ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Assessing Categorical Correlation of Network Features to Scale the Scope of Denial of Service Attacks

G.S.Shabbir Ahmmed
M.Tech (CSE)
Kottam College of Engineering
Chinnatekuru (V), Kallur (M), Kurnool District-518218

A.V.Ramakrishna Reddy
Assistant Professor
Kottam College of Engineering
Chinnatekuru (V), Kallur (M), Kurnool District-518218

Abstract:

DDOS attack recognition dependent on defects is one concerning considerable DDOS attack recognition techniques. Phenomenal progress in the amount of computer network users prospects to the significant divergence of position activities. Henceforth we posses to choose the noticeable as sociability among functions of the network and such of each network transfer. In respect to this perspective, here in this paper, we own constructed a mathematical scaling procedure to approximate if a network transaction is safe, questionable or DDOS attack. The recommended model is utilizing bipartite graph strategy to approximate the powerful associability of the attributes. The associability of the specifications is basically symbolized by the associability of that attributes specific values. The outcomes explained from the empiric study identify that the recommended product is effectively offering the consistency regarding determining the uncomfortable state of a network exchange.

Index Terms: ***Denial-of-service attack, network traffic characterization, multivariate correlations, triangle area***

Introduction

In the previous few decades Internet has endured an sudden growth. Together with the spacious expansion of new services, the amount and effects of attacks have become frequently increasing. The amount of computer systems and their exposure has been increasing, while the standard of sophistication and understanding needed to possess out an attack have become decreasing, as massive technical approach know-how is commonly presented in Web sites all more than the world.

Current advances in encoding, public key return, digital signature, and the building of associated standards have specified a basis for network safety. Nevertheless, security on a network proceeds more than these concerns. Obviously it must comprise protection of computer techniques and networks, at all stages, top to bottom.

Considering it appears difficult to assurance finalize protection to a system by indicates of prevention components (e.g. verification techniques), the utilize of an Intrusion Detection System (IDS) is of biggest significance to present intrusions in a network or in a system. IDSs are commonly categorized on the foundation of numerous criteria [1].

Express of the art in the discipline of intrusion recognition is mostly symbolized by pervert established IDSs. Thinking about that the majority attacks are recognized with acknowledged tools, obtainable on the Internet, an individual based IDS might seem a good solution.

However hackers frequently arrive up with new strategies for the attacks, that a abuse based IDS is not qualified to block. This is the primary reason why our efforts have concentrated on the development of an position dependent IDS. In specific our goal is to expose intrusions offered out using TCP bugs, by utilizing statistical model to identify the tendencies of

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

network traffic. The use of analytical techniques is a well recognized strategy to identify two exclusive kinds of "anomalies": masqueraders (evaluating the demand stream of a host) and intruders (evaluating the progression of TCP moves in the network traffic) [2].

INTRUSION DETECTION BY FEATURE ASSOCIATION

The strategy of PDDOS statistic projected in this paper is primarily accepts the information of the provided training set and offer specific values utilized in those information as two private sets and additional builds a bipartite graph among these two.

Assumptions**:**

Let set of features

$$\{f1, f2, f3, \ldots\ldots, fn \forall f_i = \{f_i v_1, f_i v_2, \ldots. f_i v_m\}\}$$

Which are providing categorical principles and worn to form the T

Here T is set of network transaction proceedings of the specified training set such that

$$T = \{t_1, t_2, t_3, \ldots\ldots t_n \forall t_i = \{val(f_1), val(f_2), \ldots. val(f_i), val(f_{i+1}), \ldots. val(f_n)\}\}$$

The position of categorical principles of features go each network transaction will be measured as transaction value set tvs, and each and every one transaction assessment sets are referred as '$STVS$'.

Here in above explanation $val(f_i)$ can be distinct as $val(f_i) \in \{f_i v_1, f_i v_2 \ldots. f_i v_m\}$

Here subsequent to the term feature refers the present categorical value of the attribute

Let two features '$val(f_i)$' and '$val(f_j)$', '$val(f_i)$' connected with '$val(f_j)$' if and only if $(val(f_i), val(f_j)) \in tvs_k$.

