Evaluation of SVM Kernel Functions to Detect DDoS attacks

Kishorebabu Dasari*

Keshav Memorial Institute of Technology, Hyderabad , India, 500029.

Srinivas Mekala

KG Reddy College of Engineering and Technology, India, 501504.

Abstract

The digital landscape undergoes a constant transformation with the rise of novel technologies. As novel advancements emerge, cyberattacks become more commonplace and cunning, exploiting these very innovations. A Distributed Denial of Service (DDoS) attack is a type of cyber-attack that aims to compromise the availability of information security, thereby disrupting services for legitimate users. Detecting DDoS attacks is essential to lessen their impact. This paper introduces a method for detecting DDoS attacks using network flow features, as opposed to the more commonly utilized network type features. The suggested method utilizes the Support Vector Machine (SVM) classification algorithm, employing different kernel functions such as linear, RBF, polynomial, and sigmoid. To identify uncorrelated feature subsets, Pearson, Spearman, and Kendall correlation methods were utilized. Experiments were conducted using the CIC-DDoS2019 dataset from the Canadian Institute for Cyber Security. The study found that using the uncorrelated feature subset identified by Pearson's method resulted in superior performance with SVM's RBF and polynomial kernel functions.

Category: Ubiquitous Computing

Keywords: DDoS attacks; Correlation; Support Vector Machine

I. INTRODUCTION

The digital landscape undergoes a constant transformation with the rise of novel technologies. This progress, however, brings a hidden cost. Cyberattacks are becoming increasingly frequent and intricate, requiring continual adaptation of security measures. Distributed Denial of Service (DDoS) [1] is a type of cyber-attack where the attacker overwhelms servers with an immense amount of traffic, utilizing resources such as zombies and botnets, thereby preventing legitimate users from accessing the server's resources. The motivations behind DDoS attacks are financial, economic benefits, cyber warfare's and personal revenge or intellectual challenge. DDoS attack architecture consists of attacker, control handler, botnets or zombies and victim server.

DDoS attacks[2] classified into three types such as volumetric, protocol and application based on attack launching approach. Volumetric DDoS attack utilizes an immense volume of network traffic to fully overwhelm and exhaust the available network bandwidth. Protocol based DDoS attacks launching approach use malicious connection requests by targeting layer 3 and 4 of OSI/ISO network inorder to completely reducing processing capacity. Application DDoS attacks launch attacks by exploit the weaknesses in layer 7 in order to consume the resources. Detecting DDoS attacks early and accurately is crucial to minimize losses in various

areas such as reputation and finances.

DDoS attack, types of DDoS attacks and DDoS attacks consequences are discussed here. Section 2 discussed related work. Section 3 of this paper discussed the methodology. The results are discussed in section 4 of this paper. Section 5 concludes this paper.

II. RELATED WORK

Wang et al. [3] presented a DDoS attack detection system that utilizes features chosen by a random forest algorithm. These features are then fed into a Support Vector Machine (SVM) for classification. Cheng et al. [4] investigated a DDoS attack detection approach that leverages a Support Vector Machine (SVM) for classification. To improve efficiency, they employed Principal Component Analysis (PCA) for feature selection. Ramamoorthi et al. [5] introduced a DDoS attack detection system that builds upon a modified Support Vector Machine (SVM) with string kernels. This approach aims to improve the accuracy of DDoS attack identification. Daneshgadeh et al. [6] proposed hybrid method with combination of Shanon entrophy, Kernel online anamoly detection and SVM for DDoS attack detection. Juneja et al. [7] presented a rule-based Support Vector Machine (SVM) model for detecting various types of DDoS attacks. This approach combines rule sets with SVM classification to enhance the system's ability to identify different DDoS attack variants. Kato et al. [8] proposed packet analysis based DDoS attack detection with SVM RBF kernel. Amir et al. [9] proposed DDoS attacks detection mechanism with different feature engineering mechanisms to collect features and different classification algorithms. Hoyos et al. [10] developed a prototype system for detecting DDoS attacks using a Support Vector Machine (SVM).

