IJESRT
INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

A SURVEY ON DETECTION OF DENIAL OF SERVICE ATTACK

Swathy Mohan*, Niyas N
*Department of Computer Science and Engineering,KMCT College of Engineering,India

ABSTRACT

Denial-of-Service attack is an attempt to make network resources or machine unavailable to its intended users. It can temporarily or indefinitely interrupt or suspend the services of a host connected to the Internet. For the Detection of denial of service attack (DoS) various detection systems are developed. The existing method is based on machine learning and statistical analysis, the proposed system changes the traffic records as images and detection of DoS attacks as a computer vision problem. In this proposed method a multivariate correlation analysis approach is developed to explain network traffic records and changes the records into the respective images. Network traffic records images are used for the proposed denial of service attack detection system that is developed based on a dissimilarity measure Earth Mover's Distance. By using this method it can also detect the unknown Denial of Service attack.

KEYWORDS:Denial of service attack (DoS), Earth mover's distance, Machine learning, Statistical learning.

INTRODUCTION

Network attack is an attempt to destroy, expose, alter disable, steal or gain unauthorized access. Denial of service attack is active attacks which attack the services of the system also gain the information from the system. Denial of service attacks is used for exploiting system vulnerabilities of a victim or flooding a victim with a large volume of useless network traffic to occupy the designated resources. The attack detection mechanisms can be divided into two namely misuse-based detection and anomaly-based detection. The misuse detection mechanism can achieve high detection rates in known attacks [1] But they are incapable of detecting any unknown attacks or even variants of existing attacks. Anomaly-based detection mechanism uses a different detection methodology that monitors and labels any network activities presenting significant deviation from the legitimate traffic profiles as suspicious objects. Thus anomaly-based detection mechanism is able to identify previously unknown attacks.

In the existing systems it suffers a common issue is that achieving high accuracy in classifying both normal traffic and attack traffic [2]. The most of the systems use only several simple network features of incoming traffic in modeling normal network traffic, and overlook the correlations between the network features. Proposed systems are based on traditional statistical correlation analysis techniques, capable of studying the correlations between the features in a given sample set [3]. The traditional statistical correlation analysis techniques make these anomaly-based detection systems which is incapable of recognizing individual attack records hidden in a sample set. Various classifiers are used to help improve detection accuracy for classifying normal and attack traffic. In computer vision tasks the technique used are the potential candidates. Image retrieval and object shape recognition are some of the commonalities in attack detection and computer vision task. The queries to image retrieval tasks or object shape recognition tasks are equivalent to normal traffic to attack detection. Object shapes or the images that do not match the queries interpreted the detection of DoS attack. Thus for detecting the attacks computer vision techniques provide effective solutions to the problem.

In the paper to avoid the above problems sophisticated anomaly based system was proposed for detecting the denial of service attacks. Various techniques are used such as to find the correlations between the features of network traffic are extracted using our previously developed method Multivariate Correlation Analysis (MCA) technique, which characterizing network traffic. To improve the detection accuracy use the principle of object shape recognition and Earth Mover's Distance (EMD) for testing the traffic records. EMD will find the distance between normal and new records if it does not match the new record become an attack.

To determine or characterize the traffic records first the basic features are generated from the network traffic packets captured at the destination .Then dimensionality Reduction Based on PCA performs dimensionality reduction using PCA for the training normal traffic records. It does not cause loss of information by the use of PCA which seeks the optimal subspace for the best representation of the data. Subspace selected are used for training and the test phase it can reduces the computational overhead for finding the attack and normal traffic records. It consists of a training phase and a test phase, in the training phase it consist of both attack and normal traffic records. In the test phase new records are comparing with the trained records. When compared with the previous method by using this method it can efficiently distinguish both known and unknown DoS attacks.

LITERATURE SURVEY

Denial of service attack is an active attack it is an attempt to make network resources or a machine unavailable to its intended users. Temporarily or indefinitely interrupt or suspend the services of a host connected to the Internet.

Zhiyuan Tan propose [4] a system for Denial-of-Service Attack detection Based on Multivariate Correlation Analysis. In the method Multivariate Correlation Analysis (MCA) is used for the DoS attack detection that it accurately characterizes the network traffic by extracting the geometrical correlations between network traffic features. It employs the principle of anomaly-based detection in attack recognition or detecting the attack traffic. That it makes capable of detecting known and unknown DoS attacks effectively .To speed up the process of MCA a triangle-area-based technique is also proposed. By using this method the system outperforms two other previously developed state-of-the-art approaches .Thus it increases the detection accuracy.