Construct a weighted graph WG with standards of features as vertices and edges connecting values of features. An edge among any two features $val(f_1), val(f_2)$ will be weighted as follows

$$ctvs = 0;$$
$$foreach\ \{tvs \forall tvs \in STVS\}$$
$$ctvs += \{1 \forall (val(f_1), val(f_2)) \subseteq tvs\}$$

Here in the exceeding equation $ctvs$ indicates the calculate of transactions, which surround both features $val(f_1), val(f_2)$. Then the edge weight between features $val(f_1)\ and\ val(f_2)$ can be considered as follows.

$$w(val(f_1) \leftrightarrow val(f_2)) = \frac{ctvs}{|STVS|}$$

In the procedure of construction a weighted graph we believe that an edge connecting any two features subsist if and only if $ctvs \geq 1$

Process

In consider exploring the procedure by an instance, let believe the total number of divergent values of features as 8 that symbolize as a set $V = \{val_1, val_2, \ldots. val_8\}$ and $|T|$ as 6, Here $|T|$ is size of the network transaction records

Table 1 binary illustration of the association connecting T and V

	val_1	val_2	val_3	val_4	val_5	val_6	val_7	val_8	
tvs_1	1	0	0	0	0	1	0	1	(val_1, val_6, val_8)
tvs_2	0	1	0	0	1	1	0	1	$(val_2, val_5, val_6, val_8)$
tvs_3	1	1	1	0	0	0	1	0	$(val_1, val_2, val_3, val_7)$
tvs_4	0	0	0	0	0	0	1	0	(val_7)
$tvs5$	0	0	0	1	0	1	1	1	$(val_4, val_6, val_7, val_8)$
$tvs6$	1	1	1	1	0	0	1	0	$(val_1, val_2, val_3, val_4, val_7)$

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

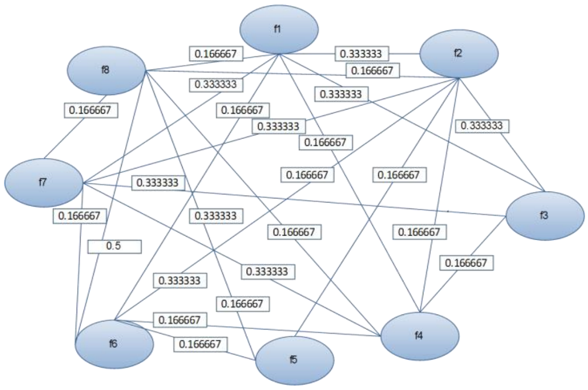

Fig1: An illustration weighted graph of categorical values set of count 8.

Here in above table1 and Figure1 each element $\{val_1, val_2, ...val_8\}$ can be $f_i v_j$ such that $\{f_i v_j \exists i \in [1,2,......n] \wedge j \in [1,2,.....m]\}$

In the procedure of detecting the alliance of each feature categorical value $f_i v_j$ referred as val_k with network transaction records, originally we build a bipartite graph among transaction value sets $STVS$ and the attribute categorical values V .

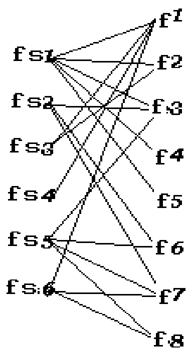

Fig 2: bipartite graph between STVS and V

If a attribute categorical value $f_i v_j$ that referred as val_1 exists in tvs_1 then the weight of the connection connecting val_1 and tvs_1 will be the sum of the weights of the edges connecting val_1 and each feature categorical value $\{f_i v_j \exists f_i v_j \in tvs_1\}$ of tvs_1 that defined in weighted graph WG .

Table 2: matrix A as follows that symbolize the connection weights connecting a attribute categorical value and each transaction value set

	val_1	val_2	val_3	val_4	val_5	val_6	val_7	val_8
tvs	0.879761	0.593871	0.719431	0.888985	0.009925	0.951906	0.596128	0.140759
tvs	0.487673	0.175749	0.473697	0.098931	0.672281	0.65032	0.688884	0.102504
tvs	0.334752	0.194757	0.436627	0.005882	0.196037	0.337898	0.703112	0.58396
tvs	0.737372	0.696016	0.147575	0.993947	0.865990	0.515846	0.629118	0.014945
tvs5	0.071456	0.59152	0.712407	0.471292	0.328539	0.693965	0.889919	0.434219
tvs6	0.00046	0.767607	0.056246	0.148297	0.924187	0.313737	0.506235	0.993027

