https://doi.org/10.48009/2_iis_2018_1-12
Issues in Information Systems
Volume 19, Issue 2, pp. 1-12, 2018

AN EFFICIENT HYBRID MODEL FOR DETECTING DISTRIBUTED DENIAL OF SERVICE (DDOS) ATTACKS IN CLOUD COMPUTING USING MULTIVARIATE CORRELATION AND DATA MINING CLUSTERING TECHNIQUES

Anteneh Girma, Robert Morris University, girma@rmu.edu
Ping Wang, Robert Morris University, wangp@rmu.edu

ABSTRACT

The distributed nature of cloud computing makes it vulnerable and prone to sophisticated distributed intrusion attacks such as Distributed Denial of Service (DDoS) attacks. In order to detect those network attacks and respond swiftly, there must be a reliable defense system designed to distinguish anomalies embedded in legitimate traffic. In order to ensure the high availability of any offered services, the data center resources must be protected from DDoS threats. The existing solutions for monitoring incoming traffic and detecting DDoS attacks have excessive false alarms and become ineffective in early detection of high level flooding attacks and resolution of the cloud service availability issues. Therefore, it is necessary to devise a model that can detect DDoS attacks and serve the legitimate users with available resources with minimal downtime. This research paper addresses this need by investigating the multivariate correlation among the selected and ranked features. This study presents the promising performance results of our proposed comprehensive hybrid solution model using DBSCAN and Entropy, discusses the research findings, and visualizes the experimental results to show the degree of parametrical dependency among the selected features and the effectiveness of our multivariate correlational approach.

Keywords: DDoS Detection, Data Mining, Multivariate Correlation, Cloud Security, DSACAN Clustering

INTRODUCTION

The vulnerability of the Internet and its open architecture makes it susceptible to various forms of network attacks. DDoS (distributed denial of service) is the major and the most serious security threat that challenges the availability of the data centers resources to Intended clients. For high availability of the offered services, data center resources must be protected from DDoS attack threats. Existing defensive schemes deployed either in centralized or decentralized (Hybrid) ways to defend the flooding DDoS attacks have some advantages and disadvantage when their features are evaluated against their deployment methods (Bhaya & EbadyManaa, 2017; D'Cruze, Wang, Sbeit, & Ray, 2018). For host-based and network-based detection mechanisms, the detection and response features are deployed at the server and routers respectively. The main disadvantage associated with the host-based defensive mechanisms is that they cannot detect the flooding attack before it reaches the victim resources (servers) and they are mainly user dependent for security. The advantage of the destination (host) based detection mechanisms is that they are cheaper and can easily control the inbound and outbound IP packets to victim resources. The disadvantage of network-based detection mechanisms include high rate of storage and processing overhead at the routers and failure to control and detect inbound traffic attack.

Defensive detection mechanisms deployed in a distributed/hybrid manner against DDoS attacks have been considered the most promising and potentially effective. The main reasons noted for the success of hybrid detection schemes include their capability of early DDoS attack detection, their ability to trace back the flooding attack to sources, and their ability to prevent the attack. This research is motivated by the need to detect and identify possible threats to the cloud computing environment both those that involve intrusions, which are the attacks from outside an organization, and those that involve misuses or attacks from within an organization. Our research uses the theory of information focusing on the idea of Entropy, which is a measure of the uncertainty or randomness associated with a random variable (data coming over the network). In our benchmark study, we have investigated the performance of our proposed integrated comprehensive detection model to determine how it performs by applying three different training and testing datasets, and by comparing varied detection prototype models. We plan to achieve the security

implementation of our proposed detection model by applying two phases: behavioral analysis and knowledge analysis. Using behavioral analysis allows us to recognize both legitimate use and any behavioral deviation. During the behavioral analysis, the network must be correctly trained to efficiently identify the intrusion; nevertheless, we focus primarily on identifying users' behavioral patterns and deviations from such patterns. Using knowledge analysis affords us the opportunity to describe a malicious behavior and add a new observed rule. Our proposed intrusion detection system gathers and analyzes data from different angles (both from the computer and the network) in order to identify and detect any possible DDoS attacks or security breaches. Such intrusions could be either attacks from within the organization (insider threats) or attacks from outside. Most importantly, the proposed solution model makes vulnerability assessment using correlation studies and data mining technologies to detect the DDoS attacks.

LITERATURE REVIEW

Cloud computing offers an efficient method to access the actual remote computing through the Internet. But its security vulnerabilities have been a major and continuous risk element for both the clients and the cloud service providers. There has been research discussion on the security threats associated with cloud computing and cloud service delivery models (Abdul, et al., 2014; Carlin & Curran, 2012; Girma, Garuba, & Li, 2015; Gonzalez, et al., 2012). The third party involvement in data processing is another security threat for the service availability and integrity and confidentiality of information in cloud computing.

Distributed Denial of Service (DDoS) has been a major threat for cloud services. DDoS attacks often target cloud services with overwhelming floods of requests that result in service down time for legitimate users (Girma, Garuba, & Goel, 2014). Various techniques have been proposed to enhance cloud computing security measures against DDoS attacks (Asha & Chitra, 2013; Ashktorab & Taghizadeh, 2012; Bace & Mell, 2001; Bhaya & EbadyManaa, 2017; Katkamwar, Puranik, & Deshpuaunde, 2012; Lonea, Popesch, & Tianfield, 2013; Shelke, Sontakke, & Gawande, 2012; Ubhale, Sahu, & Raisoni, 2013). However, these techniques are not able to resolve the sophisticated and distributed nature of the DDoS vector attacks, and the attacks are still compromising and shutting down the victims (cloud computing services) at an alarming rate.