Researchers employ machine learning classification algorithms alongside various feature selection methods to detect DDoS attacks. This study explores the utilization of a Support Vector Machine (SVM) classifier with diverse kernels to identify DDoS attacks. Proposed Feature selection is Correlation feature selection, which is filter based feature selection method. And this study use flow features instead of type features of communication network.

III. METHODOLOGY

This research utilizes the CIC-DDoS2019 dataset[11], sourced from the Canadian Institute for Cyber Security, encompassing 11 distinct types of DDoS attacks and 87 network traffic features. CICFlowmeter employed to convert the pcap files into CSV format. Correlation methods applied to select uncorrelated features, and Support Vector Machines (SVM) with various kernel functions used to distinguish DDoS attack classes from benign traffic.

Correlation methods [12-13] were employed to identify similarities among features. In this research, features with a correlation coefficient of 80 or higher were considered correlated. The study employed Pearson, Spearman, and Kendall correlation methodologies.

Pearson correlation coefficient calculated by

$$\rho(X,Y)=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\cdot\sum_{i=1}^{n}(y_i-\bar{y})^2}} \quad (1)$$

$\bar{x}$ is the mean of X sample, $\bar{y}$ is the mean of Y sample

Spearman correlation coefficient calculated by

$$\rho = 1-\frac{6\sum d_i^2}{n(n^2-1)} \quad (2)$$

d is the difference value of ranks

Kendall correlation coefficient calculated by

$$\tau = \frac{N_c-N_d}{N_c+N_d} \quad (3)$$

$N_c$is the concordant number, $N_d$is the discordant number

Support Vector Machine (SVM) [14] is a supervised machine learning classification algorithm. It classifies the data based on support vectors distance from hyper plane. Kernel functions are used for convert the non-linear data to linear data. This study applied linear, RBF, polynomial, and sigmoid kernel functions.

Linear kernel function defined as

$$k(x_i,x_j) = x_i * x_j \quad (4)$$

Radial Basis Function (RBF) kernel function defined

as

$$k(x_i, x_j) = \exp(-\gamma||x_i - x_j||^2) \quad (5)$$

Polynomial kernel function defined as

$$k(x_i, x_j) = (1 + x_i * x_j)^d \quad (6)$$

Sigmoid kernel function defined as

$$k(x_i, x_j) = \tanh(\propto x^T y + c) \quad (7)$$

IV. RESULTS AND DISCUSSION

In this study, experiments were conducted on a DDoS attack dataset sourced from the CIC-DDoS2019 dataset, developed by the Canadian Institute for Cyber Security. The primary objective was to detect DDoS attacks using only network flow features, excluding network type features. Consequently, seven network type features, such as source and destination IP addresses, were removed from the original set of 87 features. During pre-processing, records with missing values were eliminated, and target class labels were encoded as 0 and 1. Additionally, features with a variance threshold of 0 or 0.01 were discarded. The dataset initially had 12 features with a variance of 0 and 3 features with a variance of 0.01.

Uncorrelated features were identified by excluding correlated features selected by correlation methods. The study employed Pearson, Spearman, and Kendall correlation methods to define respective uncorrelated feature subsets. The dataset contained 39 Pearson, 46 Spearman, and 45 Kendall correlated features. After removing these correlated features, the uncorrelated feature subsets included 25 features for Pearson, 18 for Spearman, and 19 for Kendall. The common uncorrelated features across these methods constituted the PSK-uncorrelated feature subset, consisting of 15 features.

The study evaluated the SVM classification algorithm with linear, RBF, polynomial, and sigmoid kernel functions using Pearson, Spearman, Kendall, and PSK uncorrelated feature subsets for DDoS attack detection. The classification performance for distinguishing between DDoS attacks and benign classes was evaluated using metrics such as accuracy, K-fold cross-validation accuracy, log-loss, ROC-AUC score, and specificity.