First the basic features are generated from ingress network traffic to the internal network and also used to form traffic records for a well-defined time interval. It can analyze the destination network and it reduces the overhead of detecting the malicious activities. Then the Multivariate Correlation Analysis, in which the "Triangle Area Map Generation" method is used to extract the correlations between two distinct features within each traffic record .The triangle areas stored in Triangle Area Maps (TAMs).. Training phase used for training the records as attack and normal in the test phase based on the trained records it testing the new records.

Yossi Rubner proposes [5] The Earth Mover's Distance as a Metric for Image Retrieval. A metric between two distributions is the earth mover's distance. It can matches perceptual similarity better than other distance used for image retrieval. It also allows the partial matching also it become more robust than histogram matching techniques .In the paper compare retrieval performance of EMD based on the color and texture also compare with the other distance measure techniques. In the one-dimensional distribution the overall brightness content of a gray-scale image can be find by the image intensities, and in the three dimensional distribution it can play a similar role for the color images. The distribution of local signal energy over frequency can be described by the texture content of an image. It can be used for the image retrieval.

In color to compute the earthmover's distance between color Images. Convert the distribution of pixel colors to the CIE-Lab color space which was expressly designed so that short Euclidean distances correlate strongly with human color discrimination performance, albeit for pairs of colors on a neutral background. To compute the earth movers distance between color images use Euclidean distance in color space with underlying ground distance between individual colors. The Euclidean distance correlates strongly with human color discrimination performance. In texture, color is a purely point wise property of images, the texture involves a notion of spatial extent that a single point has no texture. The frequency content of a neighborhood carried the texture information of a point in the image when the texture is defined in the frequency domain.

Jia Xu proposed [6] a method Efficient and Effective Similarity Search over Probabilistic Data based on Earth Mover's Distance the similarity search raises new challenges to traditional relational database also for supporting the manipulation of probabilistic data. Some complicated distance operators have proven their values for better distinguishing ability in the probabilistic domain. Introduce a new database approach to answer range queries and k-nearest neighbor queries on probabilistic data. EMD is one of the most popular distance operators in probabilistic domain. It also is robust to outliers and tiny probability shifting, imp-roving the quality of similarity search on probabilistic histograms. With respect to the number of histogram bins the improvement on search quality pays an expensive cost on computation efficiency due to the cubic complexity of EMD it also relieves the efficiency issue.

Present a general approach to provide a truly scalable and highly concurrent index scheme applicable with mainstream relational databases, such as PostgreSQL and MySQL to overcome the difficulties. A B+ tree is constructed to index pointers to all probabilistic records based on the mapping values for each domain. In this method calculates a pair of lower and upper bound for each domain, guaranteeing that all of the query results. Retrieved records are joined with intersection operator and range queries are thus issued on each B+ tree. In the intersection results Verifications and refinements are then conducted on the candidates remaining.

Haibin Ling proposed[7] EMD-L1: An Efficient and Robust Algorithm for Comparing Histogram-Based Descriptors .EMD-L1, for computing the Earth Mover's Distance (EMD) between a pair of histograms.EMD-L1 is applied for comparing histograms based which is practically impossible with previous method. The various computer vision tasks such as shape matching image retrieval texture analysis, the Histogram-based descriptors are used .The bin-to-bin distance functions, such as Lp distance, Â2 statistics, and KL divergence, are most commonly used for comparing descriptors. The domain of the histograms is already aligned. The cross bin dissimilarity function that addresses the above alignment problem by solving the transportation problem as a special case of linear programming (LP). EMD is useful for more general class of histogram descriptors such as SIFT and shape context beyond the color signature application.

EMD-L1 is tested for two experiments. First it is applied to the inner distance shape context for shape matching. In the second EMD-L1 is applied to SIFT descriptors for feature matching on images. In the Shape matching with shape context EMD-L1 is tested for shape matching by applying to the inner distance shape context. ISDC is an extension of shape context by using the shortest path distance. For image feature matching EMD-L1 used for the interest point matching. It can be done on a set of ten pairs containing synthetic deformation, noise.

Rajesh Batra proposed[8] Feature Comparisons Of 3-D Vector Fields Using Earth Mover’s Distance .In this method describes the three-dimensional vector fields constructed from simple critical points. It defined a distance metric for comparing two-dimensional fields. It can rethink the representation of a critical point signature and the distance measure between the points. Problems such as grid matching and vector alignment which often complicate other comparison techniques are avoided it relies on topologically based information. The feature information is used to represent, and stored for each field, a significant amount of compression occurs for this method.