The application of clustering technology in data mining has been highly recommended and commonly used for data analytics. Density-based spatial clustering of application with noise (DBSCAN) plays an important role in data mining research by using a large dimensional dataset to explain its advantages and its shortcomings (Khan, et al., 2014; Mumtaz & Duraiswami, 2011). There have been different DBSCAN algorithms and techniques, such as the k-distance algorithm graph model (Vijayalaksmi & Punithavalli, 2012), incremental clustering technique (Wang, et al., 2013), and the parallel clustering data mining approach with modularity of sequential application and MapReduce to improve the clustering performance (Arlia & Coppola, 2001; Fu, Hu, & Wang, 2014). Even though DBSCAN has certain issues with improvements needed as discussed by Kisilevich, Mansmann, and Keim (2010) and Su, et al. (2014), researchers agree that DBSCAN is highly effective in determining clusters of different shapes compared with other clustering techniques (Erman, Arlitt, & Mahanti, 2006; Ester, et al., 1996; Ma, et al., 2014).

THEORETICAL FRAMEWORK AND RESEARCH METHODOLOGIES

The hybrid modelling approach that we use for this paper is DBSCAN clustering and entropy. The data flow architecture and design for our proposed model has been well discussed in previous research by Girma, Garuba, and Goel (2017). The major functionality of the detection scheme, its strength, limitations, and applications are discussed and explained in Girma, et al. (2015). The major advantage of applying DBSCAN clustering is its capability to recognize and differentiate the attack from normal access and cluster separately. It also helps to detect the attacks and the normal access associated for each selected feature. Entropy also has an advantage of using the entire distribution of data to estimate the entropy among data features. The DBSCAN algorithm well fits our needs as the nature of DDoS attacks is sophisticated and hackers launch vectors of flooding attacks to the victim resources. Since we are trying to come up with a solution to detect these flooding attacks early by differentiating and assigning them in different classes(clusters) as needed, we are able to significantly mitigate the risk of both network and resource depletion with the proposed model. Figure 2 below illustrates our theoretical model.

EXTRACTION OF TESTING AND TRAINING DATA

We collected two million packet samples from the data files. Then, we applied Entropy and Data Normalization computation for packet count window for 50, 100, and 200, and produced three different training and testing datasets for our proposed hybrid model. Table 1 below shows the contents of the three datasets. Each dataset has different number of packets after the Entropy and Data Normalization computation. The Entropy and Data Normalization computation for 50, 100, and 200 consecutive packets resulted in having a dataset of 40000, 20000, and 10000 packets where each of the datasets has an equal number of normal data and attack data. The final dataset for this research is extracted from CAIDA (Center for Applied Internet Data Analysis) 2007 and CAIDA 2008 source datasets.

Table 1. Description of Training and Testing Datasets

Dataset Name	Consecutive Number of Packets Used for Entropy and Data Normalization Computation	Total Number sample Data after Computation	Number of Attack Packets	Number of Normal Packets
PCW200 Normalized File	200	10,000	5,000	5,000
PCW100 Normalized File	100	20,000	10,000	10,000
PCW50 Normalized File	50	40,000	20,000	20,000