Table 1 predicts the accuracy results of SVM kernel functions for DDoS attacks with different features subsets. Pearson uncorrelated features subsets gives the better accuracy with all kernels except sigmoid kernel functions. Sigmoid kernel function gives better results with uncorrelated features subsets by Spearman correlation method. RBF kernel produce the better accuracy results with features selected by Pearson and Kendall correlation methods. Linear kernel with uncorrelated features selected by Spearman correlation method and PSK method gives better accuracy for DDoS attack detection. SVM RBF kernel function with uncorrelated features selected by Pearson correlation method gives best accuracy than others. PSK uncorrelated features subset also gives good log-loss results with all kernel functions.

Table 1. SVM Kernel functions accuracy results

SVM Kernels	Pearson	Spearman	Kendall	PSK
Linear	97.16	97.00	96.51	97.00
RBF	98.09	96.34	97.49	96.67
Poly	97.11	96.89	96.94	96.67
Sigmoid	93.29	94.10	93.07	93.18

Table 2 presents the K-fold cross-validation (KFC) accuracy results of SVM kernel functions for detecting DDoS attacks using different feature subsets. The Pearson uncorrelated feature subset achieves the highest accuracy with the RBF and polynomial SVM kernel functions. The linear kernel performs best with the Spearman uncorrelated feature subset, while the sigmoid kernel shows better KFC accuracy with the Kendall uncorrelated feature subset. Among all kernel functions, the SVM RBF kernel consistently delivers superior accuracy across all uncorrelated feature subsets. Additionally, the PSK uncorrelated feature subset demonstrates strong KFC accuracy results, comparable to other uncorrelated feature subsets acrossall kernel functions.

Table 2. SVM Kernel functions KFC accuracy results

SVM Kernels	Pearson	Spearman	Kendall	PSK
Linear	96.4642% (0.3749%)	96.9420% (0.1392%)	96.4642% (0.3749%)	96.4642% (0.3749%)
RBF	97.2287% (0.2682%)	97.0785% (0.2640%)	97.0785% (0.2640%)	97.0785% (0.2640%)
Poly	97.0375% (0.2006%)	96.9420% (0.1392%)	97.0785% (0.1580%)	97.0375% (0.2006%)

Sigmoid	93.1604% (0.6187%)	93.3379% (0.7232%)	93.5836% (0.9929%)	93.3379% (0.7232%)

Table3 presents the log loss results of SVM kernel functions for detecting DDoS attacks using various feature subsets. The Pearson uncorrelated feature subset produces the best log loss for all kernel functions, with the exception of the sigmoid kernel. The sigmoid kernel function achieves better log loss results with the Spearman uncorrelated feature subset. The SVM RBF kernel produces superior log loss results with the Pearson and Kendall feature subsets. The linear kernel shows improved log loss results with the Spearman and PSK uncorrelated feature subsets. Additionally, the PSK uncorrelated feature subset provides favorable log loss results similar to other uncorrelated feature subsets across all SVM kernel functions.

Table 3. SVM Kernel functions Log-loss results

SVM Kernels	Pearson	Spearman	Kendall	PSK
Linear	0.9801	1.0369	1.2066	1.0369
RBF	0.6599	1.2632	0.8672	1.1500
Poly	0.9992	1.0746	1.0558	1.1500
Sigmoid	2.3189	2.0361	2.3943	2.3566

Table 4 presents the ROC-AUC results of SVM kernel functions for detecting DDoS attacks using various feature subsets. The SVM polynomial kernel achieves the best ROC-AUC results across all uncorrelated feature subsets, except for the Pearson subset. The Pearson uncorrelated features yield the highest ROC-AUC value with the SVM RBF kernel. The linear kernel shows better ROC-AUC results with the Pearson uncorrelated feature subset, while the sigmoid kernel performs best with the Spearman uncorrelated feature subset. The PSK uncorrelated feature subset also delivers strong ROC-AUC results with the RBF and polynomial kernel functions. Figures 1 to 3 illustrate the ROC-AUC curves of the SVM classifier for DDoS attack detection using the Pearson, Spearman, Kendall, and PSK uncorrelated feature subsets.

