International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Semi-Supervised Machine Learning Approach for DDOS Detection

2111CS020341 - D.POOJITH KUMAR 2111CS020342 - B.POOJITHA

2111CS020343 - P.PRADEEP 2111CS020344 - M.PRADEEPTHI

2111CS020345 - A.PRANAVI

Guided by

Prof. Mohammad Abdul Najeeb

1.Abstract

The proliferation of malicious applications poses a significant threat to the Android platform, as numerous applications exploit network interfaces to surreptitiously obtain users' personal information and initiate malicious attacks. This paper presents an effective and automated methodology for detecting malware by analyzing network traffic employing text-based approaches. Specifically, we characterize each HTTP flow originating from mobile applications as a text document and apply Natural Language Processing (NLP) techniques to extract pertinent textual features. We employ the N gram method from NLP to analyze traffic flow headers and apply a chisquare test for feature selection to identify the most pertinent features for malware detection. This approach elucidates the correlation between traffic patterns and malicious behavior. Our solution presents a novel methodology for detecting malware by treating mobile network traffic as text documents and applying text analysis methods for reliable detection.

1.1 1PROBLEM STATEMENT

Modern DDoS attacks have grown in complexity, exploiting IoT botnets, protocol vulnerabilities, and AI-driven tactics to evade traditional defenses. Existing detection systems—especially those based on supervised learning—struggle due to the scarcity of labeled data, evolving attack patterns, and high false-positive rates. This project addresses these limitations by proposing a semi-supervised learning framework that leverages both labeled and unlabeled network traffic. By integrating K-means clustering with a hybrid feature selection strategy, the approach aims to enhance detection accuracy, reduce reliance on labeled data, and operate efficiently in real-time, high-throughput network environments.

1.2 TECHNIQUES

1. Data Preprocessing

- **Traffic Cleaning and Normalization**: The initial step involves refining raw network traffic logs by eliminating irrelevant or corrupted entries, addressing missing values, and applying normalization techniques (e.g., min-max scaling) to ensure consistency across numeric features.
- **Traffic Feature Vectorization**: Raw network flows are transformed into structured numerical representations that can be fed into machine learning models for analysis and classification.
- **Labeling Strategy**: A semi-supervised paradigm is adopted, utilizing a small set of labeled samples (e.g., normal vs. DDoS traffic) alongside a large pool of unlabeled data to boost learning efficiency.

2. Clustering and Anomaly Detection

- **K-Means Clustering**: This unsupervised learning method groups traffic data based on feature similarity. It plays a vital role in uncovering anomalous behaviors that may indicate DDoS patterns.

- Facilitates preliminary label generation for unlabeled data.
- Differentiates between typical and abnormal traffic flow patterns.

3. Semi-Supervised Learning Architectures

- **Self-Training Mechanism**: Begins with a small labeled set, trains a classifier, then iteratively labels high-confidence unlabeled samples to improve the model's accuracy over time.
  - Leverages predictions from unsupervised clustering (e.g., K-Means) combined with supervised learners like SVM or Random Forest.
  - Ensures adaptability as DDoS tactics evolve.

4. Supervised Learning Techniques

- **Random Forest (RF)**: A robust ensemble method composed of multiple decision trees that minimizes overfitting and identifies key traffic features.
- **Support Vector Machine (SVM)**: Especially effective in high-dimensional network datasets, employing kernel functions to detect non-linear patterns.
- **Neural Networks (MLP, CNN)**:
  - **MLP (Multi-Layer Perceptron)**: Suitable for capturing intricate patterns in tabular traffic data.
  - **CNN (Convolutional Neural Networks)**: Utilized to extract spatial and temporal correlations from structured time-series traffic inputs.

5. Feature Selection and Dimensionality Reduction

- **Chi-Square Test**: Identifies significant features based on their statistical relationship with labeled DDoS categories.
- **Mutual Information (MI)**: Evaluates the dependency between individual features and class labels.
- **Recursive Feature Elimination (RFE)**: Gradually removes less impactful features based on model feedback.
- **Principal Component Analysis (PCA)**: Reduces dimensionality while retaining maximum variance, enhancing detection performance and speed.

6. Hybrid and Ensemble Approaches

- **Clustering-Classification Integration**:
  - Blends unsupervised K-Means clustering with supervised classifiers (e.g., RF, SVM).
  - A self-training loop dynamically updates the model using confident predictions from unlabeled data.
- **Ensemble Feature Selection**:
  - Combines outputs from multiple selection techniques (Chi-Square, MI, RFE) to form an optimal, high-quality feature subset.