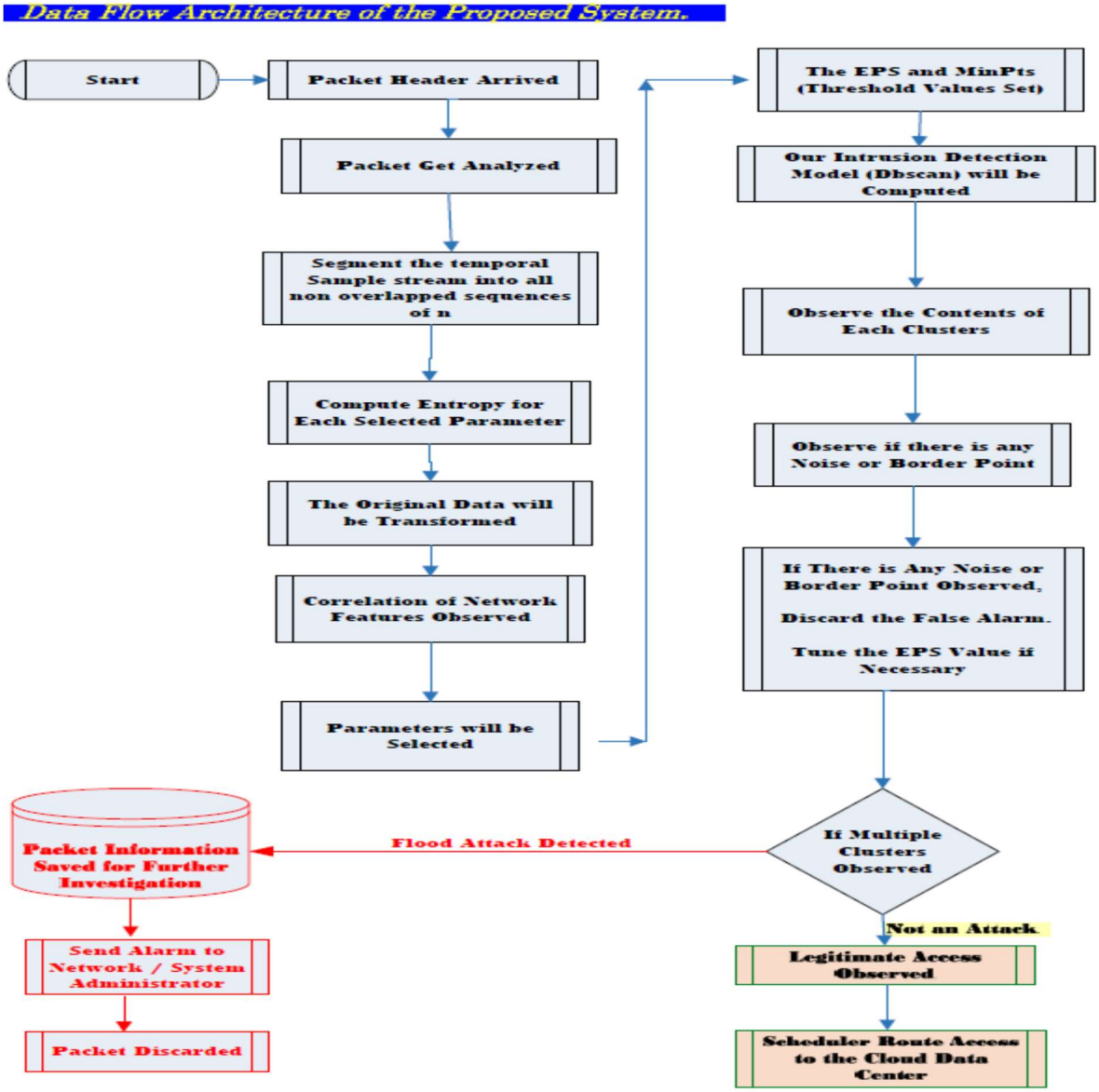

Figure 1. Data Flow Architecture of the proposed system

PERFORMANCE EVALUATION FACTORS

Many existing DDoS detection and prevention techniques lack proper selection criteria for consistent performance evaluation. The performance evaluation technique for this research is a performance measurement matrix as shown in Table 2 below. After we run our attack detection computer program, we collect the result and identify how our detection model performs in terms of distinguishing the flooding attacks from legitimate user access. There are four different possible outcomes during the detection process, and they are: True Positive, True Negative, False Positive, and False Negative.

Table 2. Performance Measurement Matrix

		Expected/Desirable DDoS Defense	
		Negative (Attack)	Positive (Normal)
DDoS Defense Model Decision	Negative (Attack)	TN	FP
	Positive (Normal)	FN	TP

After the results of the above detection function are collected, six metrics are computed to evaluate our DDoS detection model. The six metrics are accuracy, precision, reliability, false negative rate, sensitivity, and specificity.

MEASURE OF CORRELATION AND DEGREE OF DEPENDENCY

In correlational research, data analysis involves computing an index of the degree of relationship between variables. The type of correlation coefficient computed depends on the level of measures used for all variables. The observed correlation among variables provides two important functions. The first and major function is that any consistent relationship can be used to predict future events. The second function is to provide data that are either consistent or inconsistent with some currently held theory. Our correlational studies is to show the degree of dependency and relationships among the selected parameters of our different datasets. We compute the correlational coefficients using the Pearson correlation method and show its visual description to prove the degree of dependency among those selected parameters. This multivariate correlational studies significantly contribute to our detection prediction analysis and selection of our final training and testing dataset. The correlation analysis matrices for the three different datasets we use in this research are shown in Table 3 below. From the computed multivariate correlational coefficients, we can further compute the percentage of accuracy that can be used to predict the patterns and behavior of one parameter from the other. The average value of the multivariate correlational coefficient among different parameters indicates the statistical significance of our study results as shown in Table 3 and Figure 2 below.

Table 3. Analysis of Parametric Correlation for Different Packet Count Windows

Co-ordinates of Parameters From Each Correlational Matrix	Correlation Value Among Parameters For 50 Packet Count Window	Correlation Value Among Parameters For 100 Packet Count Window	Correlation Value Among Parameters For 200 Packet Count Window
Source Port - Protocol Type	**0.99494502**	**0.98666600**	**0.99432576**
Destination Port - Protocol Type	**0.98527389**	**0.99047144**	**0.99660950**
Destination Port - Source Port	**0.98437038**	**0.98816780**	**0.99277153**
Packet Length - Protocol Type	**0.98027104**	**0.97331546**	**0.98554323**
Packet Length - Source Port	**0.97787674**	**0.96984380**	**0.97954152**
Packet Length - Destination Port	**0.98375261**	**0.98032891**	**0.98821932**
Packet Window Size - Protocol Type	**0.98544875**	**0.97694332**	**0.99381846**
Packet Window Size - Source Port	**0.98294858**	**0.97172355**	**0.98759414**
Packet Window Size - Destination Port	**0.99902559**	**0.98542048**	**0.99665016**
Packet Window Size - Packet Length	**0.99473005**	**0.96925311**	**0.99152753**
Packet Acknowledgement - Protocol Type	**0.99902536**	**0.98959808**	**0.99945214**
Packet Acknowledgement - Source Port	**0.99524947**	**0.98973950**	**0.99426613**
Packet Acknowledgement - Destination Port	**0.99006067**	**0.99305845**	**0.99763720**
Packet Acknowledgement - Packet Length	**0.98472197**	**0.97520338**	**0.98647674**
Packet Acknowledgement - Packet Window Size	**0.98991973**	**0.98467248**	**0.99480983**
Average Degree of Correlation	**0.98850799**	**0.981627052**	**0.991949546**