Table 4. SVM Kernel functions ROC-AUC results

SVM Kernels	Pearson	Spearman	Kendall	PSK
Linear	0.9926	0.9893	0.9924	0.9879
RBF	0.9957	0.9910	0.9966	0.9904
Poly	0.9954	0.9923	0.9971	0.9918
Sigmoid	0.9245	0.9379	0.9111	0.9220

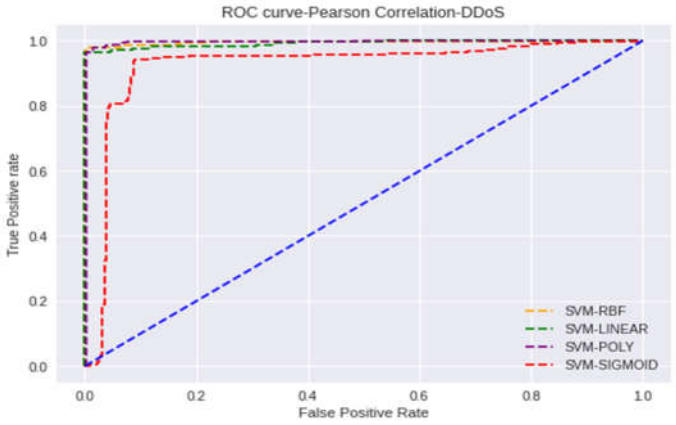

Fig. 1. functions ROC curves for SVM kernels to DDoS attacks detection with Pearson uncorrelated features

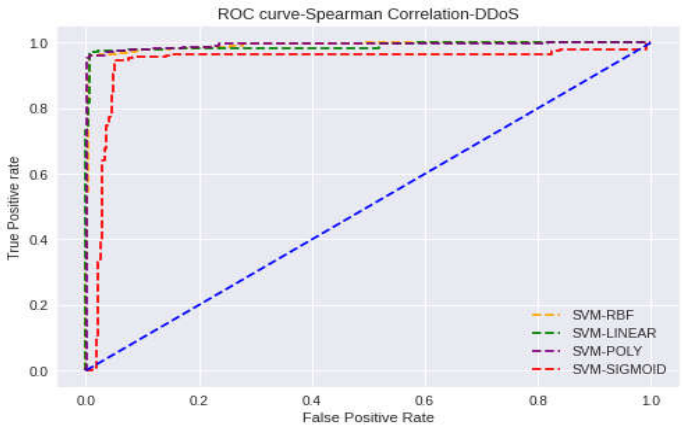

Fig. 2. ROC curves for SVM kernels to DDoS attacks detection with Spearman uncorrelated features

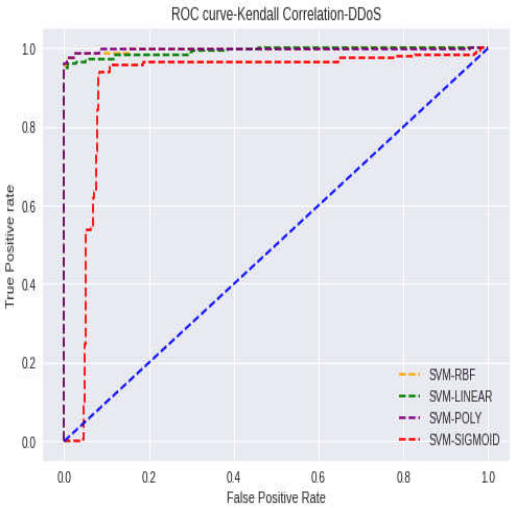

Fig. 3. ROC curves for SVM kernels to DDoS attacks detection with Kendall uncorrelated features