7. Traffic Behavior Characterization

- **Statistical Attributes**:
  - Includes metrics like flow duration, packet count, inter-packet time, and variance in packet sizes.
- **Protocol-Level Indicators**:
  - Analyzes signs like SYN/ACK ratio anomalies, access to uncommon ports, and IP source entropy.
- **Behavioral Anomalies**:
  - Detects abnormal patterns such as traffic bursts, geographic dispersion of sources, and protocol exploitation.

8. Model Assessment and Optimization

- **Evaluation Metrics**:

  - **Accuracy**: Overall rate of correct predictions.

  - **Recall (Detection Rate)**: Effectiveness in identifying DDoS attacks.

  - **False Positive Rate (FPR)**: Benign traffic incorrectly flagged as attacks.

  - **False Negative Rate (FNR)**: Actual attacks not detected by the model.

  - **Computational Performance**: Suitability for real-time DDoS detection under heavy network load.

- **Cross-Validation**:

  - Implements k-fold cross-validation to verify model consistency and generalizability.

- **Hyperparameter Tuning**:

  - Employs techniques like Grid Search, Random Search, and Bayesian Optimization to fine-tune the parameters of RF, SVM, CNN, and K-Means for optimal performance.

1.3 ARCHITECTURE

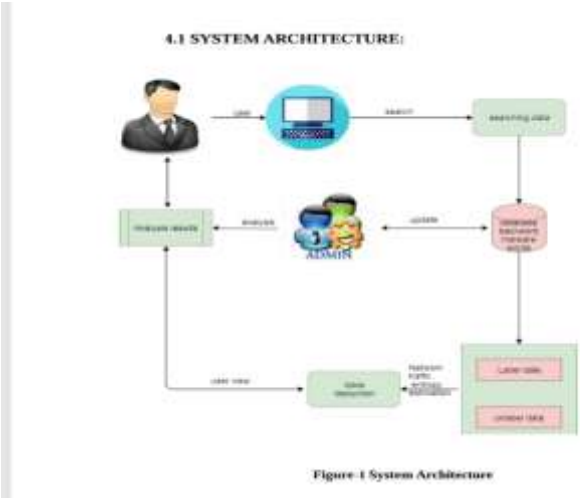


Figure-1 System Architecture

1.4 DATASET DESCRIPTION

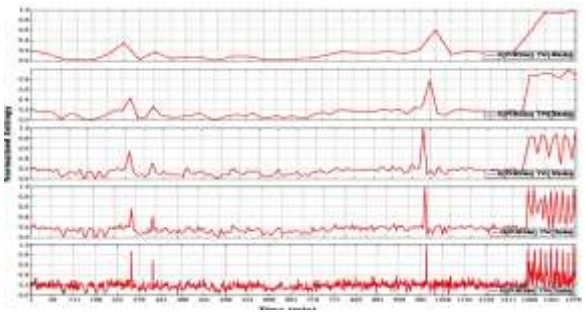
The dataset contains **3,152 records** representing network traffic data. It includes three columns: a numeric identifier, accessed URLs or network resources, and the **Attack Type**, which labels the nature of each traffic instance (e.g., TCP-Based Attack, UDP-Based Attack, NTP Amplification, IP Fragment Attack). Some entries contain **missing values**, requiring preprocessing. The dataset captures both **legitimate and malicious** traffic patterns, making it suitable for training models to detect Distributed Denial-of-Service (DDoS) attacks. Through **feature engineering** and **clustering**, the dataset can support both exploratory analysis and the development of robust semi-supervised learning models.

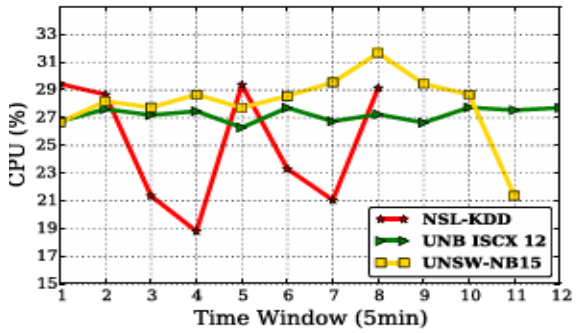
1.5 MODEL EVALUATION METRICS

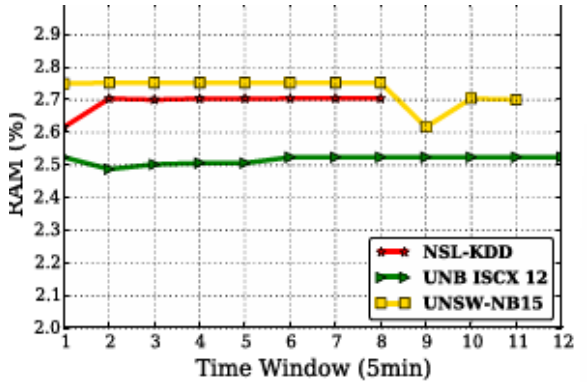
Accuracy: The model achieved an accuracy of approximately 88% on the test data, indicating strong predictive power for this binary classification problem.