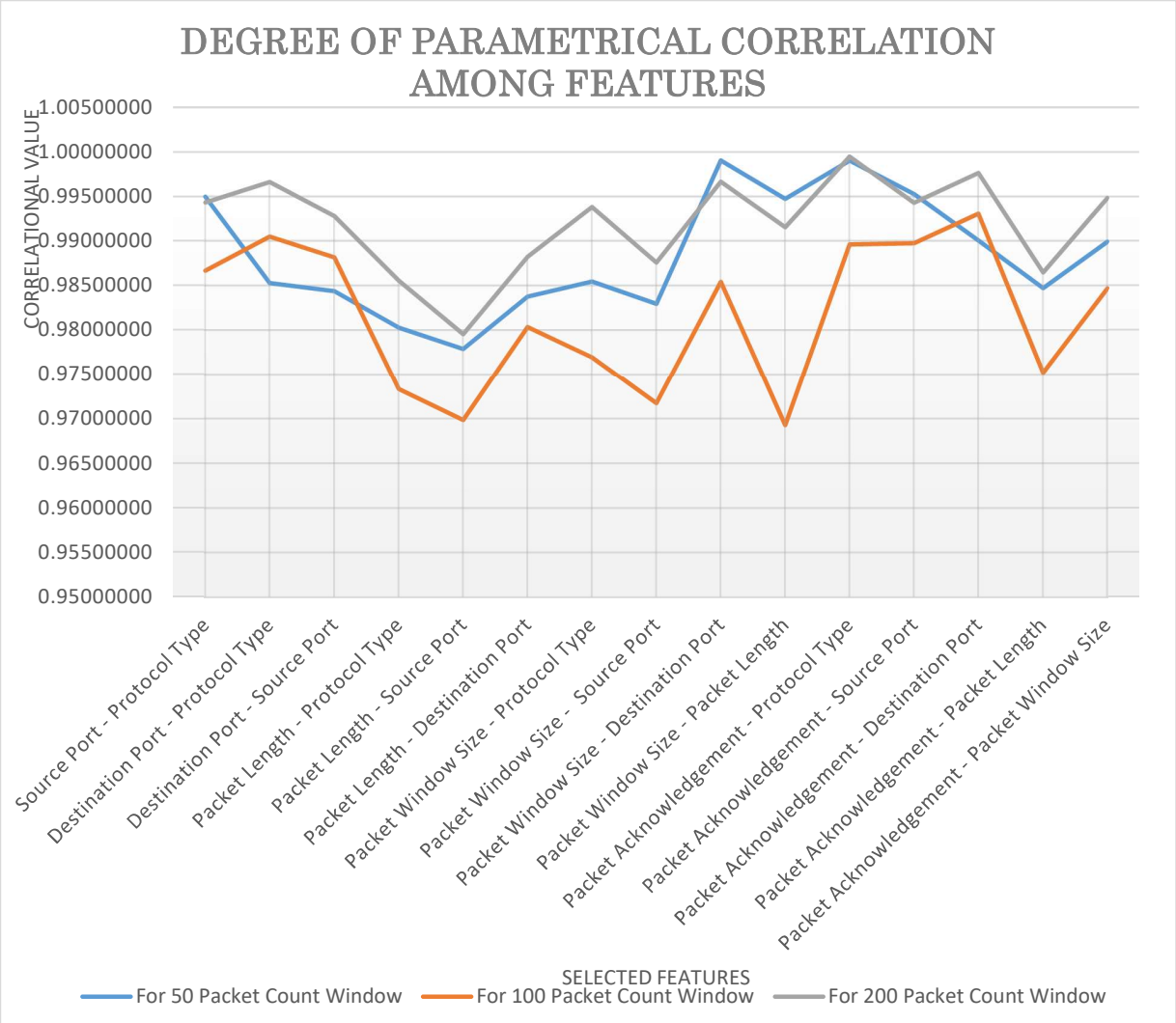

Figure 2. Degree of Parametrical Correlation among Features

EXPERIMENTAL RESULTS

We have conducted a comprehensive experiment using all three datasets extracted and listed in Table 4 below. We started from the CAIDA PCW200 and made two experiments for 60/40 and 70/30 ratio for Training/Testing dataset. Then we used the CAIDA PCW100 dataset and ran the same test run but using different EPS parameter value due to the change in dataset size (double the CAIDA PCW 200). And finally, we conducted a test ran for the CAIDA PCW50 for the same dataset ratio Training/Testing. For the CAIDA PCW50 dataset, we performed a series of tests to show how tuning the value of the EPS parameter makes a huge difference in the overall detection scenario and how our model is very friendly in-terms of making changes to the EPS parameter with tangible results. In our final assumption, we decided to have the same value for the MinPts to be constant and have a value of 50 based on the behavior of the DDoS attack. Assigning such a value could help to track some DDoS attacks initiated within the premises.