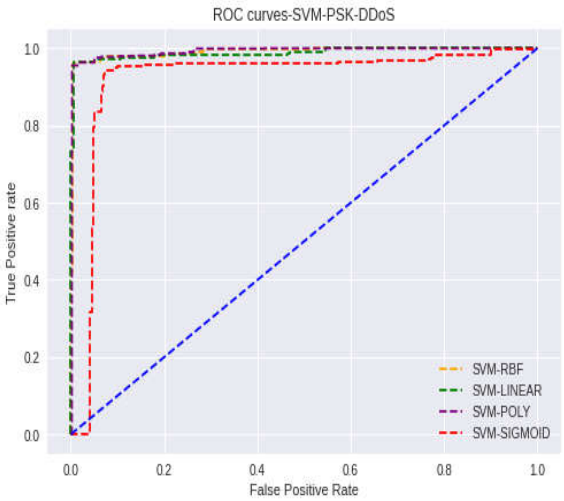

Fig. 4. ROC curves for SVM kernels to DDoS attacks detection with PSK uncorrelated features

Table 5 presents the specificity results of SVM kernel functions to detecting DDoS attacks using various feature subsets. The SVM RBF kernel achieves the highest specificity across all uncorrelated feature subsets. The Pearson and Spearman uncorrelated features provide excellent specificity values with all SVM kernel functions, except for the sigmoid kernel. The sigmoid kernel produces poor specificity results with all uncorrelated feature subsets. Additionally, the PSK uncorrelated feature subset shows strong specificity results across all kernel functions.

Table 5. SVM Kernel functions Specificity results

SVM Kernels	Pearson	Spearman	Kendall	PSK
Linear	1.00	0.99	0.97	0.98
RBF	1.00	0.99	0.99	0.99
Poly	0.99	0.99	0.99	0.98
Sigmoid	0.91	0.93	0.92	0.93

All SVM kernels with all uncorrelated feature subsets produces good the value of 1 precision, recall and F1-score values for DDoS attack detection.

IV. CONCLUSION

This study focused on using network flow features rather than network type features to detect DDoS attacks. It employed uncorrelated independent features for classifying the dependent target class feature. Pearson, Spearman, and Kendall correlation methods were used to identify uncorrelated features. The study also utilized PSK uncorrelated features, which are the common features identified by Pearson, Spearman, and Kendall methods, in conjunction with various SVM kernel functions to detect DDoS attacks. The findings indicated that the Pearson uncorrelated feature subset demonstrated superior performance across all SVM kernel functions, outperforming other uncorrelated feature subsets. Additionally, the PSK uncorrelated feature subset displayed favorable classification outcomes. Among the SVM kernels, the RBF and polynomial kernels yielded the most accurate DDoS attack classifications with all feature subsets, whereas the sigmoid kernel function showed the least favorable results. Notably, combining the Pearson uncorrelated feature subset with RBF and polynomial kernels resulted in the highest classification accuracy. Future work will explore the use of neural network classification models to further enhance this approach.

REFERENCES

1. Kishore Babu Dasari, Nagaraju Devarakonda, "Detection of DDoS Attacks Using Machine Learning Classification Algorithms", International Journal of Computer Network and Information Security(IJCNIS), Vol.14, No.6, pp.89-97, 2022. DOI:10.5815/ijcnis.2022.06.07.
2. Dasari, K.B., Devarakonda, N. (2021). Detection of different DDoS attacks using machine learning classification algorithms. Ingénierie des Systèmes d'Information, Vol. 26, No. 5, pp. 461-468. https://https://doi.org/10.18280/isi.260505.
3. C. Wang, J. Zheng and X. Li, "Research on DDoS Attacks Detection Based on RDF-SVM," 2017 10th International Conference on Intelligent

Computation Technology and Automation (ICICTA), Changsha, China, 2017, pp. 161-165, doi: 10.1109/ICICTA.2017.43.