The areas of the physical sciences including such diverse subjects as climate modeling, dynamical systems, electromagnetism, and fluid mechanics are studying by the vector field. To compare vector fields an efficient technique has not been developed. To solve the problem this method is developed. This techniques lack the quantitative capabilities for automated comparisons. The two dimensional classification techniques are used to extend the classification to three-dimensional critical points. To redefine the EMD metric allowing for a quantitative comparison between 3-D flow fields a complete categorization of 3-D simple critical points is presented is used . This method also explains the demonstrating effectiveness of the technique on a thermal convection model described by the Lorenz equations.

The feature based comparison method to three-dimensional vector fields is extended. If use the property that a 3-D critical point can be decomposed into a set of 2-D critical points with planar phase portraits the extension can be straight-forward .The redefined distance function for EMD remains a metric. The connections between the critical points are not considered in the 2D case. To reduce the number of false positives and to provide a better distance between two fields the connections should be taken into account. To understanding complicated phenomenon such as the Lorenz model demonstrated the usefulness of the method. With the EMD it captured the evolution of the thermal convection. By using this method the system is represented by quantities, fast searches can be easily constructed to locate particular patterns.

Fereshteh Nayyeri proposes [9] Image Matching Using Dimensionally Reduced Embedded Earth Mover's Distance Earth Mover's Distance (EMD) is a distance measure method. EMD requires high execution time so to overcome emd-L1 was proposed but it decreases the performance. So a new method was proposed in this method EMD embedded with three dimensional reduction methods. In this two methods are used first samplings the samples are selected from the records and used for matching. Second sketching based on selected samples are used for distance measure and characterizing the normal and attack traffic. The exact distance between images can be measure by using EMD and by this measure it can retrieve the most similar images from a database, its execution time is problematic, and this similarity measure is very time consuming. EMD to $L1$, was proposed to solve the EMD problem this method maps the image matrix to an $L1$ norm it is less time consuming, it produces distortion. The exact computation may be practically infeasible; in this situation, an approximation solution is helpful to find the exact result with some distortion. Execution time and performance are important factors in image retrieval. To solve these problem this method is introduced sampling, reduces the time but decreases performance. Sketching improve performance by sacrificing the time of execution.

Kristen Grauman proposes [10] Fast Contour Matching Using Approximate Earth Mover's Distance. To computing the exact minimum cost matching it complexity become high, previous algorithms could only run efficiently when using a limited number of features per shape. In this method propose a contour matching algorithm that quickly computes the minimum weight matching between sets of descriptive local features using a recently introduced low-distortion embedding of the Earth Mover's Distance (EMD) .The nearest neighbors in a database of embedded contours are retrieved in sub linear time via approximate nearest neighbors search with Locality-Sensitive Hashing (LSH).

Matching features from one shape to the features of another often reveals how similar the two shapes ,the cost of matching two features may be defined as how dissimilar they are in spatial location, appearance, curvature, or orientation; the minimal weight matching is the correspondence field between the two sets of features that requires the least summed cost The high-dimensional local features taken from shapes in an image database construct a subspace that captures much of the descriptive power of the rich features, yet allows us to represent them compactly. A subspace over the "shape context" feature which consists of local histograms of edge points, and successfully use it within the proposed approximate EMD shape matching method. A new fast contour matching algorithm was proposed that utilizes an approximation to EMD to judge similarity between sets of local shape descriptors. It enables fast shape-based similarity retrieval from large databases, and its run-time is only linearly dependent on the number of feature points used to represent a shape.

C.J.H. Weeïnk proposes [11] DDoS defense mechanisms: a state of the art research. The Distributed Denial-of-Service (DDoS) attack is widely available but lack of effective mechanisms that defend against such attacks in a reasonable amount of time. In this method presents a research that analyzed the security techniques currently available to reduce the result of DDoS attacks. Also a cooperative distributed defense technique with multiple local detection techniques is the most effective by using this method to analyzes the Distributed Denial-of-Service (DDoS) attacks a state of the art research was proposed in this method and defense techniques that can be used to prevent or reduce the consequences attacks. There no real comparison of different DDoS defense mechanisms can be found in existing method. This method provide an analysis of security mechanisms currently available to reduce the result of DDoS attacks and to identify which technique or combination of techniques is the most promising for the future.