Table 4. Assumptions and Parameters Setup for CAIDA PCW 200 Dataset

Original Input Dataset	CAIDA PCW 200
Total Number of Records	10,000
Ratio of Training/Testing Phase	70/30 and 60/40
Training Dataset	70 and 60 percent of our Input dataset randomly sampled using
Testing Dataset	30 and 40 percent of our Input dataset randomly sampled using R-Programming Language
EPS-Parameter value	0.5
MinPts	50

During this training phase of our experiment, 70 percent of our original dataset (CAIDA PCW200) was used to randomly select 7000 records. R-Programming Language was used to compute the sampling process. After the sampling process was completed, we observed 3512 Attack Packets and 3488 Normal Packets for legitimate access. Table 5 below shows the result of running our DBSCAN detection model.

Table 5. Training Phase Experimental Result for CAIDA PCW 200 Dataset

Expected/Desirable DDoS Defense			
		Negative (Attack)	Positive (Normal)
DDoS Defense Model Decision	Negative (Attack)	3512	0
	Positive (Normal)	0	3488

During the testing phase of our experiment, 30 percent of our original dataset (CAIDA PCW200) was used to randomly select 3000 records. R-Programming Language is used to compute the sampling process. After the sampling process was completed, we found 1483 Attack Packets and 1517 Normal Packets for legitimate access. The result of running our DBSCAN detection model is shown in Table 6 below.

Table 6. Testing Phase Experimental Result for CAIDA PCW 200 Dataset

Expected/Desirable DDoS Defense			
		Negative (Attack)	Positive (Normal)
DDoS Defense Model Decision	Negative (Attack)	1483	0
	Positive (Normal)	0	1517

VISUALIZATION OF EXPERIMENTAL RESULTS

In this section, we present the visualization charts for our experimental results to show how our DDoS detection solution approach distinctly discriminates the flush crowd (the legitimate access) from the flooding attacks. Moreover, in this section, we show how our detection solution model recognizes and differentiates the normal (legitimate access) from the flooding attacks using the correlational matrix among the selected features, by displaying the density points associated with correlated features that represent both the attacks and the normal access. In both cases, we have used R-programming languages for the computation. Figure 3 and Figure 4 below displays the visual representation of how our proposed detection model recognizes and differentiates the flush crowd from the flooding attacks, by displaying the density points associated with correlated features that represent both normal access and the attacks. The charts are based on computations from the CAIDA PCW200 Dataset.

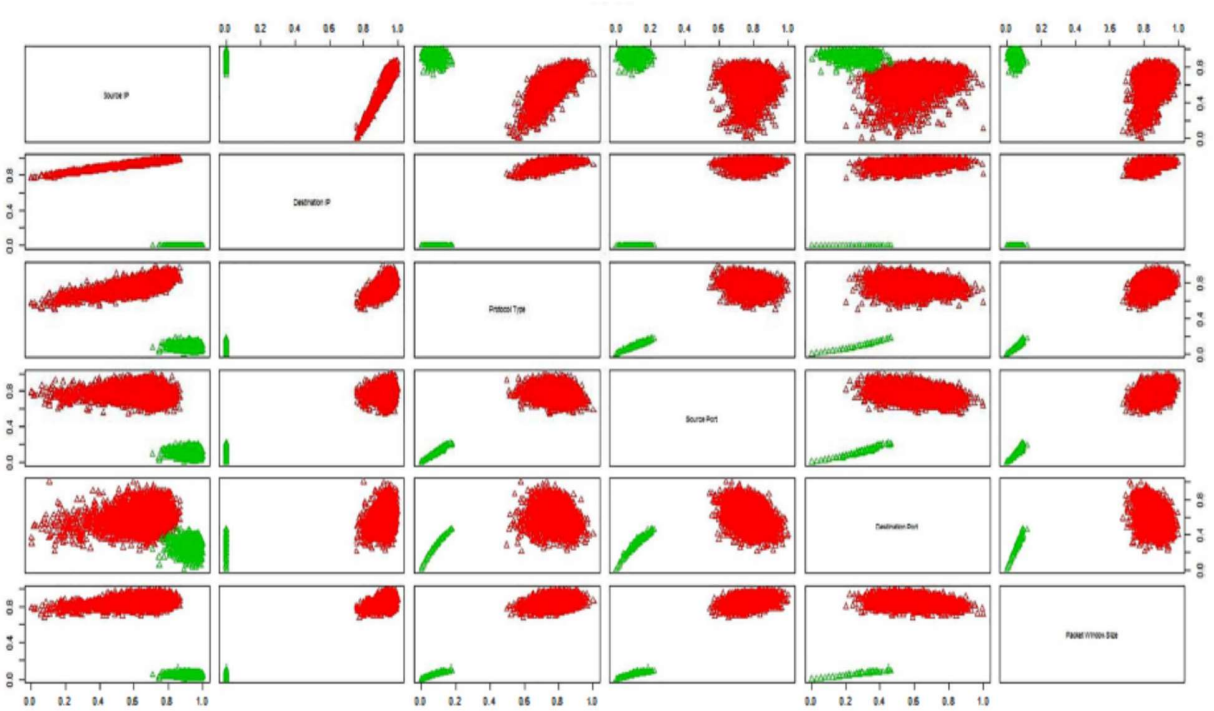

Figure 3. DBSCAN Clustering of Attack and Normal Access Associated with Multivariate Correlated Features