Table 3: Transpose matrix A' of matrix A as fallows that represents the connection connecting a transaction and each transaction level feature set fs .

	tvs_1	tvs_2	tvs_3	tvs_4	tvs5	tvs6
val_1	0.879761	0.487673	0.334752	0.737372	0.071456	0.00046
val_2	0.593871	0.175749	0.194757	0.696011	0.59152	0.767607
val_3	0.719431	0.473697	0.436627	0.147578	0.712407	0.056246
val_4	0.888985	0.098931	0.005882	0.993941	0.471292	0.148297
val_5	0.009925	0.672281	0.196037	0.865996	0.328539	0.924187
val_6	0.951906	0.65032	0.337898	0.515843	0.693965	0.313737
val_7	0.596128	0.688884	0.703112	0.629113	0.889919	0.506235
val_8	0.140759	0.102504	0.58396	0.01492	0.434219	0.993027

Let consider $STVS$ as a database and depict it as a bipartite graph without loss of information. Let $STVS = \{tvs_1, tvs_2, ..., tvs_6\}$ be a list of network transactions with feature categorical values and $V = \{val_1, val_2, ...val_8\}$ be the corresponding set of

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

feature correlation values. Then, clearly $STVS$ is equivalent to the duplex-graph $DG=(STVS,V,E)$ Here $E=\{(tvs_i,val_j):val_j \in tvs_i, tvs_i \in STVS, val_j \in V\}$.

The bipartite graph demonstration of the position of transaction value sets $SCFS$ is inspiring. It offers us the idea of creating link-based standing models for the assessment of connected sets. In this bipartite graph, the connections maintain of a transaction c is correspondent to degree of all its attributes weight. Although, it is crucial to posses assorted closeness weights for distinctive transaction benefits sets in order to reflect their assorted importance. The assessment of impact connected sets must be calculated from these weights. Following comes the question of how to obtain weights in a set of deal value sets. Naturally, a transaction level showcase set with high distance weights should possess many of the functions those belongs to the same dealing with high connections support; at the same time, a transaction with high connections support must be secured by less or zero other transaction appreciate sets high closeness weights. The reinforcing relationship of transaction appreciates sets and transactions are simply like the connection between hubs and authorities in the bipartite graph.

Additional assuming transaction value sets as untainted hubs and the feature categorical standards as pure authorities, the hub and authority principles can be calculated as follows:
Let matrix illustration of transaction value sets and mark connections as a matrix 'A'(see table 3). The value represents that a feature associated how many attribute categorical values of the same transaction

If a feature f_1 survive in feature set fs_1 then the weight of the connection between f_1 and fs_1 will be the sum of the weights of the edges connecting $f1$ and each feature of $fs1$ that distinct in weighted graph WG

Think the matrix u that representing each hub initial value as 1.
Initially consider the each recorded weights as 1 by default as fallow and represent them as matrix u.

1
1
1
1
1
1

Transpose the matrix A as A'(see table 4)

Find Feature weights by multiplying A' with u as $v=A'\times u$ (Matrix multiplication between A' and u gives a matrix v that represents the authority weights)
Now find the original recorded weights through matrix multiplication between $A and v$.
$u=A\times v$
Then the $Pddos$ of feature association value $f_i v_j$ can be measured as follows

$$Pddos(f_i v_j)=\frac{\sum_{k=1}^{|STVS|}\{u(tvs_k):(f_i v_j \rightarrow tvs_k)\neq 0\}}{\sum_{k=1}^{|STVS|}u(tvs_k)}$$

Then the $Pddos$ between feature association values $f_i v_j$ and $f_{i'} v_{j'}$ can be measured as follows

$$pddos(f_i v_j \leftrightarrow f_{i'} v_{j'})=\frac{\sum_{k=1}^{|STVS|}\{u(tvs_k)\exists(f_i v_j, f_{i'} v_{j'})\subset tvs_k\}}{\sum_{k=1}^{|STVS|}u(tvs_k)}$$

Here in the above equation descriptions, the $|STVS|$ represents total number of transaction value sets.