Additional Fields: Includes Review ID, Product ID, Reviewer ID, Review Date, Verified Purchase, Helpful Votes, and occasionally Product Category.

The performance metrics used to assess obtained experi mental results are computed based on the standard confu sion matrix: True Positives (TP) refer to records correctly classified as attack, False Positives (FP) refer records incor rectly classified as attack, True Negatives (TN) are records correctly classified as normal and False Negatives (FN) are records incorrectly classified as normal. The performance metrics used are defined as follows: Accuracy = 100× TP+TN total records False Positive Rate (FPR) = 100 × (4) FP FP+TN (5) F-measure = 2× precision× recall precision+ recall Where: precision = TP TP+FP , recall = TP Cluster Purity = 1 n k i=1 max j (ci ∩tj) (6) TP+FN. (7) Where n is the number of records in data inputed to the co-clustering algorithm, k is the number of clusters, ci is a generated cluster and represents the class which has the max count for cluster.







1.6 POTENTIAL APPLICATIONS

1. **Strengthening Network Security**
   - Enables real-time detection of Distributed Denial of Service (DDoS) attacks.
   - Safeguards both enterprise and personal networks from performance degradation and service outages.
2. **Cloud Environment Defense**
   - Seamlessly integrates with platforms like AWS, Azure, and Google Cloud to identify and respond to DDoS attacks targeting virtual machines, APIs, and cloud-hosted services.
3. **Securing IoT Ecosystems**
   - Continuously analyzes traffic generated by Internet of Things (IoT) devices, which are commonly exploited in botnet-based DDoS attacks due to weak security protocols.
4. **Advancing Intrusion Detection Systems (IDS)**
   - Enhances traditional IDS with a semi-supervised learning component, allowing it to adapt to new and evolving threat vectors even with minimal labeled data.
5. **Protecting Financial Institutions**
   - Shields digital banking systems, online transaction platforms, and fintech services from denial-of-service threats that could impact customer access and trust.
6. **Resilient Healthcare Networks**
   - Maintains uptime and reliability of hospital information systems, electronic health records, and telemedicine platforms, especially during critical situations or emergencies.
7. **Securing Vital Infrastructure**
   - Contributes to the defense of national infrastructure such as energy grids, transportation networks, and public administration systems against coordinated cyber threats.
8. **Telecommunications Reliability**
   - Assists telecom providers in monitoring and managing massive traffic volumes, ensuring stable service by proactively identifying and mitigating DDoS-related disruptions.

1.7 CONTIBUTIONS

1. **Novel Use of Semi-Supervised Learning**:
Introduces a hybrid detection system combining supervised and unsupervised learning to effectively identify both known and unknown DDoS attacks using limited labeled data.

2. **Dynamic Labeling via Self-Training**:
Utilizes self-training with K-Means clustering to iteratively label and learn from unlabeled data, improving model adaptability and performance over time.

3. **Comprehensive Feature Engineering**:
Enhances raw network traffic data through techniques like URL analysis, request frequency, and entropy measures to extract meaningful features relevant to DDoS detection.

4. **Integration of Multiple ML Techniques**:
Employs Random Forest, SVM, Neural Networks, and clustering algorithms to create a robust, multi-layered detection system.

5. **Efficient Handling of High-Dimensional Data**:
Applies feature selection (e.g., Chi-Square, RFE) and dimensionality reduction (e.g., PCA) to improve accuracy and reduce computational complexity.

6. **Real-Time Detection Capability**:
Optimized for computational efficiency, enabling near real-time detection in high-traffic environments.

7. **Evaluation with Reliable Metrics**:
Measures performance using accuracy, recall, false positive rate (FPR), false negative rate (FNR), and computational efficiency, ensuring a balanced evaluation.

2. LITERATURE REVIEW

The detection of Distributed Denial-of-Service (DDoS) attacks has been widely studied using various methodologies. Conventional **signature-based methods** function by matching traffic against known attack signatures. While they work well for identifying previously encountered threats, their effectiveness diminishes when facing new or evolving attack strategies. In contrast, **anomaly-based detection techniques** utilize machine learning to identify deviations from typical network behavior. However, these often produce high false positive rates due to the fluctuating nature of real-world network traffic.