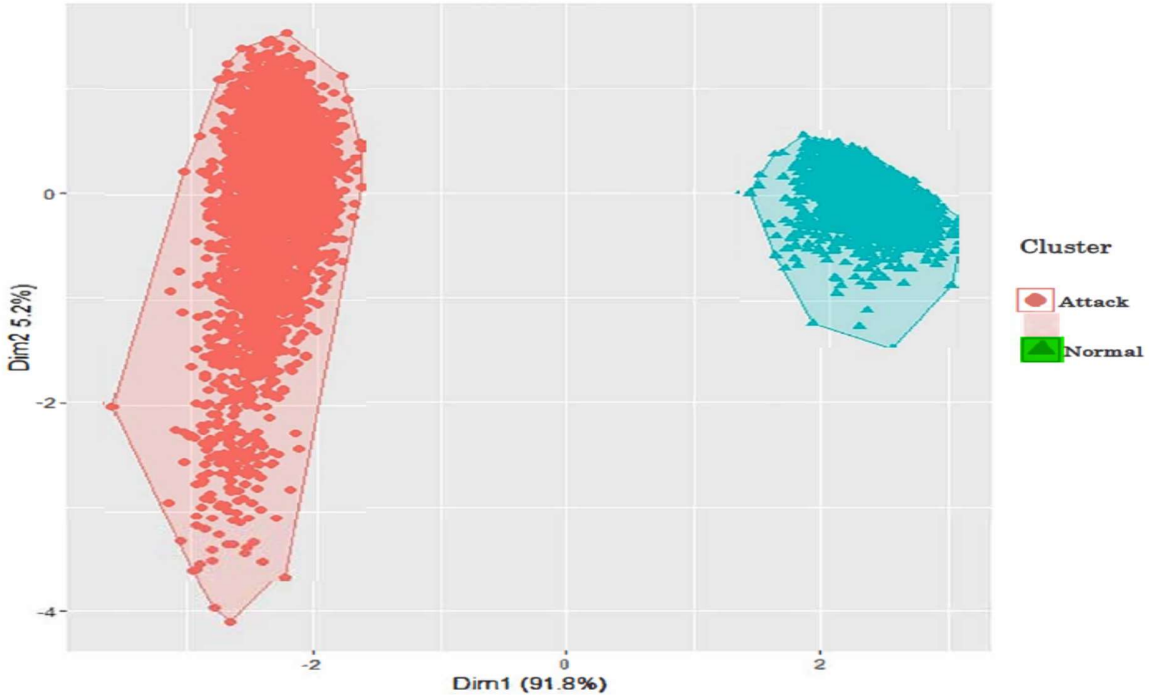

Figure 4. Effective clustering of flooding attack and legitimate access using DBSCAN, Multivariate Correlation, and Entropy

PERFORMANCE EVALUATION REPORT

After completing the experiment of our model against the three datasets, we evaluated the performance of our hybrid model against the key performance evaluation factors listed in Table 7 below. We computed the value of each of these performance factors (Accuracy, Precision, Reliability or False Alarm Rate, False Negative Rate, Sensitivity, and Specificity) for all the 20 experiments we conducted. Figure 5 below shows a graphical representation of the experimental results.

Table 7. Experimental and performance evaluation results

Experimental Data	Accuracy	Sensitivity	False Negative	F1-Score
70% training	99.99%	99.99%	0.01%	99.99%
30% testing	99.98%	99.97%	0.03%	99.98%
60% training	99.99%	99.98%	0.02%	99.99%
40% testing	99.99%	99.99%	0.01%	99.99%