Further the $Pddos$ of the each transaction value set tvs_i can be measured as follows

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

$$pddos(tvs_i) = 1 - \frac{\sum_{j=1}^{m} \{pddos(\{val_j \exists val_j \in V\}) : (val_j \subset tvs_i)\}}{|tvs_i|}$$

$$pddost = \frac{\sum_{i=1}^{|STVS|} pddos(tvs_i)}{|STVS|}$$

Here in the above equation $|STVS|$ indicates the total number of transaction value sets

The standard deviation of the *Pddos* of each transaction value set needs to be measured further, which is in regard to estimate the low, medium and high ranges of *pddost* . The exploration of mathematical notation of estimating standard deviation follows

$$sdv_{pddost} = \sqrt{\frac{\left(\sum_{i=1}^{|STVS|} (pddos(tvs_i) - pddost)^2\right)}{(|STVS|-1)}}$$

The Feature Association Impact Scale range can be explored as follows

Lower threshold of *pddost* range is

$pddost_l = pddost - sdv_{pddost}$

Higher threshold of *ddpt* range is

$pddost_h = pddost + sdv_{pddost}$

A network Transaction nt can be said as safe if and only if $pddos(nt) < pddost_l$

A Network Transaction nt can be said as suspected to be an intrusion if and only if $pddos(nt) \geq pddost_l \;\&\&\; pddos(nt) < pddost_h$

The Network Transaction nt can be confirmed as intrusion if $pddos(nt) \geq pddost_h$

PRAGMATIC ANALYSIS OF THE PROPOSED MODEL:

We considered the reliability of the projected system on prepared network transactions dataset of NSL-KDD [17]. The preceding said data set possesses 125973 selections as preparing set, and 22544 selections are obtainable as test set. The working out set is used to calculate the showcase relationship affect scale threshold and its lower, medium and upper values. The test set is utilized to forecast the scalability of the projected model. Curiously, the scientific study provided promising results. The reports explained in table 2

Table 2: Statistics of the experiment results

Total Number of Records	148517
Total number fields in a record	41
Total number of feature categorical values found	18370
Total number of edges determined	146960
Feature Association Impact Scale threshold found:	0.802100787
Feature Association Impact Scale threshold Upper Bound	0.862553922
Feature Association Impact Scale threshold Lower Bound	0.741647652

Total records Tested 22544

Total number of records found with ‘ *fais* ’ less than lower bound are 3502 (out of this false negatives are 1288)

Total number of records found with ‘ *fais* ’ greater than lower bound are 21042 (true positives are 18692 and 2350 records are false positives)

As per the results explore in table 2 and 3, the projected model is perfect to the level of 92.73%. The failure percentage is 7.26%, which is supposed and occurred due to categorical principles of the features.

Performance Analysis

We used interruption detection correctness (the portion of appropriate forecasts by the recommended) as the primary efficiency measure. In acquisition to calculating precision, the precision, recall, and F-

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

measure were utilized to measure the efficiency; these are characterized using appropriate equations.

$$pr = \frac{t_+}{t_+ + f_+}$$

Here in above Equation the pr indicates the precision, t_+ indicates the true positives and f_+ indicates the false positive

$$rc = \frac{t_+}{t_+ + f_-}$$

Here in exceeding Equation, the 'rc' indicates the recall, 'f_-' indicates the false negative.

$$F = \frac{2 * pr * rc}{pr + rc}$$

Here in the above Equation, 'F' indicates the F-measure.

Table 3: Precision, recall and F-measure values found from the results of the empirical analysis.

precision	recall	f-measure
0.888336	0.9355634	0.9113436

CONCLUSION:

A unique statistical strategy regarding anomaly based intrusion detection is projected in this paper. The endeavours to determine a proportion that assessments the affect of a network transaction if it is protected, suspicious or entrance is first in best of our information. The empirical results acquired from scientific study performed on NSL-KDD dataset is excellent and stimulating our analysis further. In upcoming a novel future connection evaluation strategy can be required that might lead to eliminate the deemed feature set and procedure difficulty, and also might stimulate the reliability towards intrusion detection scope.

References:

[1] Kemmerer, R.A., Vigna, G., Intrusion Detection: a Brief History and Overview, IEEE Security and Privacy (supplement to Computer, vol. 35, no. 4) pp 27-30, April 2002

[2] Ye, N., Yebin Zhang, Y., and Borror, C.M., Robustness of the MarkovChain for Cyber-Attack Detection, IEEE Transactions on Reliability, Vol. 53, no. 1, pp. 116-123, March 2004

[3] D. Anderson, T. Frivold, and A. Valdes,Next-generation intrusion detection expert system (NIDES): A summary, SRI International, Computer Science Laboratory, 1995.