4. L. Cheng, P. Xiao, X. Zhang, Y. Liu, L. Gao and P. Zhou, "Research on DDoS detection method based on super logarithm and SVM algorithm," 2022 5th International Conference on Data Science and Information Technology (DSIT), Shanghai, China, 2022, pp. 1-8, doi: 10.1109/DSIT55514.2022.9943866.
5. A. Ramamoorthi, T. Subbulakshmi and S. M. Shalinie, "Real time detection and classification of DDoS attacks using enhanced SVM with string kernels," 2011 International Conference on Recent Trends in Information Technology (ICRTIT), Chennai, India, 2011, pp. 91-96, doi: 10.1109/ICRTIT.2011.5972281.
6. S. Daneshgadeh, T. Kemmerich, T. Ahmed and N. Baykal, "An Empirical Investigation of DDoS and Flash Event Detection Using Shannon Entropy, KOAD and SVM Combined," 2019 International Conference on Computing, Networking and Communications (ICNC), Honolulu, HI, USA, 2019, pp. 658-662, doi: 10.1109/ICCNC.2019.8685632.
7. K. Juneja and C. Rana, "A Rule Framed SVM Model for Classification of Various DDOS Attack in Distributed Network," 2018 2nd International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE), Ghaziabad, India, 2018, pp. 96-101, doi: 10.1109/ICMETE.2018.00032.
8. Kato, Keisuke, and Vitaly Klyuev. "An intelligent ddos attack detection system using packet analysis and support vector machine." IJICR 14.5 (2014): 3.
9. Aamir, M., Zaidi, S.M.A. DDoS attack detection with feature engineering and machine learning: the framework and performance evaluation. Int. J. Inf. Secur. 18, 761–785 (2019). https://doi.org/10.1007/s10207-019-00434-1.
10. Hoyos Ll, Manuel S., et al. "Distributed denial of service (ddos) attacks detection using machine learning prototype." Distributed Computing and Artificial Intelligence, 13th International Conference. Springer International Publishing, 2016.
11. Iman Sharafaldin, Arash Habibi Lashkari, Saqib Hakak, and Ali A. Ghorbani, "Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy", IEEE 53rd International Carnahan Conference on Security Technology, Chennai, India, 2019.
12. Dasari, K.B., Devarakonda, N. (2022). TCP/UDP-based exploitation DDoS attacks detection using AI classification algorithms with common uncorrelated feature subset selected by Pearson, Spearman and Kendall correlation methods. Revue d'Intelligence Artificielle, Vol. 36, No. 1, pp. 61-71. https://doi.org/10.18280/ria.360107.
13. Dasari, K.B., Devarakonda, N. (2022). Detection of TCP-based DDoS attacks with SVM classification with different kernel functions using common uncorrelated feature subsets. International Journal of Safety and Security Engineering, Vol. 12, No. 2, pp. 239-249. https://doi.org/10.18280/ijsse.120213.
14. K. B. Dasari and N. Devarakonda, "SynFlood DDoS Attack Detection with SVM Kernels using Uncorrelated Feature Subsets Selected by Pearson, Spearman and Kendall Correlation Methods," 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA), 2022, pp. 1-6, doi: 10.1109/ICCSEA54677.2022.9936114.
15. Mekala, S., Padmaja Rani Supervisor, B., & Padmaja Rani, B. (2020). Article ID: IJARET_11_11_121 Kernel PCA Based Dimensionality Reduction Techniques for Preprocessing of Telugu Text Documents for Cluster Analysis. International Journal of Advanced Research in Engineering and Technology, 11(11), 1337–1352. https://doi.org/10.34218/IJARET.11.11.2020.121.