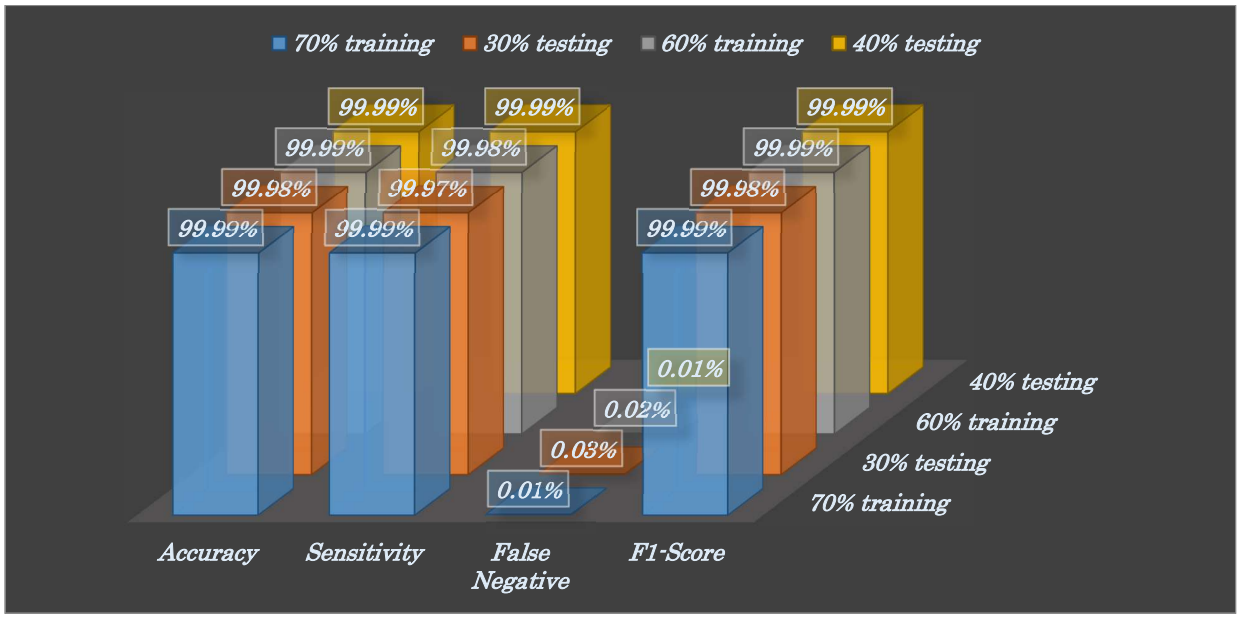

Figure 5. Graphical representation of experimental results

DISCUSSION OF FINDINGS

In our research study for this paper, we used an input file with 2,000,000 packets. We transformed our input data file into three different datasets. These different datasets were generated by applying the entropy data transformation method for 50, 100, and 200 consecutive incoming packets. The three datasets were CAIDA PCW50, CAIDA PCW100, and CAIDA PCW200. These transformations have led to a number of interesting results that support our assumption. The experimental results suggested that our hybrid model is effective and efficient in detecting the DDoS attacks. The summary of our results is presented in Table 8 below.

Table 8. The Effect of Entropy Data Transformation

Data Set Name	*Number of Records for Detection Computation*	*Computed Vector Size*	*Performance Rate*	*Complexity*
CAIDA PCW 50	40,000	5.8GB	Low	High
CAIDA PCW 100	20,000	2.1GB	Medium	Medium
CAIDA PCW 200	10,000	510MB	Very Low	Very Low

DDoS attacks are a main vulnerability for cloud computing providers. The rapid growth of cloud computing services and the exponential growth of the DDoS attacks have made it necessary to develop a comprehensive and effective solution for different kinds of DDoS attacks. Prior research has provided a comprehensive analysis of existing DDoS detection and prevention mechanisms and their limitations (Girma, Garuba, & Goel, 2014; Girma, Garuba, & Li, 2015). The experiment results of this study has demonstrated a promising result in mitigating the risks of DDoS attacks for both network and host based resources.

The results and their statistical significance in our study indicate the effectiveness of our proposed model. One of the achievements in our research besides its high level of accuracy is that the false alarm rate is significantly reduced to zero. Moreover, we have proved how our detection approach effectively and efficiently recognizes and distinguishes the legitimate access from the attacks; and hence we can detect the abnormal behavior of the incoming packets and detect the DDoS attacks at the early stage in the real time. The experimental results show that our proposed hybrid comprehensive approach to mitigate the DDoS flooding attacks is effective in running the entropy and DBSCAN models in parallel so that we could avoid having our cloud service being slowed or shut down by DDoS attacks.

The research in this paper builds upon prior studies on different techniques of intrusion detection systems. Our research contributes the following important findings to the security body of knowledge:

- It raises DDoS attack alarms at the edge of the network and host based levels and eliminates DDoS attack packets before they reach the main cloud resource.
- It discriminates DDoS attacks from the surge of legitimate traffic effectively. And hence, the legitimate traffic would not be denied access to the cloud resource.
- It demonstrates that the high volume of data in a cloud environment can be handled by applying data transformation using entropy.
- It demonstrates that both entropy and normalization for data transformation and DBSCAN for clustering techniques can be combined to detect DDoS attacks effectively.

CONCLUSION

As described above, the correlational studies we have conducted in ranking parametrical features are an important step for our detection approach. One of the major results achieved with our research besides its high level of accuracy is that, the false alarm rate is significantly reduced to zero. Moreover, we have proved how our DDoS detection approach is able to effectively and efficiently recognize and distinguish the legitimate access from the attacks; and hence we can detect the abnormal behavior of the incoming packets and detect the attacks at the early stage in the real time. Our research has also indicated that it is very effective to introduce a hybrid comprehensive approach to mitigate the DDoS flooding attacks by running the entropy and DBSCAN models in parallel so that we could avoid a slowdown or a shutdown for cloud services in the case of DDoS attacks.

Our research has made significant contributions to the detection and mitigation of DDoS attacks. We recommend a continued research in the DBSCAN clustering area to improve its capability of handling big data. It might be possible to approach the big data issue with running DBSCAN clustering in parallel using multi-thread programming to enhance performance rate and data handling. By using this parallel computational approach on

DBSCAN clustering, we could have individual clusters for different types of attacks, which could help improve the overall detection approach.

REFERENCES

Abdul, Y., Aldeen, A.S., Razzaque, M.A., & Saleh, M. (2014). A survey on security issue and its proposed solution in cloud environment. *Proceedings of the 1st International Conference of Recent Trends in Information and Communication Technologies (IRICT 2014)*, 459-470.

Arlia, D., & Coppola, D. (2001). Experiments in parallel clustering with DBSCAN. *Euro-Par 2001, LNCS 2150*, 326-331.

Ashktorab, V., & Taghizadeh, S. R. (2012). Security threats and counter measures in cloud computing. *International Journal of Application or Innovation in Engineering and Management*, *1*(2), 234-245.

Bace, R., & Mell, P. (2001). NIST Special Publication on intrusion Detection Systems. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a393326.pdf

Bhaya, W., & EbadyManaa, M. (2017). DDoS attack detection approach using an efficient cluster analysis in large data scale. *Proceedings of 2017 Annual Conference on New Trends in Information &Communications Technology Applications (NTICT),* 168-173.