Salvatore J. Stolfo proposes [12] Cost-based Modeling for Fraud and Intrusion Detection: Results from the JAM Project. Using the JAM distributed data mining system for the real world problem of fraud detection in financial information systems the results is achieved. Thus gives clear evidence that state-of-the-art commercial fraud detection systems can be substantially improved in stopping losses due to fraud by combining multiple models of fraudulent transaction shared among banks. The traditional statistical metrics use to train and evaluate the performance of learning systems are misleading and perhaps inappropriate for this application. The Cost-based metrics are more relevant in certain domains, and defining such metrics poses significant and interesting research questions both in evaluating systems and alternative models, and in formalizing the problems to which one may wish to apply data mining technologies. In this method demonstrates the techniques developed for fraud detection can be generalized and applied to the important area of Intrusion Detection in networked information systems. The main results of the JAM Project that focused the discussion on cost-sensitive modeling techniques for credit card fraud detection and network intrusion detection. It can be applying data mining techniques to build intrusion detection models. In intrusion detection it can examine the cost factors and cost models in and discussed the challenges in cost-sensitive modeling for intrusion detection.

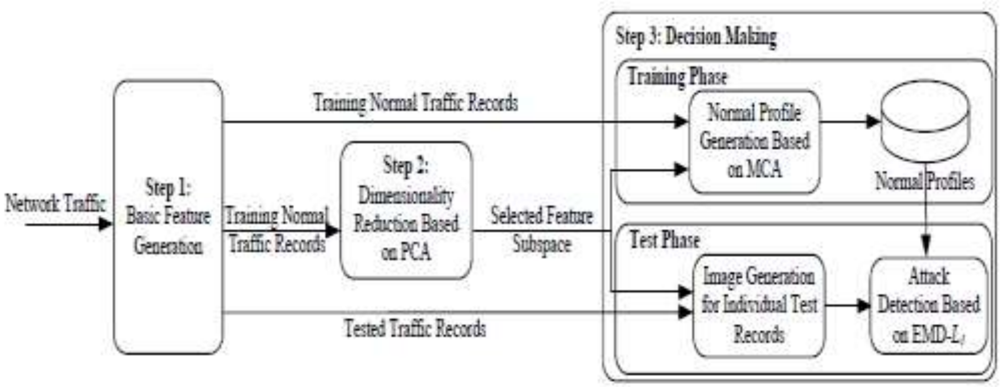
Marina Thottan proposes [13] Anomaly Detection in IP Networks .The network anomaly detection is a vibrant research area. Various techniques such as artificial intelligence, machine learning, and state machine modeling are used. In the method, first review these anomaly detection methods and then describe in detail a statistical signal processing technique based on abrupt change detection. The signal processing technique is effective at detecting several network anomalies. In the method to show the potential to apply signal processing techniques to the problem of network anomaly detection. It will provide better insight for improving existing detection tools as well as provide benchmarks to the detection schemes employed by these tools.

In this method describe the problem of IP network anomaly detection in a single administrative domain along with the types and sources of data available for analysis. To motivating the need for signal processing techniques to study this problem a special emphasis is placed. Also describe areas in which advances in signal detection theory have been useful. A technique based on abrupt change detection for addressing the challenge in this method. It also provides a review of the area of network anomaly detection. It is clear that there is a significant advantage in using the wide array of signal processing methods to solve the problem of anomaly detection. The networking and signal processing areas will help develop better and more effective tools for detecting network anomalies and performance problems.

The proposed method can accurately differentiate the attack and normal traffic. It also supports partial matching. The time complexity also be less when compared to the other method

RESULTS AND DISCUSSION

The sophisticated anomaly based system was proposed for detecting the denial of service attacks. Various techniques are used such as to find the correlations between the features of network traffic are extracted using our previously developed method Multivariate Correlation Analysis (MCA) technique, which characterizing network traffic. To improve the detection accuracy use the principle of object shape recognition and Earth Mover's Distance (EMD) for testing the traffic records. To find the distance between normal and new records if it does not match the new record become an attack the distance measure method EMD is used.


Step 3: Decision Making
Training Phase
Training Normal Traffic Records
Normal Profile Generation Based on MCA
Normal Profiles
Network Traffic
Step 1: Basic Feature Generation
Training Normal Traffic Records
Step 2: Dimensionality Reduction Based on PCA
Selected Feature Subspace
Test Phase
Image Generation for Individual Test Records
Attack Detection Based on EMD-L_1
Tested Traffic Records

Detection of denial of service attack system

To characterize the traffic records first the basic features are generated from the network traffic packets captured at the destination .Then it is converted respective images. Then dimensionality Reduction Based on PCA performs dimensionality reduction using PCA for the training normal traffic records. It does not cause loss of information by the use of PCA which seeks the optimal subspace for the best representation of the data. Subspace selected are used for training and the test phase it can reduces the computational overhead for finding the attack and normal traffic records. In the training phase it consist of both attack and normal traffic records it can be trained using multivariate correlation analysis. In the test phase new records are comparing with the trained records. It can be done by the distance measure method earth mover's distance. When compared with the previous method by using this method it can efficiently distinguish both known and unknown DoS attacks.

