloud service's resource utilization, such as CPU and memory usage, can provide valuable insights into the presence of a DDoS attack. A sudden and significant increase in resource consumption beyond normal thresholds is a strong indication of an ongoing attack.

Session Data. Examining the characteristics of user sessions, including session duration and the number of requests per session, can aid in anomaly detection. DDoS attacks may result in short-lived sessions with a high number of requests, differing from typical user behavior.

3.2.2 Pre-Processing

Before using time series data for anomaly detection in DDoS attacks on a cloud service, it is crucial to apply preprocessing steps to ensure the data's quality and reliability. Firstly, irrelevant or noisy information must be removed to enhance the accuracy of the analysis.

This involves eliminating duplicate entries, as redundant data could skew the results. Additionally, handling missing values is vital, as incomplete data can lead to biased conclusions. Outliers, which are extreme data points that deviate significantly from the overall pattern, should also be addressed, as they can distort the analysis and lead to false-positive results. By cleaning and refining the time series data, the anomaly detection system can focus on the most relevant and informative aspects, improving its ability to detect potential DDoS attacks effectively.

Another important preprocessing step is to remove short-term fluctuations and emphasize longer-term trends. Short-term fluctuations may arise due to noise in the data or temporary spikes that do not necessarily indicate an attack. To avoid false alarms, smoothing techniques can be applied to the time series data. Methods like moving averages or exponential smoothing can reduce noise and reveal underlying patterns and trends more clearly. By emphasizing longer-term trends, the anomaly detection system can better identify sustained deviations from normal behavior, which are more indicative of potential DDoS attacks. This preprocessing step enhances the system's ability to differentiate between genuine anomalies and transient variations, resulting in a more reliable and robust DDoS detection mechanism for cloud services.

3.3 System

As the streaming time series data flows into the system, the unsupervised anomaly detection model continuously analyzes it to produce anomaly scores. These scores represent the model's confidence in its predictions for each instance in the data stream. The active learning component then selects instances with high uncertainty scores, indicating that the model is unsure about their classification. These instances are candidates for querying domain experts for labeling, as they likely contain challenging or novel patterns that demand human expertise for accurate identification.

The pool of instances selected through uncertainty sampling is presented to domain experts for labeling. Leveraging their domain knowledge and insights, domain experts can efficiently identify true anomalies amidst uncertain instances, providing valuable labeled data to enhance the model's learning process. By incorporating these newly labeled instances into the training set, the unsupervised model adapts and updates its internal representations, effectively refining its anomaly detection capabilities based on the collective intelligence of human experts.

Cloud server networks are dynamic and subject to frequent changes, making them susceptible to concept drift and data drift. The human-in-the-loop feedback loop plays a pivotal role in enabling the system to adapt to evolving network conditions, data distributions, and emerging threats. By involving experts in the feedback loop, the model can continuously learn from real-world observations and rapidly assimilate new knowledge. This iterative process ensures that the model stays up-to-date and maintains its effectiveness over time, effectively tackling the challenges posed by an ever-changing network environment.

To address concept drift, the unsupervised model's performance is continuously monitored, and the system actively checks for significant decreases in its anomaly detection accuracy. Various drift detection techniques, such as Drift Detection Method (DDM) or Page-Hinkley Test, are employed to promptly identify concept drift occurrences in real-time. Once concept drift is detected, the unsupervised model is retrained using recent data, including the labeled instances provided by domain experts. This timely update enables the model to adapt to the evolving patterns in the cloud server network traffic, ensuring its effectiveness in detecting anomalies under changing conditions.

Similarly, to handle data drift, the methodology dynamically adjusts the unsupervised model's weights and thresholds. Techniques like Exponentially Weighted Moving Average (EWMA) or Kullback-Leibler divergence are leveraged to adapt the model to the evolving data distribution. This data-driven model adaptation guarantees that the anomaly detection system can maintain optimal performance despite shifts in the data, efficiently mitigating the impact of data drift on anomaly detection accuracy.

By synergistically integrating active learning, human-in-the-loop feedback, and handling concept drift and data drift, the proposed methodology ensures that the anomaly detection system remains flexible, robust, and adaptable to the ever-changing cloud server network behaviors and emerging anomalies. This iterative and feedback-driven process empowers the model to continuously improve its performance and effectively tackle real-world challenges in cloud server network traffic anomaly detection. A schematic of proposed workflow is illustrated in Figure 3.1.

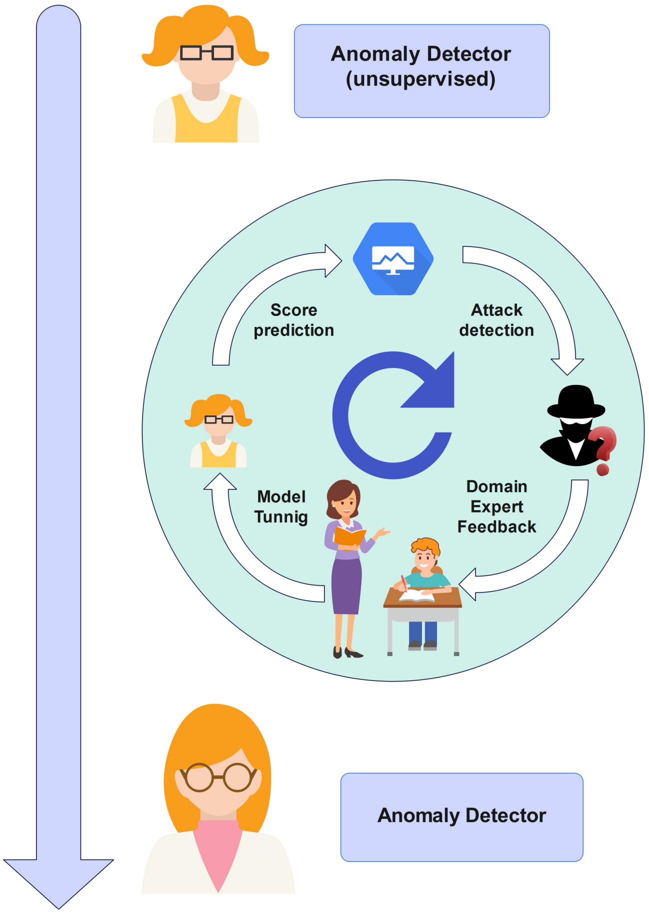

Figure 3.1: Data streaming anomaly detection workflow.