Carlin, S., & Curran, K. (2012). Cloud computing technologies. *International Journal of Cloud Computing Technologies*, *1*(2), 59-65.

D'Cruze, H., Wang, P., Sbeit, R.O., & Ray, A. (2018). A software-defined networking (SDN) approach to mitigating DDoS attacks. In S. Latifi (Eds.) Information Technology – New Generations. Advances in Intelligent Systems and Computing, vol 558 (pp.141-145). Springer, Cham.

Erman, J., Arlitt, M., & Mahanti, A. (2006). Traffic classification using clustering algorithms. *Proceedings of the 2006 SIGCOMM workshop on Mining network data,* 281-286. doi: 10.1145/1162678.1162679

Ester, M., Kriegel, H., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD-96 Proceedings,* 226-231.

Fu, X., Hu, S., & Wang, Y. (2014). Research of parallel DBSCAN clustering algorithm based on MapReduce, International *Journal of Database Theory and Application, 7*(3), 41-48.

Girma, A., & Garuba, M., & Goel, R. (2014). Cloud computing vulnerability: DDoS as its main security threat, and analysis of IDS as solution model. *Proceedings of the 11th International Conference on Information Technology: New Generations (ITNG 2014),* 307-312.

Girma, A., Garuba, M., & Goel, R. (2017). Advanced machine language approach to detect DDoS attack using DBSCAN clustering technology with entropy. In: Latifi S. (Ed). Information Technology - New Generations. Advances in Intelligent Systems and Computing, vol 558 (pp.125-131). Springer, Cham.

Girma, A., Garuba, M., & Li, J. (2015). Analysis of security vulnerabilities of cloud computing environment service models and its main characteristics. *Proceedings of the 12th International Conference on Information Technology: New Generations (ITNG 2015)*, 207-211.

Girma, A., Garuba, M., Li, J., Liu, C., & Abayomi, K. (2015). Analysis of DDoS attacks and an introduction of a hybrid statistical model to detect DDoS attacks on cloud computing environment. *Proceedings of the 12th International Conference on Information Technology: New Generations (ITNG 2015),* 212-217.

Gonzalez, N., Miers, C., Redigolo, F., Simplicio, M., Carvalho, T., Naslund, M., & Pourzandi, M. (2012). A quantitative analysis of current security concerns and solutions for cloud computing. *Journal of Cloud Computing*: *Advances, Systems, and Applications, 1*(11). doi: 10.1186/2192-113X-1-11

Katkamwar, N.S., Puranik, A.G., & Deshpuaunde, P. (2012). Securing cloud servers against flooding based DDoS attacks. *International Journal of Application or Innovation, 1*(3), 50-55.

Kisilevich, S., Mansmann, F., & Keim, D. (2010). P-DBSCAN: A density based clustering algorithm for exploration and analysis of attractive areas using collections of geo-tagged photos. *Proceedings of the 1st International Conference and Exhibition on Computing for Geospatial Research & Application, June 21-23, 2010, Washington, D.C.* doi:10.1145/1823854.1823897

Kozushko, H. (2003). Intrusion detection: Host based and network based intrusion detection system. Independent Study. Retrieved from https://pdfs.semanticscholar.org/471b/6047150e82d5b94cbcf1fed36586dcf929c1.pdf

Lonea, A.M., Popesch, D.E., & Tianfield, H. (2013). Detecting DDoS attacks in cloud computing environments. *International Journal of Computer Communication*, *8*(1), 70-78.

Ma, L., Gu, L., Li, B., Qiao, S., & Wang, J. (2014). GDBSCAN: An improved DBSCAN clustering method based on grid. *Advanced Science and Technology Letters, 74*(ASEA 2014), 23-28.

Mumtaz, K., & Duraiswami, K. (2011). An analysis on density based clustering of multi-dimensional spatial data. *Indian Journal of Computer Science Engineering, 1*(1), 8-12.

Rehman, S.U., Asghar, S., Fong, S., & Sarasvady, S. (2014). DBSCAN: Past, present and future. The Fifth International Conference on the Applications of Digital Information and Web Technologies, 232-238. doi:10.1109/ICADIWT.2014.6814687

Roschke, S., Cheng, F., & Meinel, C. (2009). Intrusion detection in the cloud. *Proceedings of the Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing,* 729-734.

Shelke, P.K., Sontakke, S., & Gawande, A.D. (2012). Intrusion detection system for cloud computing. *International Journal of Scientific and Technology Research, 1*(4), 67-71.

Su, Z., Yang, Q., Zhang, H., Xu, X., & Hu, Y. (2001). Correlation-based document clustering using web logs. *Proceedings of the 34th Hawaii International Conference on System Sciences,* 1-7.

Ubhale, P.R., Sahu, A.M., & Raisoni, G.H. (2013). Securing cloud computing environment by means of intrusion detection system. *International Journal of Computer Science and Management Research,* 2(5), 2430-2435.

Vijayalaksmi, S., & Punithavalli, M. (2012). A fast approach to clustering datasets using DBSCAN and pruning algorithms. *International Journal of Computer Applications, 60*(14), 1-7.

Wang, Y., Wang, Y., Wu, D., & Ren, J. (2013). An incremental rapid DBSCAN clustering algorithm for detecting software vulnerability. *Journal of Convergence Information Technology, 5*(1), 82-94.