[4] D. Anderson, T.F. Lunt, H. Javitz, A. Tamaru, and A. Valdes,Detecting unusual program behavior using the statistical component of the Next-generation Intrusion Detection Expert System (NIDES), SRI International, Computer Science Laboratory, 1995.

[5] S. Axelsson,Intrusion detection systems: A survey and taxonomy, Tech. Report 99-15, Chalmers University of Technology, Department of Computer Engineering, 2000.

[6] P. Barford and D. Plonka,Characteristics of network traffic flow anomalies, Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement, ACM New York, NY, USA, 2001, pp. 69–73.

[7] Paul Barford, Jeffery Kline, David Plonka, and Amos Ron,A signal analysis of network traffic anomalies, Proceedings of the second ACM SIGCOMM Workshop on Internet measurment (Marseille, France), SIGCOMM: ACM Special Interest Group on Data Communication, ACM Press New York, NY, USA, 2002, pp. 71–82.

[8] M. Dacier H. Debar and A. Wespi,A revised taxonomy for intrusion-detection systems, Tech. report, IBM Research Report, 1999.

[9] A. Jones and R. Sielken,Computer system intrusion detection: A survey, Tech. report, Department of

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Computer Science, University of Virginia, Thornton Hall, Charlottesville, VA, September 2000.

[10] Christopher Kruegel and Giovanni Vigna,Anomaly detection of web-based attacks, Proceedings of the 10th ACM conference on Computer and communication security (Washington D.C., USA), ACM Press, October 2003, pp. 251–261.

[11] S. Kumar,Classification and detection of computer intrusions, Ph.D. thesis, Purdue University, 1995.

[12] T. Lane,Machine learning techniques for the computer security domain of anomaly detection, Ph.D. thesis, Purdue University, August 2000.

[13] J. Lee, S. Moskovics, and L. Silacci,A Survey of Intrusion Detection Analysis Methods, 1999.

[14] Teresa F. Lunt,Detecting intruders in computer systems, Proceedings of the 1993 Conference on Auditing and Computer Technology, 1993.

[15] J. McHugh,Intrusion and intrusion detection, International Journal of Information Security1 (2001), no. 1, 14–35.

[16] P. G. Neumann and A. Ph. Porras,Experience with emerald to date, Proceedings of First USENIX Workshop on Intrusion Detection and Network Monitoring (Santa Clara, California), IEEE Computer Society Press, April 1999, pp. 73–80.

[17] A. Ph. Porras and P. G. Neumann,Emerald: Event monitoring enabling responses to anomalous live disturbances, Proceedings of the National Information Systems Security Conference, 1997, pp. 353–365.

[18] V.A. Siris and F. Papagalou,Application of anomaly detection algorithms for detecting SYN flooding attacks, Computer Communications29(2006), no. 9, 1433–1442.

[19] S.E. Smaha,Haystack: An intrusion detection system, Aerospace Computer Security Applications Conference, 1988., Fourth, 1988, pp. 37–44.

[20] A. Sundaram,An introduction to intrusion detection, Crossroads2(1996), no. 4, 3–7.

[21] Marina Thottan and Chuanyi Ji,Anomaly detection in ip networks, IEEE Transactions on Signal Processing51(2003), no. 8, 148–166.

[22] J. Kim. “An Artificial Immune System for Network Intrusion Detection.” http://www.cs.ucl.ac.uk/staff/J.Kim/GECCO_WS99.html

[23] M. Craymer, J. Cannady, J. Harrell. “New Methods of Intrusion Detection using Control-Loop Measurement.” In: Fourth Technology for Information Security Conference’96. May, 16, 1996.

[24] W. Lee, S. Stolfo. “Data Mining Approaches for Intrusion Detection.” In: Proceedings of the 7th USENIX Security Symposium. 1998.

[25] M. Prabhaker. “Intrusion Detection.” http://www.cs.wright.edu/~pmateti/Courses/499/IntrusionDetection/

[26] M. Gerken. “Rule-Based Intrusion Detection.” http://www.sei.cmu.edu/str/descriptions/rbid_body.html

[27] http://www.nfr.com/products/NID/

[28] http://www.checkpoint.com/products/firewall-1/realsecure.html