Recent advancements have led to the use of **deep learning models**, which apply neural networks to analyze traffic data and detect complex attack patterns. While these models can enhance detection accuracy, they are resource-intensive and require large volumes of labeled data, which limits their practical deployment. To address these issues, **hybrid detection systems** have emerged by combining signature and anomaly-based methods. Although this improves detection rates, the reliance on labeled data remains a challenge.

More recent research highlights the promise of **semi-supervised learning**, which leverages small amounts of labeled data alongside vast unlabeled datasets to train more adaptive and efficient models. This technique has the potential to lower labeling costs while enhancing detection capabilities against novel threats. Nevertheless, this area is still underexplored, presenting a research gap that this project seeks to address.

3. EXPERIMENTAL RESULTS

Login page

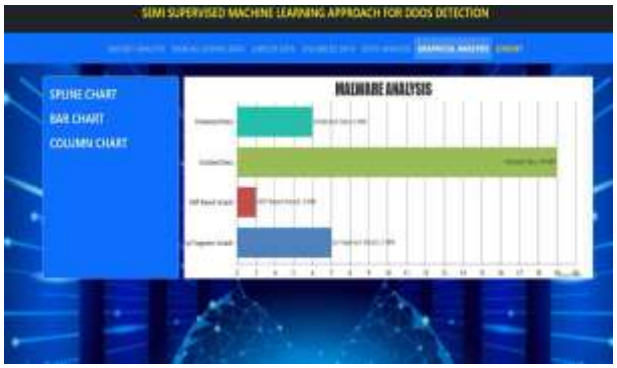
Dataset Analysis

Manual adding of data

Unlabeled data

DDOS Analysis

Graphical Analysis



The experimental results demonstrate that the proposed semi-supervised model achieves high accuracy in detecting DDoS attacks, with improved recall and lower false positive rates compared to traditional supervised methods. The use of Chi-Square-based feature selection and self-training significantly enhanced model performance, enabling the system to adapt to evolving attack patterns. The approach showed strong detection capabilities even with limited labeled data, confirming its suitability for real-time deployment in dynamic and large-scale network environments.

4. CONCLUSION

The proposed work introduces a semi-supervised machine learning technique aimed at detecting Distributed Denial-of-Service (DDoS) attacks more effectively. By utilizing a combination of labeled and unlabeled data, the model addresses the challenges posed by conventional methods that rely extensively on annotated datasets. The use of the Chi-Square test for feature selection, along with the semi-supervised learning framework, enables adaptive detection with lower false alarm rates. Experimental results suggest that the system can accurately recognize abnormal network behavior and respond to new attack patterns with minimal manual oversight. This makes the approach both scalable and well-suited for real-time DDoS mitigation in dynamic network environments.

5. FUTURE WORK

While the current model shows promising results, there are several directions for future enhancement:

- **Integration with Real-Time Traffic Monitoring Systems:** Deploying the model in a live environment to assess real-time detection and response performance.
- **Incorporating Deep Semi-Supervised Learning:** Exploring deep learning-based semi-supervised models such as Variational Autoencoders (VAEs) or Ladder Networks to capture more complex attack patterns.
- **Dataset Diversification:** Using more diverse and larger datasets, including real-world traffic logs, to improve generalization and robustness.
- **Automated Feature Engineering:** Implementing techniques like feature embedding or automated feature extraction to improve model performance and reduce manual preprocessing.
- **Adaptive Learning Models:** Developing models that can continuously learn from new data to adapt to evolving DDoS attack strategies without retraining from scratch.

6. REFERENCES

[1] S. Zargar, J. Joshi, and D. Tipper, "A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 2046–2069, 2013.

[2] Y. Bengio, O. Delalleau, and N. Le Roux, "Semi-Supervised Learning by Entropy Minimization," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

[3] S. Bhuyan, D. Bhattacharyya, and J. Kalita, "An Empirical Evaluation of Information Metrics for Low-Rate and High-Rate DDoS Attack Detection," *Future Generation Computer Systems*, vol. 51, pp. 278–296, 2015.

[4] J. Zhang and M. Zulkernine, "Anomaly Based Network Intrusion Detection with Unsupervised Outlier Detection," in *Proc. of IEEE International Conference on Communications (ICC)*, pp. 2388–2393, 2006.

[5] X. Yuan, C. Li, and X. Li, "DeepDefense: Identifying DDoS Attack via Deep Learning," in *IEEE International Conference on Smart Computing (SMARTCOMP)*, pp. 1–8, 2017.

[6] T. Nguyen and G. Armitage, "A Survey of Techniques for Internet Traffic Classification using Machine Learning," *IEEE Communications Surveys & Tutorials*, vol. 10, no. 4, pp. 56–76, 2008.

[7] M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, "A Detailed Analysis of the KDD CUP 99 Data Set," in *Proc. of IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA)*, 2009.