Using version control for deploying the described anomaly detection system in a cloud server environment would be highly beneficial. Version control systems, such as Git, allow for effective tracking and management of changes in the system's codebase and configurations. With the dynamic and iterative nature of the system, version control ensures that each update, including modifications to the unsupervised model, active learning components, and concept drift handling techniques, is systematically recorded, organized, and accessible.

Moreover, version control facilitates seamless integration with continuous integration and continuous deployment (CI/CD) pipelines. Automated testing and validation can be conducted at each stage of development, ensuring that changes to the system are thoroughly verified before deployment to production. This reduces the risk of introducing errors or regressions into the live system and helps maintain its effectiveness and reliability.

3.4 Data Streaming Pipeline

The data streaming pipeline architecture is specifically engineered to handle the vast and continuous flow of time-series data originating from various sources, including cloud servers, applications, and logs. The process begins with data ingestion, where information from these diverse sources is collected and routed into the system for further processing. Apache Kafka plays a central role in efficiently managing this data ingestion process. As a distributed and fault-tolerant message broker, Kafka employs a publish-subscribe model, enabling data producers to publish messages to specific topics, while consumer applications subscribe to these topics for data consumption. This decoupling of producers and consumers ensures scalable and adaptable data processing to cater to varying workloads.

Apache Kafka's partitioning capability is a key feature that allows data to be divided into multiple partitions distributed across nodes within the Kafka cluster. This distribution of data enables parallel processing by multiple consumers, ensuring high throughput. Moreover, partition replication ensures fault tolerance, safeguarding against data loss in the event of node failures. These characteristics make Kafka an ideal solution for efficiently handling the high volume and velocity of real-time time-series data streaming from cloud servers.

Once the data undergoes ingestion and is streamed through Kafka, it enters the data processing and anomaly detection phase. For this crucial task, Apache Spark, a powerful distributed data processing engine, is employed. Spark Streaming, a component of Spark, processes the incoming data in micro-batches, enabling low-latency and near-real-time analysis. This micro-batch processing approach ensures that data can be rapidly processed and analyzed, making it well-suited for timely anomaly detection on the time-series data generated by cloud servers. Additionally, Spark's in-memory caching capability further optimizes processing speed, enhancing the efficiency of the anomaly detection process. This pipeline is demonstrated in Figure 3.2.

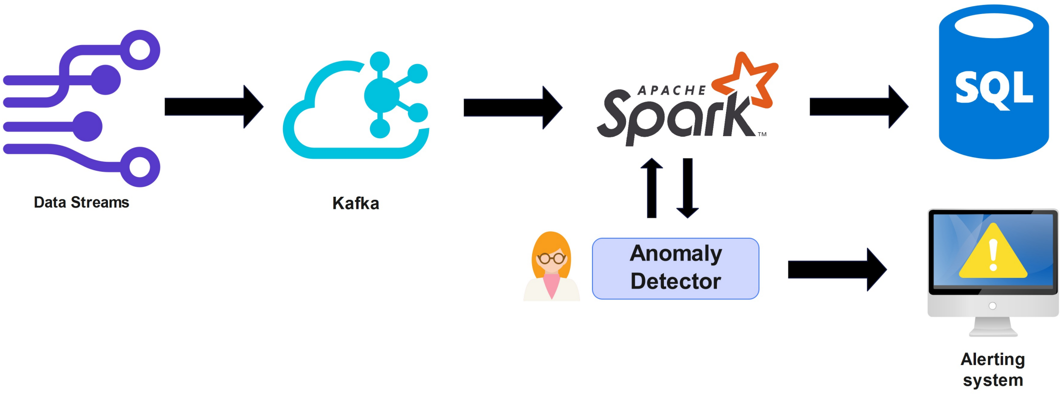

Figure 3.2: Data streaming pipeline.

Conclusion

In conclusion, our proposed anomaly detection system, integrating active learning and human-in-the-loop feedback, presents a robust and adaptive solution for detecting DDoS attacks in cloud services. The combination of machine learning algorithms and domain expert feedback achieves enhanced accuracy and efficiency in real-time anomaly detection. However, challenges related to potential bias in human labeling, scalability for large-scale data streams, and evaluation against various anomaly types remain. Future work involves exploring deep learning models, leveraging historical feedback, and addressing these challenges to enhance the system's effectiveness and applicability. Overall, our research contributes valuable insights into building resilient security systems in cloud computing through the integration of human intelligence and machine learning.

References

[1] Xingjiao Wu, Luwei Xiao, Yixuan Sun, Junhang Zhang, Tianlong Ma, and Liang He. A survey of human-in-the-loop for machine learning. *Future Generation Computer Systems*, 135:364–381, oct 2022.

[2] Xueying Zhan, Qingzhong Wang, Kuan hao Huang, Haoyi Xiong, Dejing Dou, and Antoni B. Chan. A comparative survey of deep active learning, 2022.

[3] Jianping Gou, Baosheng Yu, Stephen J. Maybank, and Dacheng Tao. Knowledge distillation: A survey. *International Journal of Computer Vision*, 129(6):1789–1819, mar 2021.