CONCLUSION

By using the proposed DoS attack detection system which is equipped with our previously developed MCA technique and the EMD-L1 is efficiently detect the attack. The technique helps to extract the correlations between individual pairs of two distinct features within each network traffic record and offers more accurate characterization for network traffic behaviors. Then it effectively distinguishes both known and unknown DoS attacks from legitimate network traffic. It also allows partial matching. In this method it can also identify the unknown attack and variant of existing attack. It requires only less time when compared with the other methods.

ACKNOWLEDGEMENTS

This work was under the guidance of Mary Linda P A Hod of CSE and my project guide Niyas N assistant professor of CSE at KMCT college of engineering. I gratefully thankful to them for their support and suggestions.

REFERENCES

[1] M. Bando, N. S. Artan, and H. J. Chao, "Scalable Lookahead Regular Expression Detection System for Deep Packet Inspection," Networking, IEEE/ACM Transactions on, vol. 20, no. 3, pp. 699-714, 2012.

[2] M. Thottan, and C. Ji, "Anomaly detection in IP networks," Signal Processing, IEEE Transactions on, vol. 51, no. 8, pp. 2191-2204.

[3] A. Patcha and J. M. Park, "An Overview of the Anomaly Detection Techniques: Existing Solutions and Latest Technological Trends ,"Computer Networks, vol. 51, pp. 3448-3470, 2007.

[4] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R. Liu, "A System for Denial of Service Attack Detection Based on Multivariate Correlation Analysis"Parallel and Distributed Systems, IEEE Transactions on, vol. 25, no. 2, pp.447-456, 2014.

[5] Y. Rubner, C. Tomasi, and L. Guibas, "The Earth Mover's Distance as a Metric for the Image Retrieval," International Journal of Computer Vision, vol. 40, no. 2, pp. 99-121, 2000/11/01, 2000.

[6] J.Xu, Z. Zhang, A. K. H. Tung, and G. Yu, "Efficient and effective of similarity search over probabilistic data based Earthmover's distance " The VLDB Journal, vol. 21, no. 4, pp. 535–559, 2012.

[7] H. Ling, and K. Okada, "An Efficient Earth Mover's Distance Algorithm For the Robust Histogram Comparison," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 29, no. 5, pp. 840-853, 2007.

[8] R. Batra and L. Hesselink, "Feature comparisons of the 3-D vector fields using Earth mover's distance," in Proceedings of the IEEE Visualization '99, pp. 105–114, IEEE Computer Society Press, October 1999.

[9] Hindawi Publishing Corporation" Image Matching Using Dimensionally Reduced Embedded Earth Mover's Distance "Journal of Applied Mathematics Volume 2013, Article ID 749429, 11 pages .

[10] K. Grauman and T. Darrell, "Fast contour matching using approximate Earth mover's distance," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04), vol. 1, pp. I220–I227, July 2004.

[11] C. Douligeris, and A. Mitrokotsa, "DDoS attacks and defense mechanisms: classification and state-of-the-art," Computer Networks, vol. 44, no. 5, pp. 643-666, 2004.

[12] S. J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. K. Chan, "Costbased modeling for fraud and intrusion detection: results from JAM project"in DARPA Information Survivability Conference and Exposition2000 DISCEX '00. Proceedings, 2000, pp. 130-144 vol.2.

[13] M. Thottan, and C. Ji, “Anomaly detection in the IP networks,” Signal Processing, IEEE Transactions on, vol. 51, no. 8, pp. 2191-2204.

AUTHOR BIBLIOGRAPHY

	Swathy Mohan **Swathy Mohan** completed her bachelor of engineering in computer science and engineering in 2014, from Paavai College of engineering Salem. She currently doing master degree in computer science and engineering in Kmct college of engineering Calicut university.
	Niyas N He completed master degree in computer science and engineering from KMCT College of Engineering in year 2014.Now he is an Assistant Professor, Department of Computer Science and Engineering, in KMCT College of Engineering, Calicut University.